Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits

Jacqueline M. Lane[1, 2, 3], Jingjing Liang[4], Irma Vlasac[1,3], Simon G. Anderson[5,6], David A. Bechtold[7,8], Jack Bowden[9,10], Richard Emsley[11], Shubhroz Gill[3], Max A. Little[12,13], AnneMarie I. Luik [14], Andrew Loudon[7,8], Frank A.J.L. Scheer[15], Shaun M. Purcell[3,16,17], Simon D. Kyle[14], Deborah A. Lawlor[9,10], Xiaofeng Zhu[4], Susan Redline[18], David W. Ray[7,8], Martin K. Rutter [7,8,19*], Richa Saxena[1,2,3,15*]

*Equal Contribution

[1]Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA

[2]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[3]Broad Institute, Cambridge, MA, USA

[4]Department of Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA.

[5]Division of Cardiovascular Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK

[6]The George Institute for Global Health, University of Oxford, Oxford Martin School, 34 Broad Street Oxford OX1 3BD, UK

[7]Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[8]Division of Endocrinology, Diabetes & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, UK

[9]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

[10]School of Social and Community Medicine, University of Bristol, Bristol, UK

[11]Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

[12]Engineering and Applied Science, Aston University, Birmingham, UK

[13]Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

[14]Sleep and Circadian Neuroscience Institute, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[15]Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, and Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA

[16]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

[17]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.

[18]Departments of Medicine, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston

[19]Manchester Diabetes Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

Corresponding author

R. Saxena, Center for Human Genetic Research, Massachusetts General Hospital,

185 Cambridge Street, CPZN 5.806,

Boston, MA, 02114, USA

E-mail: rsaxena@partners.org

Phone: 617-643-8578

Fax: 617-643-3203

Word count: 1923

Table count: 2

Figure count: 4

Supp count: Tables 12/ Figures 8

Chronic sleep disturbances, associated with cardio-metabolic diseases, psychiatric disorders and all-cause mortality[1,2], affect 25-30% of adults worldwide[3]. While environmental factors contribute importantly to self-reported habitual sleep duration and disruption, these traits are heritable[4-9], and gene identification should improve our understanding of sleep function, mechanisms linking sleep to disease, and development of novel therapies. We report single and multi-trait genome-wide association analyses (GWAS) of self-reported sleep duration, insomnia symptoms including difficulty initiating and/or maintaining sleep, and excessive daytime sleepiness in the UK Biobank (n=112,586), with discovery of loci for insomnia symptoms (near *MEIS1, TMEM132E, CYCL1, TGFBI* in females and *WDR27* in males), excessive daytime sleepiness (near *AR/OPHN1*) and a composite sleep trait (near *INADL* and *HCRTR2*), as well as replication of a locus for sleep duration (at *PAX-8*). Genetic correlation was observed between longer sleep duration and schizophrenia (r_G=0.29, p=1.90x10^{-13}) and between increased excessive daytime sleepiness and increased adiposity traits (BMI r_G=0.20, p=3.12x10^{-09}; waist circumference r_G=0.20, p=2.12x10^{-07}).

Rather than being 'secondary', evidence suggests disordered sleep may play an important role in the etiology and maintenance of physical and mental health[1,2]. Heritability has been estimated at ~40% for sleep duration[4,6-8], 25-45% for insomnia[9] and 17% for excessive daytime sleepiness[9], but few genetic factors are known. A Mendelian short sleep mutation in *BHLHE41* (P385R) has been identified, and confirmed in mouse models[10]. GWAS for sleep duration have been reported[11-14], but only an association at the *PAX8* locus reached genome-wide significance and was confirmed across ethnic groups[12]. There are several reported loci for restless legs syndrome (RLS) and narcolepsy, but no known robust genetic loci for insomnia symptoms or excessive daytime sleepiness[15,16].

We and others performed a GWAS for chronotype in the UK Biobank[17,18] and a 23&me participant sample[19]. To identify genetic variants that contribute to self-reported sleep duration, insomnia symptoms, and excessive daytime sleepiness and link them with other conditions, we performed GWAS using phenotype measures in UK Biobank participants of European ancestry. Variation in sleep duration, insomnia symptoms and excessive daytime sleepiness was associated significantly with age, sex, principal components of ancestry (PCs), genotyping array, depression, psychiatric medication use, self-reported sleep apnea, and BMI (**Supplementary Table 1)**, as previously reported[20-23]. Together age, sex, and PCs explained 0.4%, 3.0% and 1.3% of variation in sleep duration, insomnia symptoms, and excessive daytime sleepiness respectively. Strong and significant pair-wise phenotypic correlation was seen between the traits overall and within each sex, with limited correlation observed with chronotype. (**Fig. 1a; Supplementary Fig. 1**).

GWAS analyses of sleep duration, insomnia symptoms and excessive daytime sleepiness were performed using linear/logistic regression adjusting for age, sex, 10 PCs and genotyping array. Nine genome-wide significant ($p<5x10^{-8}$) and 14 suggestive ($p<5x10^{-7}$ to $p=5x10^{-8}$) loci were identified (**Fig. 2, Table 1, Supplementary Figs. 2 and 3**). For sleep duration (n=111,975), the strongest association was observed at the *PAX-8* locus (rs62158211T, β(se)=2.34(0.30) mins/allele, $p=4.7x10^{-14}$, effect allele frequency (EAF) 0.213, **Fig. 2a**), confirming a previously reported association ($r^2=0.96$, D'=1 to lead SNP rs1823125 in 1KG CEU)[12]. For insomnia symptoms (n=32,155 cases, 26,973 controls), significant associations were observed within *MEIS1* (rs113851554T, OR [95%CI]=1.26[1.20-1.33], $p=9.1x10^{-19}$, EAF 0.057, **Fig. 2b**), a homeobox gene implicated in motor neuron connectivity in Drosophila[24,25], retinal and lens

development in mouse[26], and Substance P expression in the amygdala[27], near *TMEM132E* (rs145258459C, 1.23[1.13-1.35], p=2.1x10^{-8}, EAF 0.983, **Fig. 2c**), a gene family with roles in brain development[28], panic/anxiety[29] and bipolar disorder[30], suggesting a link between insomnia symptoms and an underlying broader sensitivity to anxiety and stress, and near *CYCL1* (rs5922858G, OR [95%CI]=1.12[1.07-1.16], p=1.28 x10^{-8}, EAF 0.849, **Fig 2d**) a locus previously associated (p=10^{-6}) with alcohol dependence co-morbid with depressive symptoms[31]. Sex-stratified analyses identified an additional female-specific signal near *TGFBI (*rs3792900C 1.10[1.07-1.14], p=2.16x10^{-8}*,* EAF 0.470; **Table 1, Supplementary Fig. 3q, 3r, Supplementary Table 2***),* an extracellular matrix protein responsible for human corneal dystrophy[32] and a male-specific signal near *WDR27,* a scaffold protein (rs13192566G OR [95%CI]=1.14[1.09-1.20], p=3.2x10^{-8}**,** EAF 0.860**)(Table 1**, **Supplementary Fig. 3s, 3t, 4**; **Supplementary Table 2**). Independent associations at both loci are observed with type 1 diabetes, suggesting an immune role[33-35]. For excessive daytime sleepiness (n=111,648), we identified a signal near the androgen receptor *AR* (rs73536079T, β=0.634, p=3.94x10^{-8}, EAF 0.002, **Fig. 3e)**, with no sex-specific effects. Secondary analyses after additional adjustment for depression or BMI identified a signal near *ROBO1,* (depression adjustment n=107,440, rs182765975T, beta=0.099, p=3.33x10^{-8}, EAF 0.003, **Table 1, Supplementary Figure 3o**), a neuronal axon guidance receptor previously implicated in dyslexia[36], and a signal near another member of the TMEM132 family, *TMEM132B (*BMI adjustment n=75,480, rs142261172A, β=0.106, p=9.06x10^{-9}, EAF 0.004, **Table 1, Supplementary Figure 3p**). Conditional analyses did not identify independent association signals (**Supplementary Table 3**). Sensitivity analyses adjusting for factors influencing sleep traits, including self-reported sleep apnea, depression, psychiatric medication use, smoking, socio-economic status, employment status, marital status, and snoring did not significantly alter results for primary association signals (**Supplementary Table 4**).

The leading associations overlap interesting candidate genes enriched in murine/zebrafish hypocretin expressing neurons[37,38], differentially expressed in sleep-deprived rats[39], and/or regulate sleep in *Drosophila*[40]. Credible set analyses[41] highlighted a number of potential causal variants at each locus (**Table 1**) and future experimental studies will be necessary. Bioinformatic annotations[42] offer an initial opportunity at *in silico* functional interpretation (**Supplementary Table 5; Supplementary Fig. 5**). For example, multiple variants for all three traits are predicted to disrupt binding of *FOXP1*, a neural transcriptional repressor implicated in intellectual disability, autism and language impairment[43]. Interestingly, the *PAX-8* sleep duration association is adjacent to the only chromosomal fusion site since divergence of humans from other hominids ~5 million years ago[44,45], and the novel genomic structure created by this unique evolutionary history may play a causal role. Pathway analysis[46] of significant and suggestive loci revealed enrichment of genes associated with immune, neuro-developmental, pituitary and communication disorders (p<0.01), and enriched for transcription factor-binding sites for stress-responsive heat-shock-factor 1 (HSF1) and endoplasmic reticulum stress/unfolded protein-responsive factor HERPUD1 **(Supplementary Tables 6&7**).

Aside from the lead *PAX-8* SNPs and a *DRD2* region variant[47] for sleep duration, limited evidence of association was observed for previously published candidate gene or GWAS signals (p_{meta}<5x10^{-5}; **Supplementary Table 8**), or for regions encompassing core clock genes (**Supplementary Fig. 6**). Our findings for sleep duration GWAS largely overlap with Jones et al.[18], despite differences in exclusion criteria and analytic approach. Particularly, our study excluded shift workers (n=6,557), sleep medication users (n=1,184) and first-to-third degree relatives (n=7,980), whereas the linear mixed-model analyses by Jones et al. included these populations, leading to a larger sample size (n=127,573). Likely due to this increase in power, Jones et al. identified two additional signals at *VRK2* that did not attain genome-wide

154 significance in our study (rs1380703A β(se)=1.5(0.30) mins/allele, p=8.43x10^{-8} and
155 rs17190618T, β(se)=1.60(0.34) mins/allele, p=3.80x10^{-6}).

156

157 Trait heritability calculated as the proportion of trait variance due to additive genetic factors
158 measured here (observed scale SNP heritability, h^2 (S.E.)) was 10.3 (0.006)% for sleep
159 duration, 20.6 (0.011)% for insomnia symptoms and 8.4 (0.006)% for sleepiness (BOLT-REML
160 variance components analysis[48]). LD-score regression analysis[49] confirmed no residual
161 population stratification (Intercept (SE): Sleep Duration 1.012 (0.008), Insomnia Symptoms
162 1.003 (0.008), Excessive Daytime Sleepiness 1.005 (0.007). Tests for enrichment of heritability
163 by functional class using an LD-score regression approach[50] identified excess heritability across
164 active transcriptional regions for insomnia symptoms and genomic regions conserved in
165 mammals for all three sleep traits. Consistently, heritability enrichment in conserved regions
166 was seen for traits demonstrating significant genetic correlation with sleep (**Fig. 3,**
167 **Supplementary Table 9**).

168

169 Sleep duration, insomnia symptoms, excessive daytime sleepiness, and chronotype, are
170 significantly correlated both at the phenotype and genetic level (**Fig. 1**), with greater pair-wise
171 correlations in males as compared to females (Supplementary Fig.1). Thus, in order to find loci
172 common to sleep traits, we performed a multi-trait GWAS[51]. We identified two novel association
173 signals near *HCRTR2* and *INADL*, and revealed that *PAX-8* and *MEIS-1* associations influence
174 multiple sleep traits (**Fig. 2; Table 2, Supplementary Fig. 7**). *HCRTR2* encodes hypocretin
175 receptor 2, the main receptor of two receptors for wake-promoting orexin neuropeptides[52]
176 involved in narcolepsy and regulation of sleep. Notably, the minor allele at rs3122163 (C)
177 showed sub-threshold association with shorter sleep duration and morningness chronotype,

suggesting gain of function, but no association with insomnia symptoms. Assessment of objective sleep measures, functional and physiologic follow-up should yield important insights into orexin receptor signaling, a pathway important for the pharmacological treatment of narcolepsy[53] and insomnia[54]. *INADL* encodes a membrane protein involved in the formation of tight junctions, and is implicated in photoreception in mice and *Drosophila*[55,56]. The INADL protein is reported to interact with HTR2A[57], a serotonin receptor with a known role in sleep regulation[58,59].

Our strongest association for insomnia symptoms fell within *MEIS1*, a locus previously associated with RLS in GWAS[60]. Our lead SNP rs113851554 and the correlated 3'UTR variant rs11693221 (pair-wise r^2=0.69, D'=0.90 in 1KG EUR) represent the strongest known genetic risk factor for RLS and were identified in follow-up sequencing studies of MEIS1[61,62] of the original RLS GWAS signal rs2300478[60,63]. Conditional analysis suggests that only one underlying signal detected by the lead SNP rs113851554 in our GWAS explains the association of all three SNPs with insomnia symptoms (**Supplementary Fig. 8; Supplementary Table 10).** To further investigate the extent of overlap between RLS and insomnia symptoms, we tested if a weighted genetic risk score (GRS) for RLS[64,65] was also associated with insomnia symptoms with concordant direction of allelic effects (OR [95%CI]= 1.06[1.05-1.07] per RLS risk allele, p=1.17x10^{-21}; **Supplementary Table 11**). Weighting of RLS GWAS alleles by SNP effects on periodic limb movements (PLMs) did not substantially alter overall results (**Supplementary Table 11**). Interestingly, recent data indicating increased thalamic glutamatergic activity in RLS provides evidence for an underlying propensity for hyperarousal in RLS[66], which is also a core feature of insomnia. Future analyses of pair-wise bidirectional causal effects for all three traits will be necessary to determine if shared genetic associations represent causality, partial mediation or pleiotropy.

Strong epidemiologic associations of sleep duration, insomnia symptoms and sleepiness have been observed with disease traits, but the extent to which the underlying genetics is shared is unknown. Therefore, we tested for genome-wide genetic correlation between our sleep GWAS and publicly available GWAS for 20 phenotypes spanning a range of cognitive, neuropsychiatric, anthropometric, cardio-metabolic and auto-immune traits using LD-score regression[67] (**Fig. 4** and **Supplementary Table 12)**.

Genetic correlations demonstrated a strong biological link between longer sleep duration and risk of schizophrenia (r_G=0.29, p=10^{-13}), as suggested by previous reports[18,47,68]. Furthermore, a schizophrenia GRS (96 variants) was associated with longer sleep duration (β(se)=1.44(0.36) mins/allele, p=2.56x10^{-4} [2.3 hr inter-quartile range], although a variety of sleep behaviors are seen in schizophrenia patients[69-71]. Significant genetic correlation between low birth weight and longer sleep duration (r_G= -0.27, p=10^{-4}) may reflect shared links between genetically-determined insulin secretion or action pathways underlying fetal growth[72,73] and long sleep duration. In support, significant genetic correlation was observed by Jones et al.[18] between over-sleepers and both fasting insulin and risk of type 2 diabetes in UK Biobank. Genetic correlation between sleep duration and Crohn's disease risk (r_G=0.18, p=10^{-3}) is also consistent with epidemiologic observations[74].

Significant genetic correlation was also found between increased insomnia symptoms and major depression, adverse glycemic traits, increased adiposity and fewer years of education, and between excessive daytime sleepiness and increased adiposity (all p<10^{-3}), further highlighting biological overlap of sleep traits with metabolism, psychiatric traits, and educational attainment[17]. In support, studies have shown that experimentally suppressing slow wave sleep

leads to decreased insulin sensitivity and impaired glucose tolerance[75,76]. Notably, a fasting insulin GRS was not significantly associated with insomnia symptoms (7 SNPs, OR =1.01, p=0.51). Finally, consistent with a well-established but poorly-understood link between excessive daytime sleepiness and obesity[77,78], a BMI GRS was associated with excessive daytime sleepiness (95 SNPs, β(se) 0.002(0.0004) sleepiness category/allele, p=1.67x10^{-4}), but not with insomnia symptoms (OR=1.00, p=0.73).

Moving forward, replication and systematic testing of genetic correlations in larger samples will be needed. Importantly, genetic correlation testing between insomnia and RLS should be examined, but was not possible here because RLS consortium GWAS results were not available. Additionally, identifying causal relationships between genetically correlated traits may be difficult, and findings using Mendelian randomization approaches will need cautious interpretation given potential selection biases in UK Biobank[79-81].

In summary, in a GWAS of sleep traits, we identified new genetic loci that point to previously unstudied variants might modulate the hypocretin/orexin system, retinal development, and influence cerebral cortex genes. Furthermore, genome-wide analysis suggests that sleep traits share underlying genetic pathways with neuropsychiatric and metabolic disease. This work should advance understanding of molecular processes underlying sleep disturbances, and open new avenues of treatment for sleep disorders and related disorders

Methods

Population and study design

Study participants were from the UK Biobank study, described in detail elsewhere[80-82]. In brief, the UK Biobank is a prospective study of >500,000 people living in the United Kingdom. All people in the National Health Service registry who were aged 40-69 and living <25 miles from a study center were invited to participate between 2006-2010. In total 503,325 participants were recruited from over 9.2 million mailed invitations. Self-reported baseline data was collected by questionnaire and anthropometric assessments were performed. For the current analysis, individuals of non-white ethnicity were excluded to avoid confounding effects. All participants provided informed consent to the UK Biobank.

Sleep quality, quantity and covariate measures

Study subjects self-reported sleep duration, insomnia symptoms, excessive daytime sleepiness, depression, medication use, age, sex, height and weight on a touch-screen questionnaire. For sleep duration, subjects were asked, “About how many hours sleep do you get in every 24 hours? (please include naps)?” with responses in hour increments. To assess insomnia symptoms, subjects were asked, “Do you have trouble falling asleep at night or do you wake up in the middle of the night?" with responses “never/rarely”, “sometimes”, “usually”, “prefer not to answer”. To assess daytime sleepiness, subjects were asked “How likely are you to doze off or fall asleep during the daytime when you don't mean to? (e.g. when working, reading or driving)?" with responses “never/rarely”, “sometimes”, “often”, “all the time”, “don’t know”, “prefer not to answer”. Approximately 500,000 subjects answered these questions, but only the 120,286 unrelated individuals with genetic data and European ancestry were considered for this analysis. Subjects with self-reported shift work (n=6,557) or sleep medication use (n=1,184) were excluded. Subjects who responded “Do not know” or “Prefer not to answer” were set to missing. Sleep duration and excessive daytime sleepiness were untransformed and treated as continuous variables, with daytime sleepiness coded 1-4. The insomnia symptom trait was dichotomized into controls (“never/rarely”) and cases (“usually”). Covariates used in sensitivity analyses included self-reported sleep apnea, BMI, depression, psychiatric medication use, socio-economic, smoking, employment and marital status, and snoring, and secondary GWAS for sleepiness included adjustment for BMI or depression. Sleep apnea cases were defined based on ICD10 diagnosis code (391 cases). BMI at baseline visit was calculated from entries of height and weight (n=75,540 with available data). Depression was reported in answer to the question “How often did you feel down, depressed or hopeless mood in last 2 weeks?” (cases, n=4,242 based on answers “more than half the days”, or “nearly every day”). Medication use was self-reported as part of the initial UK Biobank interview. Our list of psychiatric medication for sensitivity analysis included the four most widely used: fluoxetine (Prozac), citalopram (Cipranol), paroxetine (Seroxat), and sertraline (Lustral). Our list of sleep medications included the 21 most widely used sleep medications in the UK Biobank: oxazepam, meprobamate, medazepam, bromazepam, lorazepam, clobazam, chlormezanone, temazepam, nitrazepam, lormetazepam, diazepam, zopiclone, triclofos, methyprylone, prazepam, triazolam, ketazolam, dichloralphenazone, clomethiazole, zaleplon, butobarbital. Smoking status was self-reported as past smoking behavior and current smoking behavior, and classified into “current”, “past”, or “never” smoked. Socio-economic status was represented by the Townsend deprivation index,

based on national census data immediately preceding participation in the UK Biobank. Employment status was self-reported (cases=retired, controls=currently employed). Marital status was derived from self-reported household occupancy and relatedness data. Snoring was reported in answer to the question "Does your partner or a close relative or friend complain about your snoring?".

Genotyping, quality control and imputation

Of the ~500,000 subjects with phenotype data in the UK Biobank, ~153,000 are currently genotyped. Genotyping was performed by the UK Biobank, and genotyping, quality control, and imputation procedures are described in detail at the UK Biobank website (http://biobank.ctsu.ox.ac.uk/). In brief, blood, saliva, and urine was collected from participants, and DNA was extracted from the buffy coat samples. Participant DNA was genotyped on two arrays, UK BiLEVE and UKB Axiom with >95% common content. Genotypes were called using Affymetrix Power Tools software. Sample and SNP quality control were performed. Samples were removed for high missingness or heterozygosity (480 samples), short runs of homozygosity (8 samples), related individuals (1,856 samples), and sex mismatches (191 samples). Genotypes for 152,736 samples passed sample QC (~99.9% of total samples). SNPs were excluded if they did not pass QC filters across all 33 genotyping batches. Batch effects were identified through frequency and Hardy-Weinberg equilibrium tests (p-value $<10^{-12}$). Before imputation, 806,466 SNPs pass QC in at least one batch (>99% of the array content). Population structure was captured by principal component analysis on the samples using a subset of high quality (missingness <1.5%), high frequency SNPs (>2.5%) (~100,000 SNPs) and identified the sub-sample of European descent. Imputation of autosomal SNPs was performed to a merged reference panel of the Phase 3 1000 Genome Project and the UK10K using IMPUTE2[83]. Data were prephased using SHAPEIT3[84]. In total, 73,355,677 SNPs, short indels and large structural variants were imputed. X-chromosome data were imputed separately, using Eagle 2.0 for pre-phasing with the –X chromosome flag (no reference panel) in the entire cohort[85] and IMPUTE2[83] with the Phase 3 1KG Project reference panel for imputation using the –chrX flag on 500kb chunks in randomly assigned subsets of 30,000 individuals. Post-imputation QC was performed as previously outlined (http://biobank.ctsu.ox.ac.uk/) and an imputation info score cut-off of 0.8 was applied. For GWAS, we further excluded SNPs with MAF <0.001, maximum per SNP missingness of 10%, and maximum per sample missingness of 40%. In total, up to 112,586 samples of European descent with high quality genotyping and complete phenotype/covariate data were used for these analyses.

Statistical Analysis

Phenotypic correlation analysis was performed using the Spearman test in R using the Hmisc package. Genetic association analysis for autosomes was performed in SNPTEST[86,87] with the "expected" method using an additive genetic model adjusted for age, sex, 10 PCs and genotyping array. Genome-wide association analysis was performed separately for sleep duration, insomnia symptoms, and excessive daytime sleepiness with a genome-wide significance threshold of $5x10^{-8}$ for each GWAS. We are 80% powered to detect the following

effects: sleep duration β=0.045 hrs (2.7 mins), insomnia symptoms OR=1.07, and excessive daytime sleepiness β=0.021 units (assuming a MAF 0.1, p=5x10^{-7}) and 80% powered to detect the following effects: sleep duration β= 0.048 hrs (2.9 mins), insomnia symptoms OR=1.08 and excessive daytime sleepiness β=0.023 units (assuming a MAF 0.1, p=5x10^{-8}). X-chromosome analysis was performed in PLINK 1.9[88] using linear/logistic regression with separate analysis of the pseudoautosomal regions using the split chromosome flag, adjusting for sex, age, 10 PCs and genotyping array. For the X chromosome signal at rs73536079, we verified using principal components analysis that all carriers of the minor allele fall within the major European ancestry cluster. Follow-up analyses on genome-wide suggestive and significant loci in the primary analyses included covariate sensitivity analysis individually adjusting for sleep apnea, depression, psychiatric medication use, socio-economic, smoking, employment and marital status, and snoring, or BMI (on top of the baseline model adjusting for age, sex, 10 PCs and genotyping array). Sensitivity analysis was conducted only in the subset of subjects with all secondary covariates (n=75,477 for sleep duration, n=39,812 for insomnia symptoms and n=75,640 for excessive daytime sleepiness). Enrichment for disease associated gene sets and transcription factors was performed in WebGestalt[46] using the human genome as the reference set, the Benjamini Hochberg adjustment for multiple testing, and a minimum number of 2 genes per category. Sex specific GWAS were performed in PLINK 1.9[88] using linear/logistic regression stratified by sex adjusting for age, 10 principal components of ancestry, and genotyping array. We used a hard-call genotype threshold of 0.1 (calls with greater than 0.1 are treated as missing), SNP imputation quality threshold of 0.80, and a MAF threshold of 0.001. Regional association plots were made using Locuszoom with the HG19 Nov2014 EUR reference panel for background linkage disequilibrium[89].

Trait heritability was calculated as the proportion of trait variance due to additive genetic factors across the autosomes measured in this study using BOLT-REML[48], to leverage the power of raw genotype data together with low frequency variants (MAF≥0.001). For multi-trait genome-wide association analysis we applied the CPASSOC package developed by Zhu et al.[51] to combine association evidence of chronotype, sleep duration, insomnia symptoms and excessive daytime sleepiness. CPASSOC provides two statistics, SHom and SHet. SHom is similar to the fixed effect meta-analysis method[90] but accounting for the correlation of summary statistics because of the correlated traits. SHom uses a sample size of a trait as a weight instead of variance, so that it is possible to combine traits with different measurement scales. SHet is an extension of SHom but power can be improved when the genetic effect sizes are different for different traits. The distribution of SHet under the null hypothesis was obtained through an estimated beta distribution. To calculate statistics SHom and SHet, a correlation matrix is required to account for the correlation among traits or induced by overlapped or related samples from different cohorts. In this study, we directly provide the correlation matrix calculated from the residuals of four sleep traits after adjusting for age, sex, PCs of ancestry and genotyping array. Post-GWAS genome-wide genetic correlation analysis of LD Score Regression (LDSC)[67] was conducted using all UK Biobank SNPs also found in HapMap3[89] and included publicly available data from 20 published genome-wide association studies, with a significance threshold of p=0.0026 after Bonferroni correction for all 20 tests performed. As expected, the observed heritability estimates from LDSC[67] using summary statistics for HapMap3 are lower (5.7

(0.0065)% for sleep duration, 13.3 (0.0123)% for insomnia symptoms and 5.3 (0.005)% for sleepiness) than those calculated by Bolt-REML[48] using primary data (10.3 (0.006)% for sleep duration, 20.6 (0.011)% for insomnia symptoms and 8.4 (0.006)% for sleepiness), because the HapMap3 panel restricts to variants with >5% MAF. LDSC estimates genetic correlation between two traits from summary statistics (ranging from -1 to 1) using the fact that the GWAS effect-size estimate for each SNP incorporates effects of all SNPs in LD with that SNP, SNPs with high LD have higher X^2 statistics than SNPs with low LD, and a similar relationship is observed when single study test statistics are replaced with the product of z-scores from two studies of traits with some correlation[67]. Furthermore, genetic correlation is possible between case/control studies and quantitative traits, as well as within these trait types. We performed a weighted genetic risk score analysis using risk scores for restless legs syndrome, schizophrenia, body mass index, and fasting insulin. Risk score SNPs passed the genome-wide significance threshold ($p<5x10^{-8}$) from recent large-scale genome-wide association studies and were present in the UK Biobank (restless legs syndrome 7 SNPs **Supp Table 11**[65]; schizophrenia 96 SNPs[91]; BMI 95 SNPs[92]; fasting insulin 7 SNPs[93]). Independent SNPs were identified and beta estimates recorded for calculation of the weighted risk score. The genetic risk score was calculated by summing the products of the risk allele count multiplied by the effect reported in the discovery GWAS paper. The additive genotype model was used for all SNPs. We performed partitioning of heritability using the 25 pre-computed functional annotations available through LDSC, which were curated from large-scale robust datasets[50]. Enrichment both in the functional regions and in an expanded region (+500bp) around each functional class was calculated in order to prevent the estimates from being biased upward by enrichment in nearby regions. The multiple testing threshold was determined using the conservative Bonferroni correction (p of 0.05/25 classes). Summary GWAS statistics will be made available at the UK Biobank web site (http://biobank.ctsu.ox.ac.uk/).

Author Contributions

The study was designed by JML, MKR, and RS. JML, JL, IV and RS performed genetic analyses. JML and RS wrote the manuscript and all co-authors helped interpret data, reviewed and edited the manuscript, before approving its submission. RS is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgements

This research has been conducted using the UK Biobank Resource. We would like to thank the participants and researchers from the UK Biobank who contributed or collected data. This work was supported by NIH grants R01DK107859 (RS), R21HL121728 (RS), F32DK102323 (JML), R01HL113338 (JML, SR and RS), R01DK102696 (RS and FS), R01DK105072 (RS and FS), T32HL007567(JL), HG003054 (XZ), The University of Manchester (Research Infrastructure Fund), the Wellcome Trust (salary support for DWR and AL) and UK Medical Research Council MC_UU_12013/5 (DAL). Data on glycemic traits have been contributed by MAGIC investigators and have been downloaded from www.magicinvestigators.org. Data on coronary artery disease / myocardial infarction have been contributed by CARDIo-GRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMPLUSC4D.ORG. We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses.

The authors have no competing financial interests to declare.

Table 1. Genome-wide significant ($p<5x10^{-8}$) and suggestive ($p<5x10^{-7}$) loci associated with sleep duration, insomnia symptoms, and excessive daytime sleepiness in subjects of European ancestry in the UKBiobank.

Trait	SNP	Chr:position NCBI 37	Nearest Gene(s)	Alleles (E/A)	EAF	Imputation Quality	Beta (SE)	SE	*p-val*	Most likely causal SNPs (probability)†
Sleep Duration (n=111,975)										
	rs62158211	**2:114106139**	***PAX8***	**T/G**	**0.213**	0.99	**0.039**	**0.005**	**4.72×10^{-14}**	rs62158211 (0.16), rs62158213 (0.16), rs4618068 (0.16), rs1807282 (0.16), rs56093896 (0.16)
	rs1380703	2:57941287	*VRK2/LOC647016/LOC100131953*	A/G	0.618	0.89	0.025	0.005	8.44×10^{-8}	rs1380703 (1)
	rs10953765	7:114291435	*FOXP2*	G/A	0.447	0.98	0.022	0.004	2.96×10^{-7}	rs10953765 (0.27), rs1456031 (0.14)
	rs146977851	10:56570954	*PCDH15*	C/T	0.971	0.97	0.065	0.013	3.53×10^{-7}	rs146977851 (0.85), rs75334053 (0.14)
	rs61980273	14:94218949	*PRIMA1/UNC79*	A/G	0.039	1.00	0.058	0.011	1.30×10^{-7}	rs61980273 (1)
Insomnia Symptoms (n up to 31,767 cases and 26,935 controls)							OR	95% CI		
	rs576106307	1:18007282	*ARHGEF10L*	C/CT	0.934	0.89	1.07	1.10-1.04	2.66×10^{-7}	rs576106307 (1)
	rs113851554	**2:66750564**	***MEIS1***	**T/G**	**0.057**	1.00	**1.26**	**1.20-1.33**	**9.11×10^{-19}**	rs113851554 (0.98)
	rs376775068	8:145604659	*ADCK5*	G/C	0.934	0.67	1.11	1.16-1.06	6.81×10^{-8}	rs376775068 (1)
	rs145258459	**17:32986155**	***TMEM132E***	**C/T**	**0.983**	0.69	**1.23**	**1.13-1.35**	**2.13×10^{-8}**	rs145258459 (1.0)
	rs531814036	17:43219921	*ACBD4*	C/CT	0.419	0.91	1.06	1.03-1.08	2.92×10^{-7}	rs531814036 (1)
	rs5922858	**X:82971008**	***CYCL1***	**G/T**	**0.849**	**0.99**	**1.12**	**1.07-1.16**	**1.28×10^{-8}**	rs5922858 (1)
Males	**rs13192566**	**6:169961635**	***WDR27***	**G/C**	**0.860**	**0.99**	**1.14**	**1.09-1.20**	**3.17×10^{-8}**	rs13192566 (0.50), rs13208844 (0.50)
Females	**rs3792900**	**5:135393754**	***TGFBI***	**C/T**	**0.470**	**0.99**	**1.1**	**1.07-1.14**	**2.16×10^{-8}**	rs3792900 (0.14), rs6894815 (0.07)

Excessive Daytime Sleepiness (n<111,648)							Beta	SE		
rs192315283	1:59531543	*HSD52*	C/T	0.010	0.76	0.126	0.025	3.55×10^{-7}	rs192315283 (1)
rs76645968	2:53827686	*ASB3*	G/C	0.977	0.99	0.073	0.014	1.79×10^{-7}	rs76645968 (0.26), rs12328289 (0.26)
rs920065	3:5893776	*MRPS35P1/ MRPS36P1*	C/G	0.824	0.96	0.028	0.006	4.25×10^{-7}	rs920065 (0.49)
rs115320831	4:159178375	*TMEM144*	A/G	0.702	0.98	0.024	0.005	3.68×10^{-7}	rs115320831 (0.58)
rs35309287	5:146775386	*DPYSL3*	TA/T	0.970	0.94	0.067	0.013	1.25×10^{-7}	rs35309287 (0.45), rs34398961 (0.45)
rs189689339	6:82375372	*FAM46A*	T/C	0.003	0.67	0.226	0.044	2.13×10^{-7}	rs189689339 (1)
rs17507216	15:83226925	*CPEB1*	A/G	0.232	1.00	0.026	0.005	1.59×10^{-7}	rs17507216 (0.20), rs72751643 (0.11)
rs73536079	**X:67154206**	***AR/OPHN1***	**T/G**	**0.002**	**0.90**	**0.634**	**0.115**	**3.94×10^{-8}**	rs73536079 (1)
rs182765975*	**3:78538431**	***ROBO1***	**T/G**	**0.003**	**0.86**	**0.099**	**0.018**	**3.33×10^{-8}**	rs182765975 (0.33), rs191435135 (0.33), rs182979911 (0.33)
rs142261172**	**12:126049981**	***TMEM132B***	**A/G**	**0.004**	**0.92**	**0.106**	**0.018**	**9.06×10^{-9}**	rs142261172 (0.50), rs189248622 (0.50)

E=effect allele, A=alternative allele, Chr=chromosome, OR=Odds Ratio, CI=confidence interval, INFO=imputation quality from Impute2. EAF=effect allele frequency. Note, increasing beta and Odds Ratio indicate longer sleep duration in hours, increased insomnia symptoms, and increased sleepiness. Analyses are adjusted for age, sex, genetic ancestry and genotyping array. * denotes secondary analysis with additional adjustment for depression. **denotes secondary analysis with additional adjustment for body mass index. Bold denotes genome-wide significant signals ($p<5x10^{-8}$). † Using PICS.

Table 2. Genome-wide significant ($p<5x10^{-8}$) loci associated with a multiphenotype model of sleep duration, insomnia symptoms, excessive daytime sleepiness and categorical chronotype in subjects of European ancestry in the UKBiobank.

SNP	Chr:position NCBI 37	Nearest Gene	Alleles (E/A)	EAF	Imputation Quality	Multitrait *p*-val	Sleep Duration		Insomnia Symptoms		Excessive Daytime Sleepiness		Chronotype		Causal SNPs (probability)
							Beta	SE	OR	95% CI	Beta	SE	Beta	SE	
							p-val		*p*-val		*p*-val		*p*-val		
rs12140153	1:62352479	*INADL*	T/G	0.099	0.93	$1.06x10^{-10}$	-0.009	0.007	1.039	0.999-1.08	-0.036	0.007	0.036	0.008	rs12140153 (1)
							0.22		0.05		6.60x10-7		$2.59x10^{-6}$		
rs76681500	1:77247749	*AK5*	A/G	0.159	0.99	$1.03x10^{-9}$	-0.002	0.006	0.98	0.950-1.011	-0.008	0.006	-0.043	0.006	rs76681500 (0.5732)
							0.79		0.27		0.15		$1.50x10-^{12}$		
rs694383	1:180834827	*RGS16*	C/G	0.030	1.00	$2.72x10^{-11}$	0.018	0.012	0.98	0.917-1.048	-0.009	0.012	0.099	0.013	rs694383 (0.2207), rs509476 (0.2207), rs1144566 (0.2207), rs12743617 (0.2207)
							0.14		0.75		0.47		$2.61x10^{-14}$		
rs113851554	2:66523432	*MEIS1*	T/G	0.056	1.00	$3.97x10^{-16}$	0.001	0.009	1.264	1.202-1.329	-0.002	0.009	0.033	0.01	rs113851554 (0.9619)
							0.95		$9.11x10^{-19}$		0.85		$5.64x10^{-4}$		
rs62158211	2:113822609	*PAX8*	T/G	0.214	0.99	$8.18x10^{-13}$	0.039	0.005	0.943	0.917-0.969	0.005	0.005	0.014	0.005	rs62158211 (0.1547),rs62158213 (0.1547), rs4618068 (0.1547), rs1807282 (0.1547), rs56093896 (0.1547)
							$4.72x10^{-14}$		$1.31x10^{-5}$		0.37		$7.93x10^{-3}$		
rs3122163	6:55164327	*HCRTR2*	T/C	0.768	0.99	$4.18x10^{-10}$	0.019	0.005	0.984	0.957-1.011	-0.023	0.005	0.021	0.005	rs3122163 (0.0833), rs34694541 (0.0833), rs3122170 (0.0833)
							$9.97x10^{-5}$		0.52		5.51x10-6		$8.68x10^{-5}$		

E=effect allele, A=alternative allele, Chr=chromosome, OR=Odds Ratio, CI=confidence interval. EAF=effect allele frequency. Note, increasing beta and Odds Ratio indicate longer sleep duration, increased insomnia symptoms, increased daytime sleepiness, and later chronotype.

Figure 1. Sleep traits are phenotypically and genetically correlated. **a.** Phenotypic correlation between the reported sleep traits, using Spearman correlation (r). **b.** Genetic correlation (r_G) between the reported sleep traits, using LD-score regression[67]. Color scale represents the strength of the correlation. Chronotype ranges from extreme morning types to extreme evening types.

Figure 2. Regional association plots for genome-wide significant loci. Panel **a** sleep duration, **b-d** insomnia symptoms, **e** excessive daytime sleepiness, **f-g** composite trait of sleep duration, insomnia symptoms, excessive daytime sleepiness, and chronotype. Chromosomal position is indicated on the x-axis and $-\log_{10}$ *p*-values for each SNP (filled circles/squares) is indicated on the y-axis, with the lead SNP shown in purple (400kb window around lead SNP shown). Genes within the region are shown in the lower panel. The blue line indicates the recombination rate. Additional SNPs in the locus are colored according to linkage disequilibrium (r^2) with the lead SNP (estimated by LocusZoom based on the CEU HapMap haplotypes or within UK Biobank (panel **c**). Squares represent directly genotyped SNPs, and circles represent imputed SNPs.

Figure 3. Partitioning of genetic architecture of sleep duration, insomnia symptoms, and excessive daytime sleepiness across functional annotation categories. Fold enrichment estimates for the main annotations of LD-score regression[50] are indicated on the y-axis across functional annotation class on the x-axis for each trait. Error bars represent the 95% confidence interval around the estimate. 25 functional annotations were tested, and annotations passing the multiple testing threshold ($p<0.005$) are shown. For context, the average enrichment across functional annotation categories is shown for 9 traits with significant genetic correlation to at least one sleep trait (GWAS traits correlated with Sleep: includes GWAS for BMI, waist circumference, birth weight, depression, educational attainment, three glycemic traits in non-diabetics, and schizophrenia) or for 5 traits with no significant genetic correlation to any sleep traits (GWAS traits uncorrelated with Sleep: includes GWAS for Alzheimer's Disease, Type 2 Diabetes, autism, rheumatoid arthritis, and height). Abbreviations: H3K9=histone H3 lysine 9.

Figure 4. Shared genetic architecture between sleep duration, insomnia symptoms, or excessive daytime sleepiness and 20 behavioral and disease traits. LD-score regression[67] estimates of genetic correlation (r_G) of sleep duration, insomnia symptoms, and excessive daytime sleepiness are compared with the summary statistics from 20 publicly available genome-wide association studies of psychiatric and metabolic disorders, immune diseases, and other traits of natural variation. Blue, positive genetic correlation; red, negative genetic correlation, r_g values displayed for significant correlations. Larger squares correspond to more significant P values. Genetic correlations that are significantly different from zero after Bonferroni correction are marked with an asterisk, after Bonferroni correction *p*-value cut-off is 0.0025. All genetic correlations in this report can be found in tabular form in **Supplementary Table 12**. Abbreviations: BMI=body mass index, BMD=bone mineral density, HOMA-IR= Homeostatic model assessment of insulin resistance. * $p<10^{-3}$, ** $p<10^{-5}$, *** $p<10^{-7}$.

1. Fernandez-Mendoza, J. & Vgontzas, A.N. Insomnia and its impact on physical and mental health. *Curr Psychiatry Rep* **15**, 418 (2013).
2. Luyster, F.S. *et al.* Sleep: a health imperative. *Sleep* **35**, 727-34 (2012).
3. Stranges, S., Tigbe, W., Gomez-Olive, F.X., Thorogood, M. & Kandala, N.B. Sleep problems: an emerging global epidemic? Findings from the INDEPTH WHO-SAGE study among more than 40,000 older adults from 8 countries across Africa and Asia. *Sleep* **35**, 1173-81 (2012).
4. de Castro, J.M. The influence of heredity on self-reported sleep patterns in free-living humans. *Physiol Behav* **76**, 479-86 (2002).
5. Evans, D.S. *et al.* Habitual sleep/wake patterns in the Old Order Amish: heritability and association with non-genetic factors. *Sleep* **34**, 661-9 (2011).
6. Heath, A.C., Eaves, L.J., Kirk, K.M. & Martin, N.G. Effects of lifestyle, personality, symptoms of anxiety and depression, and genetic predisposition on subjective sleep disturbance and sleep pattern. *Twin Res* **1**, 176-88 (1998).
7. Heath, A.C., Kendler, K.S., Eaves, L.J. & Martin, N.G. Evidence for genetic influences on sleep disturbance and sleep pattern in twins. *Sleep* **13**, 318-35 (1990).
8. Partinen, M., Kaprio, J., Koskenvuo, M., Putkonen, P. & Langinvainio, H. Genetic and environmental determination of human sleep. *Sleep* **6**, 179-85 (1983).
9. Wing, Y.K. *et al.* Familial aggregation and heritability of insomnia in a community-based study. *Sleep Med* **13**, 985-90 (2012).
10. He, Y. *et al.* The transcriptional repressor DEC2 regulates sleep length in mammals. *Science* **325**, 866-70 (2009).
11. Gottlieb, D.J., O'Connor, G.T. & Wilk, J.B. Genome-wide association of sleep and circadian phenotypes. *BMC Med Genet* **8 Suppl 1**, S9 (2007).
12. Gottlieb, D.J. *et al.* Novel loci associated with usual sleep duration: the CHARGE Consortium Genome-Wide Association Study. *Mol Psychiatry* **20**, 1232-9 (2015).
13. Byrne, E.M. *et al.* A genome-wide association study of sleep habits and insomnia. *Am J Med Genet B Neuropsychiatr Genet* **162B**, 439-51 (2013).
14. Allebrandt, K.V. *et al.* A K(ATP) channel gene effect on sleep duration: from genome-wide association studies to function in Drosophila. *Mol Psychiatry* **18**, 122-32 (2013).
15. Gehrman, P.R., Keenan, B.T., Byrne, E.M. & Pack, A.I. Genetics of Sleep Disorders. *Psychiatr Clin North Am* **38**, 667-81 (2015).
16. Sehgal, A. & Mignot, E. Genetics of sleep and sleep disorders. *Cell* **146**, 194-207 (2011).
17. Lane, J.M. *et al.* Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. *Nat Commun* **7**, 10889 (2016).
18. Jones, S.E. *et al.* Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. *PLoS Genet* **12**, e1006125 (2016).
19. Hu, Y. *et al.* GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat Commun* **7**, 10448 (2016).
20. Pemberton, R. & Fuller Tyszkiewicz, M.D. Factors contributing to depressive mood states in everyday life: A systematic review. *J Affect Disord* **200**, 103-10 (2016).
21. Foral, P., Knezevich, J., Dewan, N. & Malesker, M. Medication-induced sleep disturbances. *Consult Pharm* **26**, 414-25 (2011).
22. Rosenberg, R.P. Clinical assessment of excessive daytime sleepiness in the diagnosis of sleep disorders. *J Clin Psychiatry* **76**, e1602 (2015).
23. Gonnissen, H.K. *et al.* Sleep duration, sleep quality and body weight: parallel developments. *Physiol Behav* **121**, 112-6 (2013).

24. Kurant, E. *et al.* Dorsotonals/homothorax, the Drosophila homologue of meis1, interacts with extradenticle in patterning of the embryonic PNS. *Development* **125**, 1037-48 (1998).
25. Casares, F. & Mann, R.S. Control of antennal versus leg development in Drosophila. *Nature* **392**, 723-6 (1998).
26. Hisa, T. *et al.* Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. *EMBO J* **23**, 450-9 (2004).
27. Davidson, S., Miller, K.A., Dowell, A., Gildea, A. & Mackenzie, A. A remote and highly conserved enhancer supports amygdala specific expression of the gene encoding the anxiogenic neuropeptide substance-P. *Mol Psychiatry* **11**, 323, 410-21 (2006).
28. Oh-hashi, K., Naruse, Y., Amaya, F., Shimosato, G. & Tanaka, M. Cloning and characterization of a novel GRP78-binding protein in the rat brain. *J Biol Chem* **278**, 10531-7 (2003).
29. Erhardt, A. *et al.* Replication and meta-analysis of TMEM132D gene variants in panic disorder. *Transl Psychiatry* **2**, e156 (2012).
30. Sklar, P. *et al.* Whole-genome association study of bipolar disorder. *Mol Psychiatry* **13**, 558-69 (2008).
31. Edwards, A.C. *et al.* Genome-wide association study of comorbid depressive syndrome and alcohol dependence. *Psychiatr Genet* **22**, 31-41 (2012).
32. Han, K.E. *et al.* Pathogenesis and treatments of TGFBI corneal dystrophies. *Prog Retin Eye Res* **50**, 67-88 (2016).
33. Bradfield, J.P. *et al.* A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet* **7**, e1002293 (2011).
34. Patry, M. *et al.* betaig-h3 Represses T-Cell Activation in Type 1 Diabetes. *Diabetes* **64**, 4212-9 (2015).
35. Han, B. *et al.* TGFBI (betaIG-H3) is a diabetes-risk gene based on mouse and human genetic studies. *Hum Mol Genet* **23**, 4597-611 (2014).
36. Poelmans, G., Buitelaar, J.K., Pauls, D.L. & Franke, B. A theoretical molecular network for dyslexia: integrating available genetic findings. *Mol Psychiatry* **16**, 365-82 (2011).
37. Dalal, J. *et al.* Translational profiling of hypocretin neurons identifies candidate molecules for sleep regulation. *Genes Dev* **27**, 565-78 (2013).
38. Yelin-Bekerman, L. *et al.* Hypocretin neuron-specific transcriptome profiling identifies the sleep modulator Kcnh4a. *Elife* **4**, e08638 (2015).
39. Mackiewicz, M. *et al.* Macromolecule biosynthesis: a key function of sleep. *Physiol Genomics* **31**, 441-57 (2007).
40. Takahama, K. *et al.* Pan-neuronal knockdown of the c-Jun N-terminal Kinase (JNK) results in a reduction in sleep and longevity in Drosophila. *Biochem Biophys Res Commun* **417**, 807-11 (2012).
41. Farh, K.K. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337-43 (2015).
42. Ward, L.D. & Kellis, M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* **44**, D877-81 (2016).
43. Hamdan, F.F. *et al.* De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet* **87**, 671-8 (2010).
44. Fan, Y., Newman, T., Linardopoulou, E. & Trask, B.J. Gene content and function of the ancestral chromosome fusion site in human chromosome 2q13-2q14.1 and paralogous regions. *Genome Res* **12**, 1663-72 (2002).

45. Fan, Y., Linardopoulou, E., Friedman, C., Williams, E. & Trask, B.J. Genomic structure and evolution of the ancestral chromosome fusion site in 2q13-2q14.1 and paralogous regions on other human chromosomes. *Genome Res* **12**, 1651-62 (2002).
46. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res* **41**, W77-83 (2013).
47. Cade, B.E. *et al.* Common variants in DRD2 are associated with sleep duration: the CARe consortium. *Hum Mol Genet* **25**, 167-79 (2016).
48. Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* **47**, 284-90 (2015).
49. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
50. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
51. Zhu, X. *et al.* Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am J Hum Genet* **96**, 21-36 (2015).
52. Mignot, E. Sleep, sleep disorders and hypocretin (orexin). *Sleep Med* **5 Suppl 1**, S2-8 (2004).
53. Thompson, M.D., Xhaard, H., Sakurai, T., Rainero, I. & Kukkonen, J.P. OX1 and OX2 orexin/hypocretin receptor pharmacogenetics. *Front Neurosci* **8**, 57 (2014).
54. Herring, W.J. *et al.* Suvorexant in Patients With Insomnia: Results From Two 3-Month Randomized Controlled Clinical Trials. *Biol Psychiatry* **79**, 136-48 (2016).
55. Shieh, B.H. & Niemeyer, B. A novel protein encoded by the InaD gene regulates recovery of visual transduction in Drosophila. *Neuron* **14**, 201-10 (1995).
56. Peirson, S.N. *et al.* Microarray analysis and functional genomics identify novel components of melanopsin signaling. *Curr Biol* **17**, 1363-72 (2007).
57. Becamel, C. *et al.* The serotonin 5-HT2A and 5-HT2C receptors interact with specific sets of PDZ proteins. *J Biol Chem* **279**, 20257-66 (2004).
58. Sharpley, A.L., Elliott, J.M., Attenburrow, M.J. & Cowen, P.J. Slow wave sleep in humans: role of 5-HT2A and 5-HT2C receptors. *Neuropharmacology* **33**, 467-71 (1994).
59. Rosenberg, R. *et al.* APD125, a selective serotonin 5-HT(2A) receptor inverse agonist, significantly improves sleep maintenance in primary insomnia. *Sleep* **31**, 1663-71 (2008).
60. Winkelmann, J. *et al.* Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet* **39**, 1000-6 (2007).
61. Xiong, L. *et al.* MEIS1 intronic risk haplotype associated with restless legs syndrome affects its mRNA and protein expression levels. *Hum Mol Genet* **18**, 1065-74 (2009).
62. Schulte, E.C. *et al.* Targeted resequencing and systematic in vivo functional testing identifies rare variants in MEIS1 as significant contributors to restless legs syndrome. *Am J Hum Genet* **95**, 85-95 (2014).
63. Spieler, D. *et al.* Restless legs syndrome-associated intronic common variant in Meis1 alters enhancer function in the developing telencephalon. *Genome Res* **24**, 592-603 (2014).
64. Moore, H.t. *et al.* Periodic leg movements during sleep are associated with polymorphisms in BTBD9, TOX3/BC034767, MEIS1, MAP2K5/SKOR1, and PTPRD. *Sleep* **37**, 1535-42 (2014).
65. Winkelmann, J. *et al.* Genome-wide association study identifies novel restless legs syndrome susceptibility loci on 2p14 and 16q12.1. *PLoS Genet* **7**, e1002171 (2011).
66. Allen, R.P., Barker, P.B., Horska, A. & Earley, C.J. Thalamic glutamate/glutamine in restless legs syndrome: increased and related to disturbed sleep. *Neurology* **80**, 2028-34 (2013).
67. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* (2015).

68. Byrne, E.M., Gehrman, P.R., Trzaskowski, M., Tiemeier, H. & Pack, A.I. Genetic Correlation Analysis Suggests Association between Increased Self-Reported Sleep Duration in Adults and Schizophrenia and Type 2 Diabetes. *Sleep* **39**, 1853-1857 (2016).
69. Wulff, K., Dijk, D.J., Middleton, B., Foster, R.G. & Joyce, E.M. Sleep and circadian rhythm disruption in schizophrenia. *Br J Psychiatry* **200**, 308-16 (2012).
70. Poulin, J. *et al.* Sleep habits in middle-aged, non-hospitalized men and women with schizophrenia: a comparison with healthy controls. *Psychiatry Res* **179**, 274-8 (2010).
71. Chouinard, S., Poulin, J., Stip, E. & Godbout, R. Sleep in untreated patients with schizophrenia: a meta-analysis. *Schizophr Bull* **30**, 957-67 (2004).
72. Hattersley, A.T. & Tooke, J.E. The fetal insulin hypothesis: an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet* **353**, 1789-92 (1999).
73. Horikoshi, M. *et al.* New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat Genet* **45**, 76-82 (2013).
74. Ananthakrishnan, A.N. *et al.* Sleep duration affects risk for ulcerative colitis: a prospective cohort study. *Clin Gastroenterol Hepatol* **12**, 1879-86 (2014).
75. Tasali, E., Leproult, R., Ehrmann, D.A. & Van Cauter, E. Slow-wave sleep and the risk of type 2 diabetes in humans. *Proc Natl Acad Sci U S A* **105**, 1044-9 (2008).
76. Nedeltcheva, A.V. & Scheer, F.A. Metabolic effects of sleep disruption, links to obesity and diabetes. *Curr Opin Endocrinol Diabetes Obes* **21**, 293-8 (2014).
77. Vgontzas, A.N. *et al.* Obesity without sleep apnea is associated with daytime sleepiness. *Arch Intern Med* **158**, 1333-7 (1998).
78. Bixler, E.O. *et al.* Excessive daytime sleepiness in a general population sample: the role of sleep apnea, age, obesity, diabetes, and depression. *J Clin Endocrinol Metab* **90**, 4510-5 (2005).
79. Swanson, J.M. The UK Biobank and selection bias. *Lancet* **380**, 110 (2012).
80. Collins, R. What makes UK Biobank special? *Lancet* **379**, 1173-4 (2012).
81. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
82. Allen, N.E., Sudlow, C., Peakman, T., Collins, R. & Biobank, U.K. UK biobank data: come and get it. *Sci Transl Med* **6**, 224ed4 (2014).
83. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
84. O'Connell, J. *et al.* Haplotype estimation for biobank-scale data sets. *Nat Genet* **48**, 817-20 (2016).
85. Loh, P.R., Palamara, P.F. & Price, A.L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* **48**, 811-6 (2016).
86. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
87. Wellcome Trust Case Control, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
88. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
89. International HapMap, C. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52-8 (2010).
90. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
91. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).

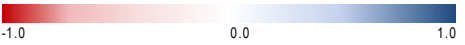
92. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
93. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* **42**, 105-16 (2010).

a.

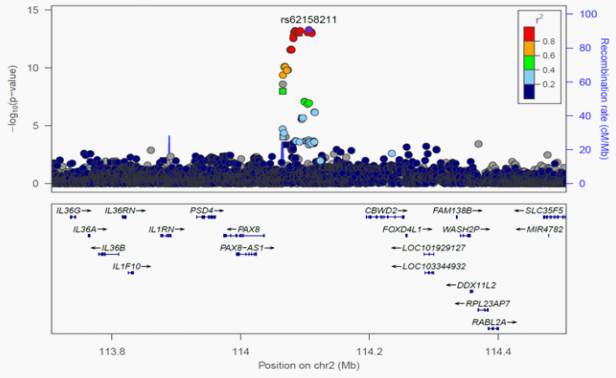
	Sleep Duration	Insomnia Symptoms	Excessive Daytime Sleepiness	Chronotype
Sleep Duration		r=-0.25	r=-0.03	r=0.03
Insomnia Symptoms	$p<0.001$		r=0.08	r=0.00
Excessive Daytime Sleepiness	$p<0.001$	$p<0.001$		r=-0.01
Chronotype	$p<0.001$	$p=0.2$	$p=3\times10^{-3}$	

b.

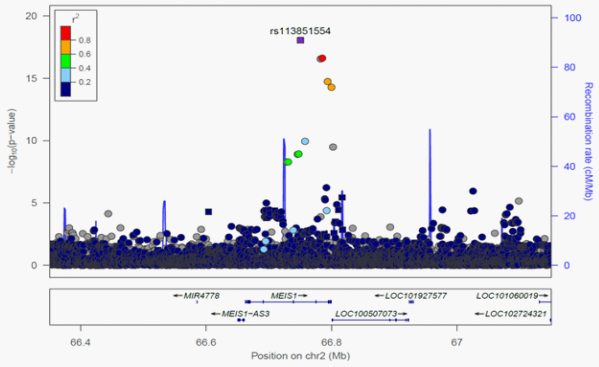
	Sleep Duration	Insomnia Symptoms	Excessive Daytime Sleepiness	Chronotype
Sleep Duration		$r_g=-0.50$	$r_g=-0.22$	$r_g=0.04$
Insomnia Symptoms	$p=6\times10^{-17}$		$r_g=0.27$	$r_g=0.02$
Excessive Daytime Sleepiness	$p=6\times10^{-4}$	$p=2\times10^{-6}$		$r_g=-0.06$
Chronotype	$p=0.6$	$p=0.7$	$p=0.2$	



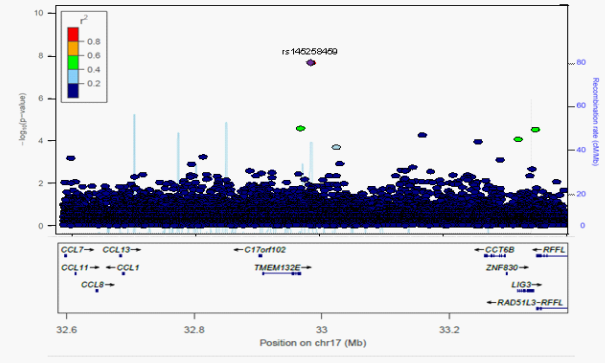
a.


Sleep Duration
rs62158211
Position on chr2 (Mb)

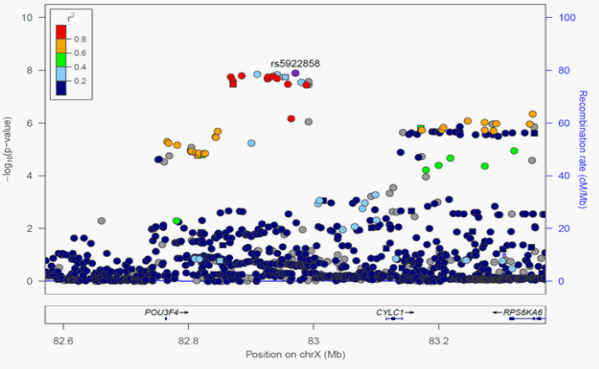
b.


Insomnia Symptoms
rs113851554
Position on chr2 (Mb)

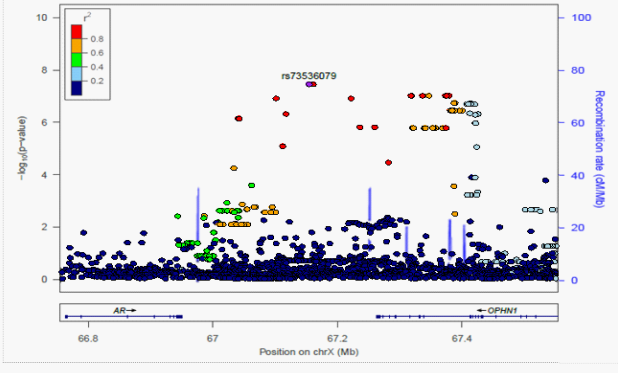
c.


Insomnia Symptoms
rs146258459
Position on chr17 (Mb)

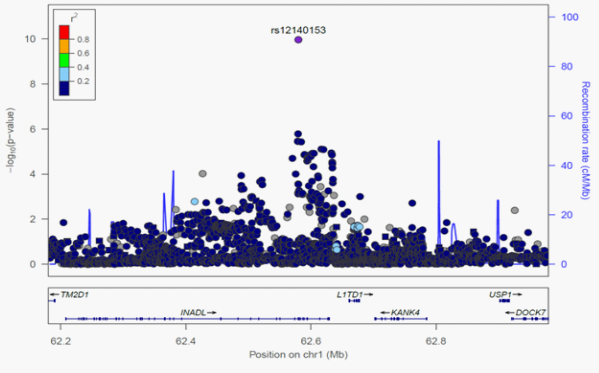
d.


Insomnia Symptoms
rs5922858
Position on chrX (Mb)

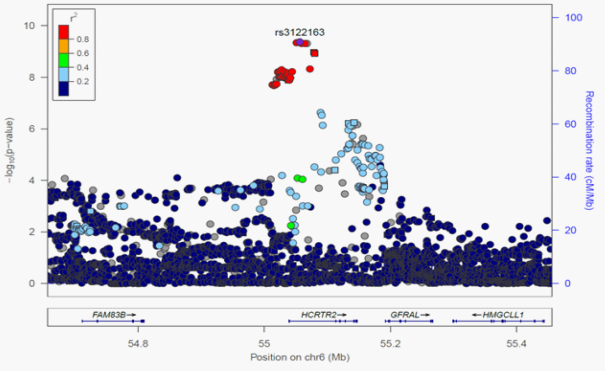
e.

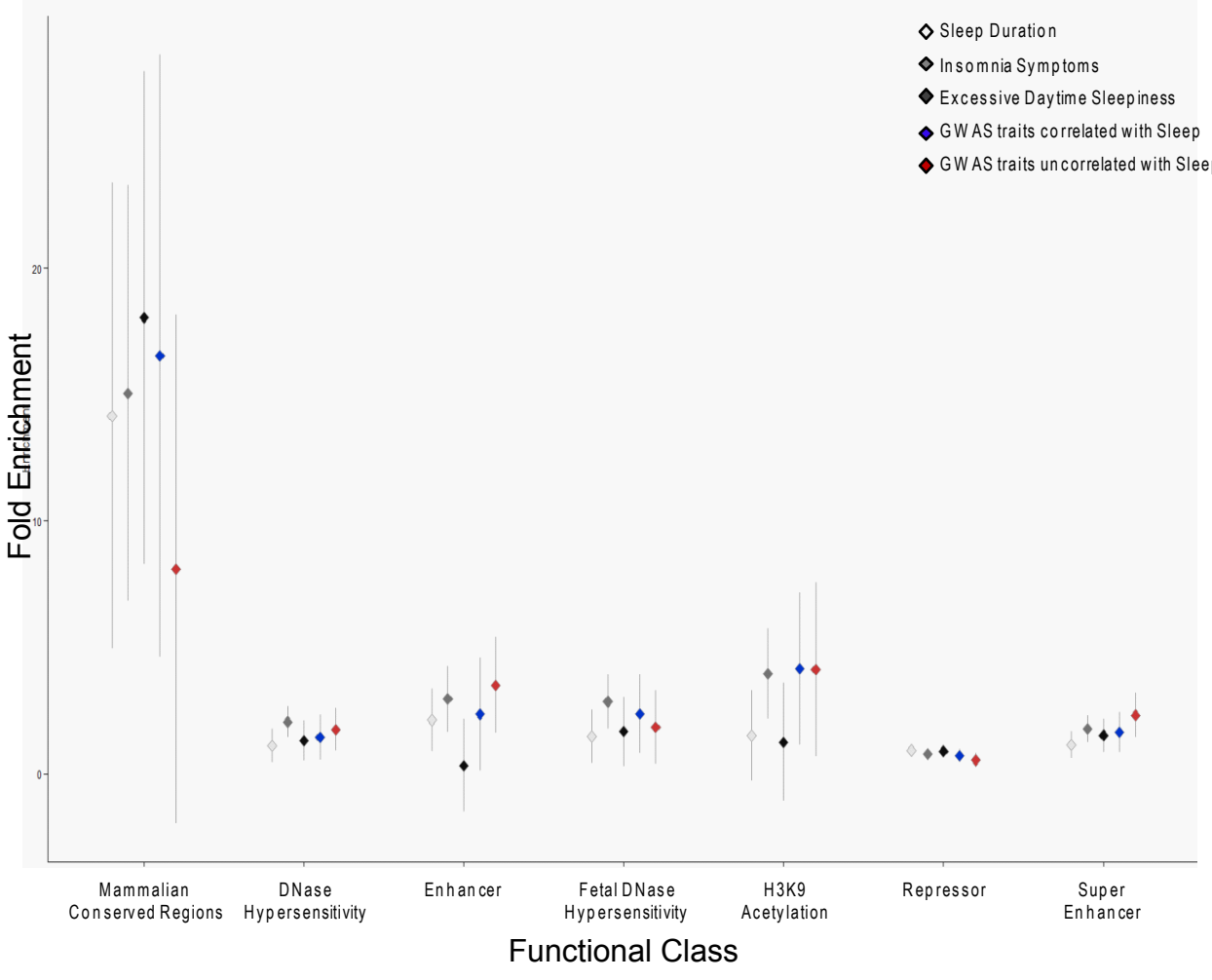

Excessive Daytime Sleepiness
rs73536079
Position on chrX (Mb)

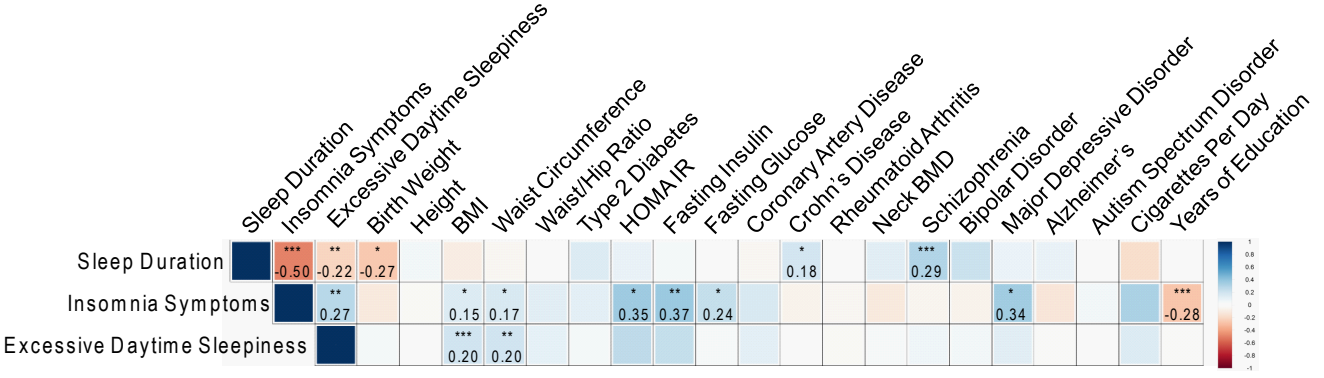
f.


Sleep Multitrait
rs12140153
Position on chr1 (Mb)

g.


Sleep Multitrait
rs3122163
Position on chr6 (Mb)


Sleep Duration
Insomnia Symptoms
Excessive Daytime Sleepiness
GWAS traits correlated with Sleep
GWAS traits uncorrelated with Sleep
Fold Enrichment
20
10
0
Mammalian Conserved Regions
DNase Hypersensitivity
Enhancer
Fetal DNase Hypersensitivity
H3K9 Acetylation
Repressor
Super Enhancer
Functional Class


Sleep Duration
Insomnia Symptoms
Excessive Daytime Sleepiness
Birth Weight
Height
BMI
Waist Circumference
Waist/Hip Ratio
Type 2 Diabetes
HOMA IR
Fasting Insulin
Fasting Glucose
Coronary Artery Disease
Crohn's Disease
Rheumatoid Arthritis
Neck BMD
Schizophrenia
Bipolar Disorder
Major Depressive Disorder
Alzheimer's
Autism Spectrum Disorder
Cigarettes Per Day
Years of Education
Sleep Duration
Insomnia Symptoms
Excessive Daytime Sleepiness
*** -0.50
** -0.22
* -0.27
* 0.18
*** 0.29
** 0.27
* 0.15
* 0.17
* 0.35
** 0.37
* 0.24
* 0.34
*** -0.28
*** 0.20
** 0.20