NEUROSCIENCE

Variant-to-gene mapping followed by cross-species genetic screening identifies GPI-anchor biosynthesis as a regulator of sleep

Justin Palermo[1]†, Alessandra Chesi[2,3]†, Amber Zimmerman[2,4]†, Shilpa Sonti[2], Matthew C. Pahl[2], Chiara Lasconi[2], Elizabeth B. Brown[1], James A. Pippin[2], Andrew D. Wells[2,3,4], Fusun Doldur-Balli[4], Diego R. Mazzotti[5,6], Allan I. Pack[4], Phillip R. Gehrman[4]*‡, Struan F.A. Grant[2,7,8,9]*‡, Alex C. Keene[1]*‡

Genome-wide association studies (GWAS) in humans have identified loci robustly associated with several heritable diseases or traits, yet little is known about the functional roles of the underlying causal variants in regulating sleep duration or quality. We applied an ATAC-seq/promoter focused Capture C strategy in human iPSC-derived neural progenitors to carry out a "variant-to-gene" mapping campaign that identified 88 candidate sleep effector genes connected to relevant GWAS signals. To functionally validate the role of the implicated effector genes in sleep regulation, we performed a neuron-specific RNA interference screen in the fruit fly, *Drosophila melanogaster*, followed by validation in zebrafish. This approach identified a number of genes that regulate sleep including a critical role for glycosylphosphatidylinositol (GPI)–anchor biosynthesis. These results provide the first physical variant-to-gene mapping of human sleep genes followed by a model organism–based prioritization, revealing a conserved role for GPI-anchor biosynthesis in sleep regulation.

INTRODUCTION

Dysregulation of sleep duration, timing, and quality are associated with substantial disease risk and public health burden (*1*, *2*). Sleep duration and quality vary markedly between individuals, suggesting the presence of complex genetic factors that distinctly regulate characteristics of sleep (*3*). Despite this recognized concern, variable sleep differences across the population have a poorly understood biological basis, particularly from the genetic standpoint (*4*).

Virtually all physiologic processes are affected by sleep, strongly suggesting that its function extends beyond the brain to affect diverse cell types and physiological processes (*4*). In recent decades, substantial progress has been made in understanding the mechanisms underlying circadian rhythms, including the identification of many genetic loci that affect interindividual variability, yet much less is known about variability in sleep disorders such as insomnia across human populations (*5*). A number of genome-wide association studies (GWAS) efforts have been conducted for insomnia-related phenotypes. Initial efforts in relatively smaller datasets ($N \leq 10{,}000$) failed to achieve genome-wide significant associations with self-reported insomnia symptoms (*6*, *7*). However, more recent studies have combined data from the U.K. Biobank and 23andMe for an insomnia GWAS of >1.3 million individuals that yielded 202 associated loci significant at the genome-wide level (*8*).

A central impediment to interpreting GWAS studies for complex traits is determining whether the nearest gene to an associated single-nucleotide polymorphism (SNP) functionally contributes to the observed phenotype (*9*, *10*). Even when the most obvious gene at the locus would appear functionally linked a priori, perhaps those genes represent a "red herring," and the actual causative gene remains to be found, or equally likely, there may be more than one effector gene at a given locus (*11*–*13*). While there is a relative paucity of public domain genomic data relevant to sleep-related tissue, such as expression quantitative trait loci (eQTL) data, related techniques can be leveraged to identify sleep-influencing effector genes. We elected to carry forward established and novel insomnia GWAS signals to such a next level of investigation. The application of chromatin conformation capture–based ultra-high resolution promoter "interactome" have the ability to determine whether chromatin "looping" contributes to human disease at key locations associated with complex traits (*14*–*18*).

Given the need for functional insight into reproducible genetic associations with sleep traits, our goal was to provide the first comprehensive physical variant-to-gene mapping for insomnia GWAS–implicated loci by taking advantage of our data generated on neural progenitor cells (NPCs). The leveraging of GWAS findings to find genetic variation that affects sleep requires first defining the effector genes affected by the key regulatory regions harboring the associated putative causal noncoding SNPs, then localizing expression to defined brain regions or cell types, and, lastly, characterization of impact on sleep duration and timing in vivo. Here, we integrated Assay for Transposase-Accessible Chromatin with high-throughput

[1]Department of Biology, Texas A&M University, College Station, TX 77843, USA. [2]Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. [3]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Division of Sleep Medicine, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Pennsylvania, PA 19104, USA. [5]Division of Medical Informatics, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS 66103, USA. [6]Division of Pulmonary Critical Care and Sleep Medicine, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS 66103, USA. [7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [8]Divisions of Human Genetics and Endocrinology and Diabetes, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. [9]Department of Genetics, University of Pennsylvania, Philadelphia, PA 19104, USA.
*Corresponding author. Email: gehrman@upenn.edu (P.R.G.); grants@chop.edu (S.F.A.G.); keenea@tamu.edu (A.C.K.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.

sequencing (ATAC-seq)/promoter-focused Capture C data with GWAS findings to implicate effector genes affected by regulatory regions coinciding with key insomnia-associated SNPs with cell-type specificity. These data provide a list of candidate sleep regulators and provide the basis for in vivo analysis of gene function in genetically amenable model systems. We first used a genome-wide RNA interference (RNAi) library in the fruit fly, *Drosophila melanogaster*, to test whether the candidate genes function in neurons to regulate sleep. These experiments were followed by CRISPR-based mutagenesis in zebrafish to determine whether the effects identified in flies are conserved in a vertebrate model. These efforts identified numerous previously unidentified regulators of sleep, including a role for phosphatidylinositol glycan (*PIG*)-*Q* and glycosylphosphatidylinositol (GPI) anchoring in sleep regulation. Furthermore, this integrative approach provides a framework for high-throughput validation of candidate genes implicated through the integration of GWAS signals with variant-to-gene mapping in a relevant human cell model followed by in vivo phenotypic analyses in animal models.

RESULTS

To search for potential regulators of human sleep, we first leveraged genome-wide significant signals from published insomnia GWAS derived from a combination of the U.K. Biobank cohort and individuals who were genotyped by 23andMe and consented to participate in research for insomnia (*8*). A total of 202 genome-wide significant loci previously implicated 956 genes through positional, eQTL, and chromatin mapping that were enriched for neural cell types. These genetic associations provided the basis for variant-to-gene mapping and functional validation of sleep genes (Fig. 1A).

While sleep affects tissues throughout the body, it is largely defined by physiological changes in brain activity that drive sleep-associated behaviors (*19*, *20*). To examine the effects of loci identified through GWAS for insomnia in brain-related cell types, we leveraged data derived from both genome-wide ATAC-seq and high-resolution promoter-focused Capture C data to implicate insomnia effector genes contacted directly by regulatory regions harboring the GWAS-associated variants. Because both neurons and glia regulate many aspects of sleep function, we focused our initial analysis on induced pluripotent stem cell (iPSC)–derived NPCs, which are the precursors from which most of the glial and neuronal cell types of the central nervous system originate (*21*–*23*). iPSCs derived from two healthy individuals [Children's Hospital of Philadelphia (CHOP)WT10 and CHOPWT14] were differentiated to NPC and cultured using standard techniques (*21*). We used a high-resolution genome-scale, promoter-focused Capture C–based approach (*14*) that uses a four-cutter restriction enzyme [Dpn II; mean fragment size, 433 base pairs (bp); median, 264 bp], achieving higher resolution than the more commonly used six-cutter Hi-C–related approaches (Hind III; mean fragment size, 3697 bp; median, 2274 bp) (*14*, *24*). We have previously reported that Hind III–based approaches lack the required resolution, and we observed biases against implicating the nearest genes (*24*).

We leveraged Capture C and ATAC-seq data generated from the same cell lines and sequenced on the Illumina platform (*21*). The ATAC-seq data were analyzed with the ENCODE ATAC-seq pipeline (https://github.com/kundajelab/atac_dnase_pipelines) with the "optimal" Irreproducible Discovery Rate (IDR) calling strategy, yielding 100,067 open chromatin peaks. We then ran a comparable linkage disequilibrium (LD) score regression between our NPC dataset and what we carried out previously for neurodevelopmental traits (*21*) with insomnia GWAS summary statistics. The insomnia GWAS was significantly enriched by 5.25-fold ($P = 0.0185$). Motivated by this observation, we sought to determine the informative genetic variants associated with insomnia; we extracted 11,348 proxy SNPs for each of 246 independent signals coinciding with 200 informative insomnia GWAS loci where proxies could be identified [coefficient of determination (r^2) > 0.7 to sentinel SNP in Europeans] and overlapped those variants with the positions of the open chromatin regions (ATAC-seq peaks). We identified 321 informative proxy SNPs corresponding to 100 of the insomnia loci in high LD with the sentinel SNP at each locus investigated. This effort substantially shortened the list of candidate causal variants from the initial GWAS discoveries, since ATAC-seq permitted us to focus on variants residing within open chromatin regions in cells that are relevant for sleep-wake regulation.

Leveraging the Capture C dataset (*21*), we mapped the informative variants from the insomnia GWAS loci to their target genes in NPCs. Of the insomnia GWAS loci investigated, 36 were implicated in a chromatin loop, with proxy SNPs residing in open chromatin (not in a baited promoter region) contacting one or more open gene promoters. A total of 135 open baited regions corresponding to the promoters of 141 genes (88 coding) were connected to 76 open chromatin regions harboring one or more insomnia proxy SNP through 148 distinct non–bait-to-bait chromatin looping interactions (table S1). Some chromatin loops pointed to the nearest gene (such as rs13033745 at *MEIS1*; Fig. 1B), while others to a gene or multiple genes further away from the candidate regulatory open SNP (such as rs9914123, which resides in an intron of *COPZ2* but loops to the promoters of several genes in a ~700-kb region; Fig. 1C). The chromatin loops involving three insomnia-associated SNPs (rs3752495, rs8062685, and rs9932282; r^2 with sentinel SNP rs3184470 = ~1) and the promoters of *PIG-Q*, *NHLRC4*, and NME4 are shown in Fig. 1D. These analyses identified 88 candidate target-coding genes, including *MEIS1*, which has already been widely implicated in sleep and restless leg syndrome (*25*, *26*). Mining our RNA sequencing (RNA-seq) data on the same cell line, we observed that almost all the identified target-coding genes (80 of 88) were expressed at moderate or high level [percentile of expression, >50%; transcripts per million (TPM) of >1.5] (table S1).

The genes and neural mechanisms regulating sleep are highly conserved from flies to mammals, and powerful genetics in non-mammalian models can be used to screen for previously unidentified regulators of sleep (*27*, *28*). In fruit flies, sleep can be identified through behavioral inactivity bouts lasting for 5 min or longer, and the *Drosophila* activity monitor (DAM) system detects activity through infrared beam crossing and is widely used to quantify sleep (*29*, *30*). To determine whether the candidate genes from our three-dimensional genomics effort contribute to sleep regulation, we expressed RNAi targeted in candidate genes selectively in neurons under control of nSyb-GAL4 and screened genes for sleep (Fig. 2A). Of the 88 insomnia-associated coding genes identified through our variant-to-gene mapping analyses, we could identify 66 genes with moderate to strong orthologs in fruit flies as defined by DIOPT, which integrates results from multiple prediction tools (*31*). Of these genes, 54 had available RNAi lines in the Vienna *Drosophila* Stock Center or the DRSC/TRiP collection

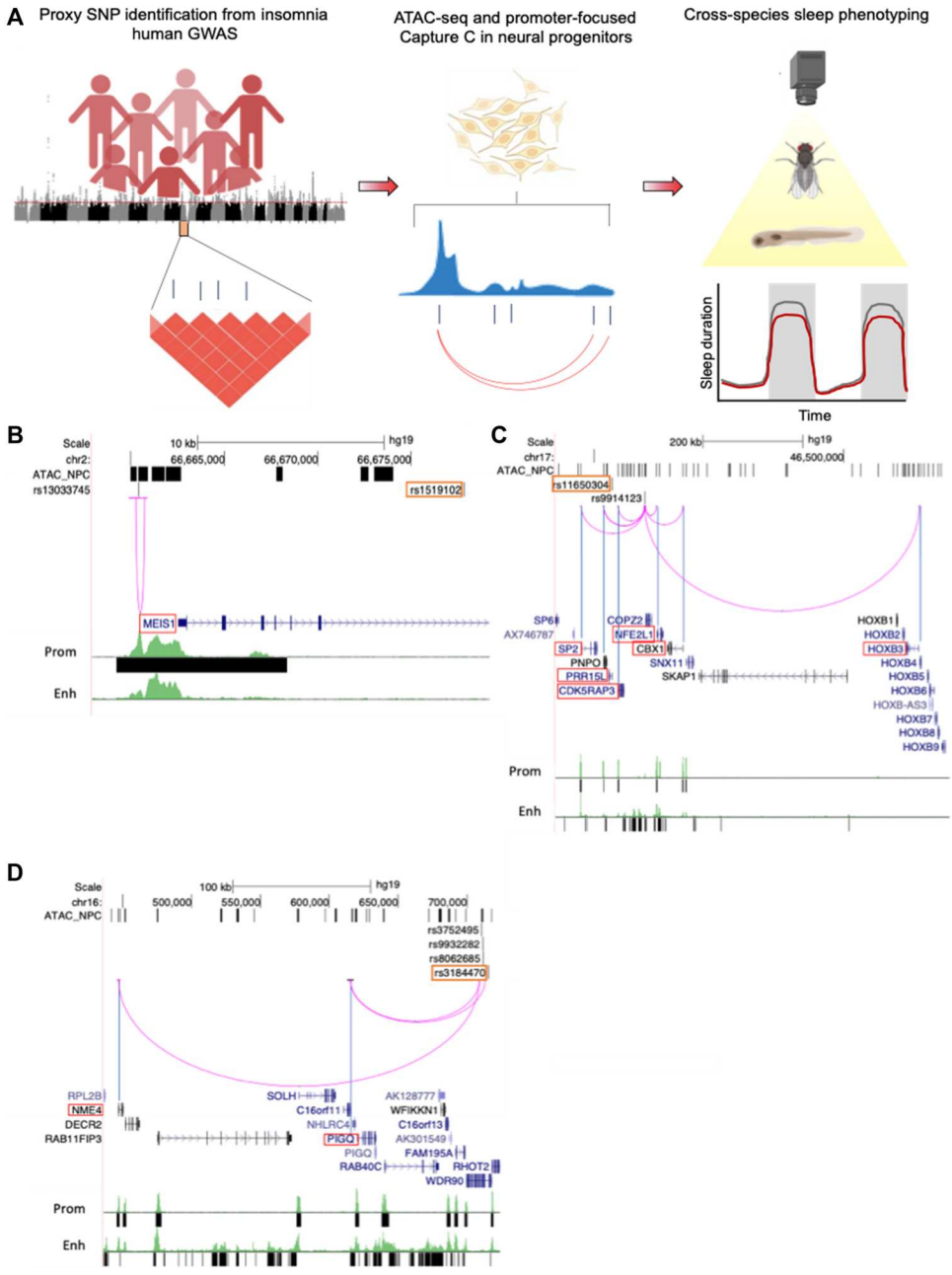

Fig. 1. Translating human GWAS signals to functional outcomes with variant-to-gene mapping. (**A**) Leveraging existing insomnia human GWAS loci, we identified proxy SNPs in strong linkage disequilibrium with sentinel SNPs using both genome-wide ATAC-seq and high-resolution promoter-focused Capture C data from iPSC-derived NPCs and then performed high-throughput sleep and activity screening using *Drosophila* RNAi lines with confirmation in a vertebrate zebrafish (*Danio rerio*) model. (**B** to **D**) Three examples of chromatin loops linking insomnia associate SNPs to candidate effector genes in NPCs. (B) rs13033745 [coefficient of determination (r^2) with sentinel SNP rs1519102 = 0.84] loops to the *MEIS1* promoter region. (C) rs9914123 (r^2 with sentinel SNP rs11650304 = 0.76) loops to the promoters of *SP2*, *PRR15L*, *CDK5RAP3*, *NFE2L1*, *CBX1*, and *HOXB3* in a ~700-kb region. (D) rs3752495, rs8062685, and rs9932282 (r^2 with sentinel SNP rs3184470 = ~1) loop to the promoters of *PIG-Q*, *NHLRC4*, and *NME4*. Orange box, sentinel SNP. Black bars, open chromatin peaks from ATAC-seq. Magenta arcs, chromatin loops from promoter-focused Capture C. Neuronal enhancer and promoter tracks are from (*81*).

(Fig. 2A). There were no significant differences in total sleep duration between control flies from each RNAi library, and therefore, all lines were tested and analyzed together (Fig. 2B). This initial analysis identified a number of short and long sleeping lines. For example, knockdown of the genes encoding the cell adhesion molecule *connectin* (ortholog of CHADL) and the basic helix-loop-helix transcription factor *daughterless* (ortholog of TCF12) resulted in short-sleeping phenotypes (Fig. 2B and table S2). In addition, we identified a short-sleeping phenotype for the Hox cofactor, *homothorax* (*hth*), and ortholog of mammalian *Meis1*, which have already been implicated in sleep and human restless leg syndrome (*26*). The screen also identified a number of genes associated with long-

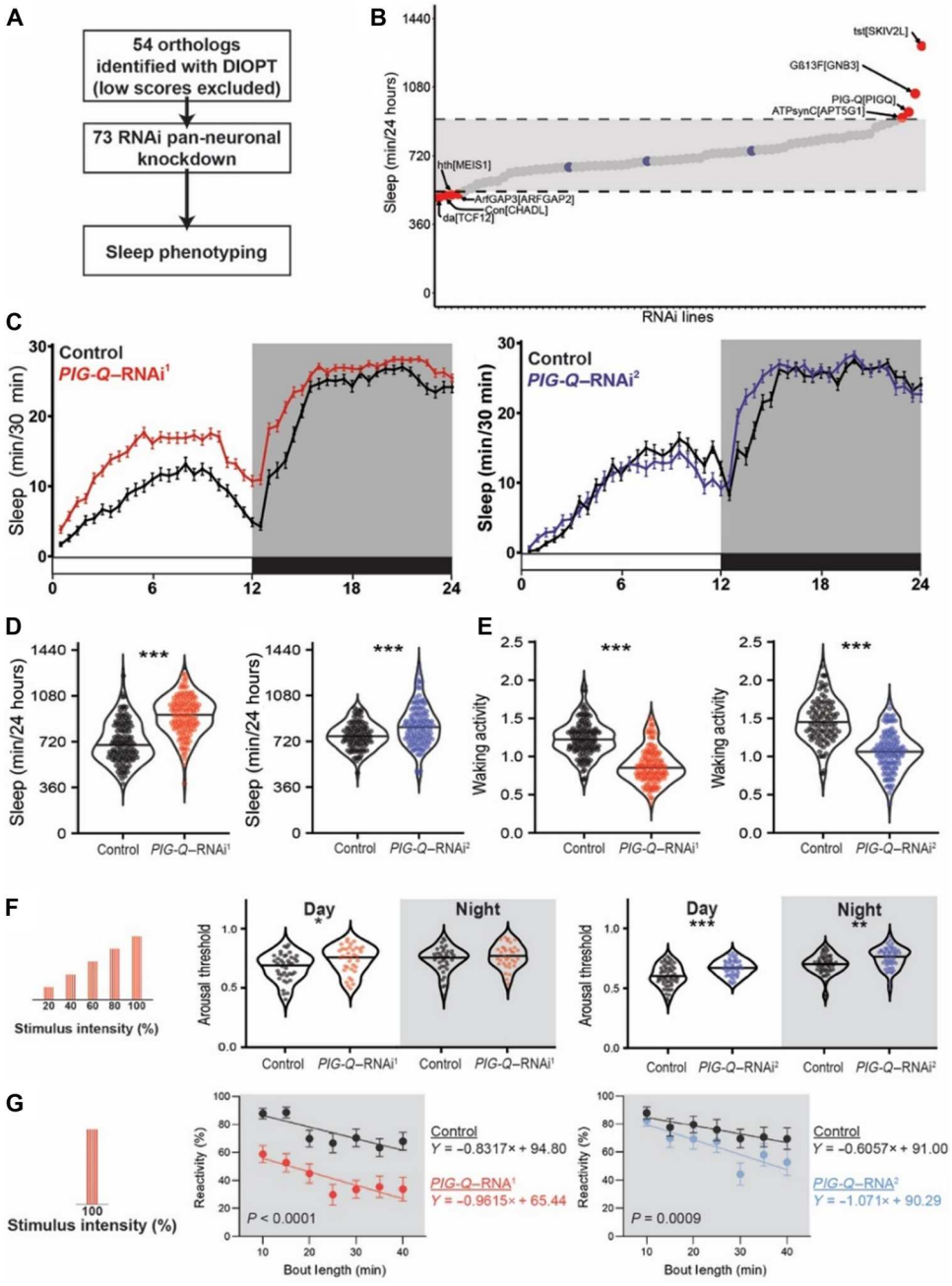

Fig. 2. *PIG-Q* knockdown increases sleep duration and sleep depth. (**A**) Design of orthologous gene screen. (**B**) Total sleep minutes over a 24-hour period in viable RNAi crosses (73 lines, $n > 16$ per line). Dashed lines and grayed area indicate two SDs from the mean for every animal tested in either direction. Blue dots indicate control sleep responses, while red dots indicate sleep responses of RNAi lines that fall outside two SDs. (**C**) Sleep profiles of two independent RNAi lines targeting *PIG-Q* (*PIG-Q*–RNAi^1, red; *PIG-Q*–RNAi^2, blue). (**D**) Knockdown of *PIG-Q* significantly increases total sleep (*t* test, *PIG-Q*–RNAi^1 $t_{260} = 11.42$, $P < 0.0001$; *PIG-Q*–RNAi^2 $t_{184} = 4.282$, $P < 0.0001$). (**E**) Knockdown of *PIG-Q* significantly decreases waking activity (*t* test: *PIG-Q*–RNAi^1 $t_{260} = 11.45$, $P < 0.0001$; *PIG-Q*–RNAi^2 $t_{184} = 11.09$, $P < 0.0001$). (**F** and **G**) The *Drosophila* arousal tracking system records fly movement while simultaneously controlling periodic mechanical stimuli. (F) Arousal threshold was measured on sleeping flies using mechanical stimuli of increasing strength. Knockdown of *PIG-Q* significantly increases arousal threshold (restricted maximum likelihood: *PIG-Q*–RNAi^1 $F_{1,73} = 4.267$, $P = 0.0424$; *PIG-Q*–RNAi^2 $F_{1,102} = 16.42$, $P < 0.0001$). This occurs during the day for both independent RNAi lines (*PIG-Q*–RNAi^1, $P = 0.0127$; *PIG-Q*–RNAi^2, $P = 0.0002$), while an increase in arousal threshold only occurred in one line during the night (*PIG-Q*–RNAi^1, $P = 0.4308$; *PIG-Q*–RNAi^2, $P = 0.0020$). (G) Reactivity was measured by assessing the proportion of flies that react to a single mechanical stimulus for each bin of immobility. Knockdown of *PIG-Q* significantly decreases nighttime reactivity (analysis of covariance with bout length as covariate: *PIG-Q*–RNAi^1, $F_{1,661} = 107.1$, $P < 0.0001$; *PIG-Q*–RNAi^2, $F_{1,594} = 24.87$, $P < 0.0001$). For sleep profiles, error bars represent ± SEM. For violin plots, the median (solid black line) is shown. White background indicates daytime, while gray background indicates nighttime. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

sleeping phenotypes including *Gß13F* (ortholog of GNB3), the RNA helicase *twister* (ortholog of SKIV2L), and the GPI-anchoring biosynthesis protein, *PIG-Q* (*32*). *PIG-Q* encodes an enzyme *N*-acetylglucosaminyl transferase that localizes to the endoplasmic reticulum and is required for synthesis of the GPI anchor that regulates the cellular localization of ~150 proteins. Together, these findings revealed complex sleep phenotypes associated with individual genes identified through human GWAS studies. Given the high conservation of *PIG-Q* and its critical role in GPI biosynthesis, we chose to focus on this gene for further analyses.

To validate the screening results, we repeated the sleep analyses using additional genetic controls including a second, independently derived RNAi line. Flies with pan-neuronal *PIG-Q* were compared to controls harboring the GAL4 driver or the RNAi line alone. Both RNAi lines significantly increased sleep over control flies, fortifying the notion that loss of *PIG-Q* in neurons promotes sleep during the daytime and nighttime (Fig. 2. C and D, and fig. S1). Analysis of each phase of sleep revealed that while knockdown of *PIG-Q*–RNAi^1 significantly increases sleep both during the day and night phases (day: $t_{260} = 10.49$, $P < 0.0001$; night: $t_{260} = 8.581$, $P < 0.0001$), *PIG-Q*–RNAi^2 is not significantly different during the day phase ($t_{184} = 0.2797$, $P = 0.78$); however, knockdown of *PIG-Q*–RNAi^2 did significantly increase sleep at night ($t_{184} = 8.853$, $P < 0.0001$). A direct comparison between the two RNAi lines revealed significant differences in effectiveness between the two RNAi lines with *PIG-Q*–RNAi^1 increasing sleep duration to a greater degree than *PIG-Q*–RNAi^2 [analysis of variance (ANOVA), *PIG-Q*–RNAi^1 versus *PIG-Q*–RNAi^2, $P < 0.0001$], and no significant difference was detected between the controls ($P = 0.50$).

The waking activity, defined as the average amount of activity while the animal is awake, was reduced in *PIG-Q* knockdown

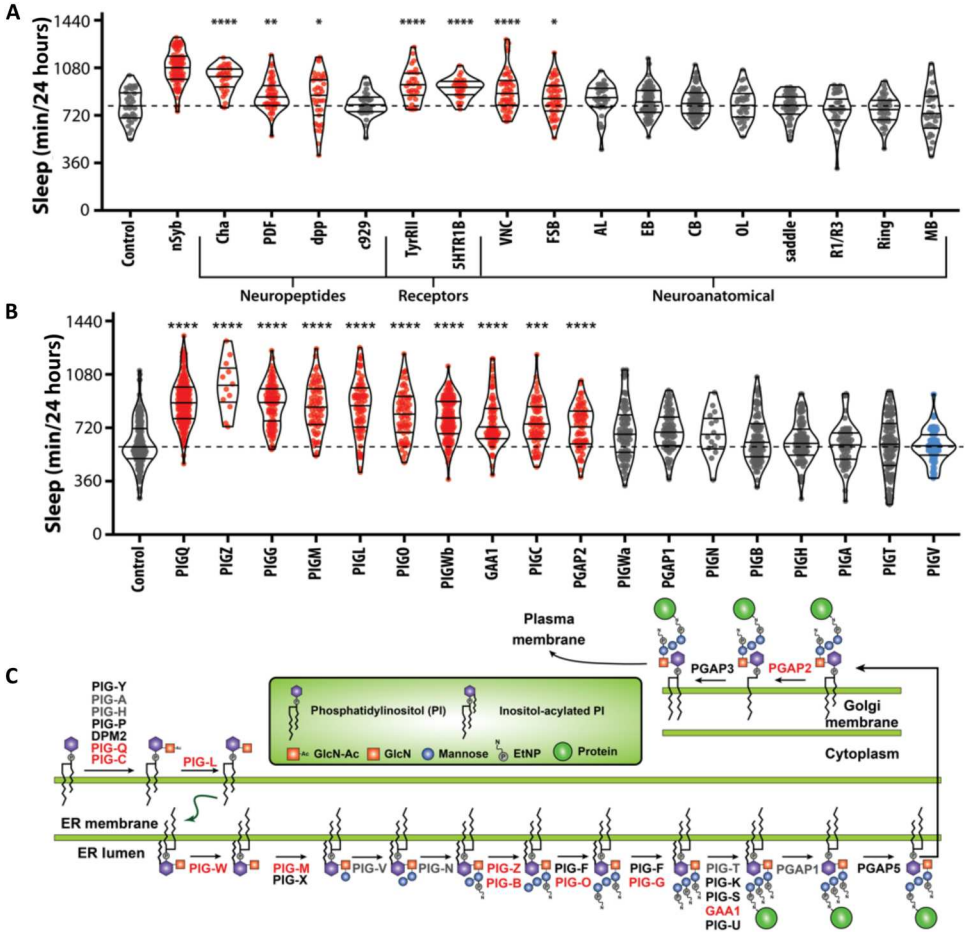

Fig. 3. Localization of *PIG-Q* and characterization of the GPI-anchor biosynthesis genes in sleep regulation. (**A**) Knockdown of *PIG-Q* in multiple *Drosophila* neuronal subpopulations affects sleep duration (ANOVA: $F_{18,756} = 25.21$, $P < 0.0001$). The dashed line represents the mean of the control line. *PIG-Q* significant increased sleep when knocked down pan-neuronally (*nSyb*; $P < 0.0001$); cholinergic neurons (*Cha*; $P < 0.0001$); dpp-expressing neurons (*dpp*; $P = 0.0334$); circadian pacemaker neurons (*PDF*; $P = 0.0025$); tyramine II receptor neurons (*TyRII*; $P < 0.0001$); serotonin receptor IB neurons (*5HT1b*; $P < 0.0001$); the ventral nerve cord (*VNC*; $P < 0.0001$); and the fan-shaped body (*fsb*; $P = 0.0478$). (**B**) Knockdown of multiple genes in the PIG pathway affects sleep duration (ANOVA: $F_{18,1251} = 39.63$, $P < 0.0001$). The dashed line indicates the mean of the control line. For violin plots, the median and 25th and 75th percentiles are shown (solid black lines). Each dot represents an individual fly; red indicates sleep duration that is significantly higher than the control, gray indicates sleep that is not significantly different, and blue indicates sleep that is significantly lower than the control as revealed by Dunnett's multiple comparisons test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$. (**C**) PIG anchor biosynthesis pathway. Genes highlighted in red represent genes that show long sleep phenotypes when knocked down pan-neuronally in *Drosophila* as described below, while genes in gray exhibited no or short sleep phenotype. Genes in black were untested because there were no available RNAi lines.

flies, suggesting a role in activity, in addition to sleep regulation (Fig. 2E). Furthermore, the identified phenotypes were present in male flies, revealing that the effect of *PIG-Q* knockdown on sleep is not sex specific (fig. S2). Together, these results confirmed that knockdown of *PIG-Q* in neurons promotes sleep.

Across phyla, sleep is defined by a homeostatic rebound following deprivation and reduced responsiveness to external stimuli. To determine whether sleep homeostasis is disrupted in *PIG-Q* knockdown flies, we sleep-deprived flies for 12 hours during the lights off period [zeitgeber time (ZT)12 to 24] by mechanical stimulation and measured recovery sleep. Following deprivation, both control and *PIG-Q* knockdown flies significantly increased sleep. A direct comparison revealed a similar percent increase in flies, suggesting that homeostatic rebound is intact in *PIG-Q*–deficient flies (fig. S3). To further investigate the role of *PIG-Q* in sleep regulation, we quantified arousal threshold in knockdown flies using the *Drosophila* Arousal Threshold (DART) system (*33*, *34*). Analysis of video recordings in this system confirmed the increased sleep phenotype of *PIG-Q* knockdown flies from infrared tracking (fig. S4). To probe for sleep depth, sleep was recorded and analyzed by video tracking before and following exposure to mechanical shaking that increase in strength. There was an increase in daytime arousal threshold in flies with pan-neuronal expression of either *PIG-Q* RNAi line, suggesting that loss of *PIG-Q* increases sleep depth (Fig. 2F). Nighttime arousal threshold was increased in flies expressing *PIG-Q*–RNAi^{2}, but not *PIG-Q*–RNAi^{1}, possibly because of a ceiling effect. Nighttime reactivity, a second measure of arousal threshold that takes into account the amount of time an animal has been asleep, was reduced in both lines, suggesting a role for *PIG-Q* in sleep depth (Fig. 2G). No differences in reactivity were identified during the daytime in *PIG-Q* knockdown flies (fig. S5). These results strengthen the finding that knockdown of *PIG-Q* expression increases sleep depth, particularly during longer sleep bouts.

In *Drosophila* and mammals, sleep-regulating neurons are found in numerous brain regions (*22*, *27*). To localize the effects of *PIG-Q*, we selectively knocked down function in different populations of neurons within the brain and measured the effects on sleep. There was no effect of knockdown in a number of canonical sleep areas including the mushroom body (R69F08) and the c929 driver that labels numerous sleep-regulating peptidergic neurons (*35*). Therefore, *PIG-Q* is unlikely to generally affect cellular function within sleep-regulating circuits. However, knockdown in cholinergic neurons within the brain (Cha) significantly increased sleep, phenocopying pan-neuronal knockdown (nSyb) (Fig. 3A), suggesting that *PIG-Q* modulates the function of these excitatory neurons. We also found increased sleep when *PIG-Q* was knocked down in a number of neuronal types including the circadian pacemaker pigment-dispersing factor (PDF) neurons, serotonin receptor (5-HTR) neurons, the ellipsoid body, and fan-shaped body (Fig. 3A and table S3). Two drivers that label the small ventral lateral neurons (sLNvs) pacemakers cells (PDF-GAL4 and dpp-GAL4 that label PDF-expressing neurons) both increase sleep (*35*–*37*). These findings suggest that *PIG-Q* is required in diverse subsets of sleep-regulating neurons for normal sleep. Therefore, *PIG-Q* is likely to function in multiple subsets of neuromodulatory circuits to regulate sleep.

PIG-Q functions in the GPI biosynthesis pathway that is highly conserved and critical for the function of GPI-anchored proteins (*37*). Given the role of GPI-anchored proteins in sleep regulation (*38*), we sought to determine whether additional components of this pathway are involved in sleep regulation. We knocked down 18 genes individually in different experiments in the GPI-biosynthesis pathway pan-neuronally and measured the effect on sleep. Sleep was significantly increased in flies with loss of PIG-Z, PIG-L, PIG-O, PIG-C, PIG-G, and PIG-M, where all slept longer than control flies expressing the RNAi line alone or the nsyb-GAL4 driver alone (Fig. 3, B and C, and table S4). Most genes targeted for pan-neuronal knockdown in this pathway resulted in increased sleep. Together, these findings suggest that generalized disruption of PIG-mediated GPI biosynthesis promotes sleep.

PIG-Q is a conserved gene across species with 44% amino acid sequence similarity between *Drosophila* and humans (*39*). Conservation is higher among vertebrates with a sequence similarity of 77% between zebrafish and humans (*39*). To determine whether the functional effects of sleep in *Drosophila* are conserved in vertebrates, we examined the role of *PIG-Q* on sleep in zebrafish, a leading vertebrate model of sleep (*40*). We disrupted *PIG-Q* expression using CRISPR-Cas9 gene editing in zebrafish (*Danio rerio*). Targeted biallelic genetic mutations producing high-efficiency knockouts (KOs) were generated in F0 larvae (Fig. 4, A to C) (*41*). We selectively targeted exon 7, which is conserved across both zebrafish *PIG-Q* transcripts and is a highly conserved region across species (Fig. 4A) (*42*). Exon 7 is also part of the *N*-acetylglucosaminyl transferase component, a major functional component of the PIG-Q protein (*43*). Five days after fertilization, *PIG-Q* KO larvae were screened for sleep phenotypes compared to control zebrafish [scrambled guide RNA (gRNA)–injected] larvae (Fig. 4A). Behavioral analyses were performed using standardized methodology in zebrafish that has been previously used for genetic and pharmacological screens (*44*, *45*). The fish were genotyped immediately following behavioral analysis, which confirmed a mutation efficiency criterion for inclusion of >90% (Fig. 4, B and C). As with *Drosophila*, loss of *PIG-Q* significantly increased sleep duration during the night ($P < 0.001$) compared to scrambled gRNA-injected controls (Fig. 4, H and I). Daytime sleep duration was also significantly ($P < 0.05$) increased compared to controls (Fig. 4, E and H) with an increase in sleep bout number ($P < 0.05$) (Fig. 4G). The sleep differences at night were due to increased sleep bout length (Fig. 4K) rather than sleep bout number (Fig. 4J). This further supports the notion that loss of *PIG-Q* function increases sleep consolidation at night. However, there was not a significant change in overall activity during the day (Fig. 4, D and F), indicating that in zebrafish, *PIG-Q* exerts its effects primarily on sleep regulation rather than locomotion. Together, these findings confirm that loss of *PIG-Q* increases sleep across phylogeny.

DISCUSSION

We have conducted the first physical variant-to-gene mapping for insomnia GWAS by identifying putative causal variants and their associated effector genes leveraging data from an ATAC-seq/chromatin conformation capture–based approach, followed by assessing functional effects on sleep/wake regulation in *Drosophila* and zebrafish. The detailed behavioral platforms to characterize sleep duration and intensity, availability of RNAi libraries that allow for genome-wide in vivo analysis of sleep function, and high-throughput assays make *Drosophila* an excellent system for validating the role of putative regulators of sleep (*22*, *28*). The candidate genes

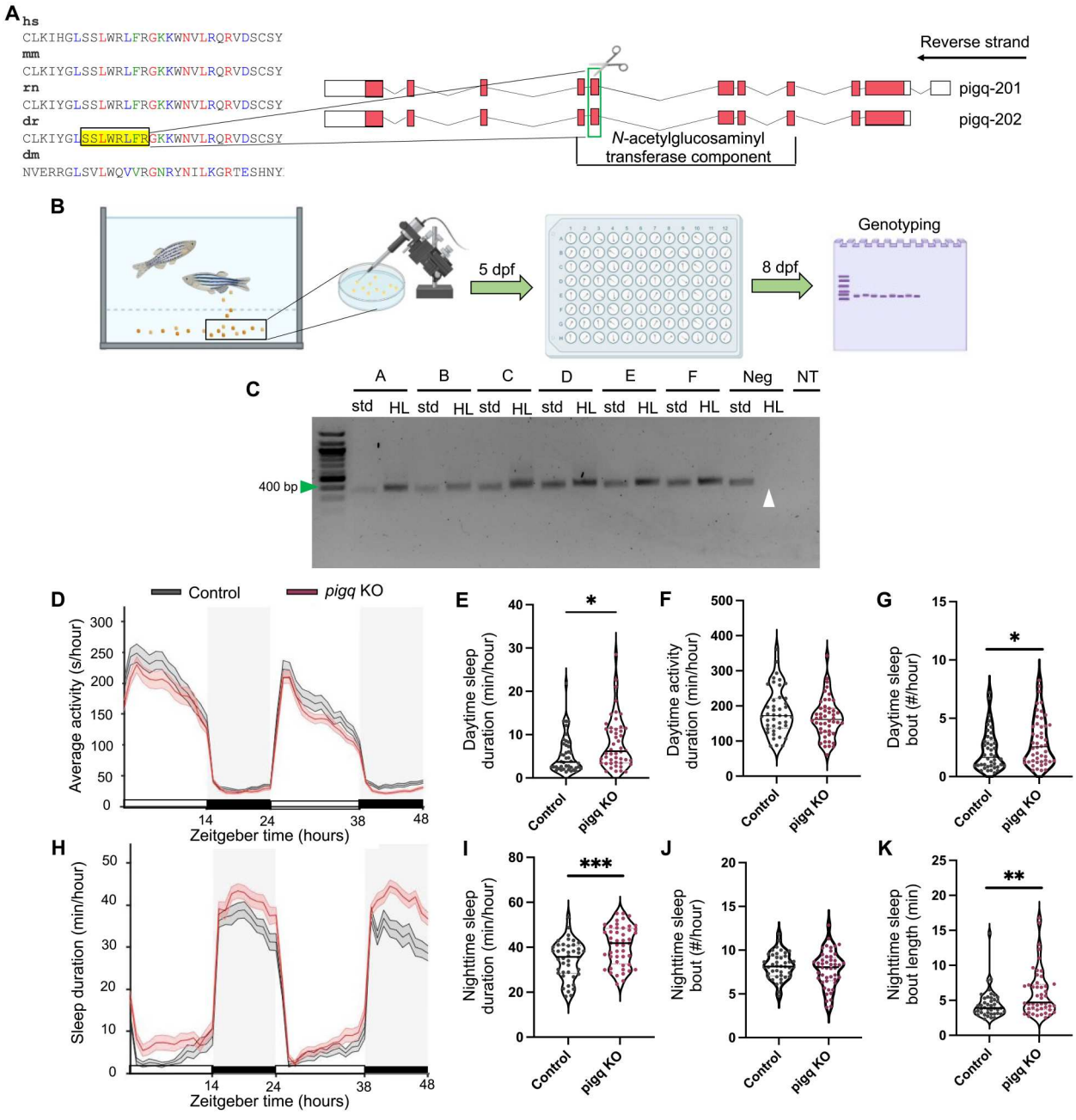

Fig. 4. CRISPR mutation of *PIG-Q* in zebrafish increases sleep. (**A**) CRISPR sgRNA design. (**B**) Schematic of embryo injection and CRISPR mutation confirmation. (**C**) Representative gel used for genotyping. Green arrow indicates 400 bp on the ladder. Expected PCR product was 366 bp. White arrow indicates wild-type DNA suppression using HL PCR as a negative control. (**D**) Average (± SEM) activity for 48 hours beginning at lights on (9:00 a.m.). (**E**) Cumulative daytime sleep across both light periods was increased in *PIG-Q* KOs (mean difference: 2.83 ± 1.09, $t_{85.2} = 2.59$, $P = 0.04$). (**F**) No difference was found in daytime activity (mean difference: −19.3 ± 13.01, $t_{88} = 1.48$, $P = 0.14$). (**G**) Daytime sleep bout number was increased in *PIG-Q* KOs (mean difference: 0.83 ± 0.40, $t_{88} = 2.06$, $P = 0.04$). (**H**) Average (± SEM) sleep duration across 48 hours beginning at lights on (9:00 a.m.). (**I**) Cumulative nighttime sleep duration was increased in *PIG-Q* KOs across both dark periods (11:00 p.m. to 9:00 a.m.) (mean difference: 6.38 ± 1.82, $t_{88} = 3.5$, $P = 0.0007$). (**J** and **K**) Nighttime sleep bout number did not differ between groups [(J) mean difference: −0.41 ± 0.38, $t_{85.2} = 1.08$, $P = 0.28$], but nighttime sleep bout length was increased in *PIG-Q* KOs [(K) mean difference: 1.60 ± 0.58, $t_{79.7} = 2.75$, $P = 0.007$]. Gray boxes indicate night, while white represents day. $N = 42$ scramble-injected controls, $N = 48$ *PIG-Q* KOs. Independent Student's *t* test was used to compare *PIG-Q* KOs and controls. Welch's correction was applied to (E), (J), and (K) because unequal variances between groups were determined. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. For (C), samples A to F represent individual larvae samples. Neg, wild-type negative control DNA; std., standard forward and reverse primers; HL, headloop primers; NT, no template; dpf, days post fertilization.

derived from our analysis were subjected to a neuron-specific RNAi screen in *D. melanogaster* followed by in-depth sleep phenotyping in *Drosophila* and zebrafish to assess the impact of such genetic perturbation on sleep/wake regulation. As a consequence, a number of short- and long-sleeping lines were identified. Therefore, this approach provides proof of principle for the use of genetic models to interrogate the functional roles of genes implicated through human GWAS studies on complex behavior.

Screening identified multiple genes with short- or long-sleep phenotypes including numerous transcription factors. These genes provide candidates for further validation in flies, including verifying phenotypes in classic genetic mutants and localizing the effects of the genes. We focused functional validation on the *PIG-Q* gene because of the robustness of the phenotype and a previously identified role for GPI-anchored genes in sleep regulation (*38*). Knockdown of *PIG-Q* significantly increased both daytime and nighttime sleep. Restricting knockdown to a number of different neuromodulatory neurons, including cholinergic and tyraminergic neurons that have previously been implicated in sleep (*46*, *47*) phenocopying pan-neuronal knockdown, supports the notion that *PIG-Q* modulates sleep through its effect in these neuronal groups. To examine how *PIG-Q* regulates sleep, we examined sleep and circadian regulation across a number of contexts. We subsequently localized *PIG-Q* function to numerous populations of neurons including cholinergic neurons, tyraminergic neurons, and neurons expressing the serotonin receptor 5HT1B, revealing that *PIG-Q* is likely to act in broad classes of neurons to regulate sleep. A number of neuronal populations where *PIG-Q* modulates sleep have been previously implicated in sleep such as cholinergic neurons, PDF-expressing neurons, and neurons of the fan-shaped body (*48*, *49*). Because many of these neurons suppress sleep, it is possible that *PIG-Q* functions to reduce neuronal activity within defined populations of neurons. Future work examining the effects of *PIG-Q* on neural activity and cellular function within defined classes of sleep-regulated neurons will be critical to understand its sleep-promoting effects.

The sleep phenotype was subsequently recapitulated in the zebrafish model, demonstrating conservation of *PIG-Q* function in regard to sleep function. The identification of *PIG-Q* implicates GPI-linked proteins in sleep regulation. Mutations in the GPI-anchored cell surface protein *sleepless* leads to robust reductions in sleep (*38*, *50*). In total, the *Drosophila* genome encodes ~150 GPI-linked proteins (*37*, *51*), and systematically testing the role of these in sleep regulation may uncover genes that are downstream of the GPI biosynthesis pathway and a broader role for GPI anchoring in sleep regulation.

In line with our similar work in other traits (*14*, *17*, *18*, *21*), we applied a physical variant-to-gene mapping approach to identify candidate regulators of sleep using loci derived from GWAS studies. While a number of studies have used human GWAS to develop candidate regulators of sleep that can be used for genetic screening, this approach may lead to the incorrect genes being implicated. For example, GWAS efforts by others for obesity have shown a pronounced association with variation within the *FTO* gene that associate with obesity (*52*). This robust association signal resides within an intronic region of this gene (*52*) and has gone on to be widely replicated in other ethnicities (*53*–*55*) plus children (*56*). Although many publications have now studied the role of the *FTO* locus in the context of obesity, a number of studies demonstrated that *FTO* is, in fact, likely not the principal causal effector gene for obesity at this locus, but rather it is *IRX3* and *IRX5* (*11*–*13*), suggesting that the genetic variant resides in an enhancer embedded in one gene that influences the expression of others. Hence, despite a great deal of data implicating *FTO* as the gene involved in obesity, in fact through refined methodologies (similar to what we propose here) in the absence of eQTL support, other genes that are physically located near *FTO* are actually the physiologically relevant effector genes (*11*–*13*). Similarly, the insomnia-associated candidate regulatory variants at GWAS locus number 170 identified by our variant-to-gene mapping reside in an intron of *WDR90* but loop across ~90 kb to the promoter region of two candidate effector genes farther away, *NHLRC4* and *PIG-Q*, and to the *NME4* promoter 258 kb away. While further experiments, such as CRISPR-Cas9 editing of the candidate variants in human cells, are required to validate a regulatory role on these target genes, our *Drosophila* phenotypic screen identified *PIG-Q* as the likely culprit gene at this locus.

Large-scale genetic screens have been applied in a number of animal models including *Caenorhabditis elegans*, *Drosophila*, zebrafish, and mice to identify genetic regulators of sleep (*38*, *44*, *57*–*60*). Despite the ability for high-throughput behavioral screening, unexpectedly few studies have used these models to validate genes identified in human GWAS studies. In fruit flies, the voltage-gated Ca^{2+} channel cacophony (*61*) and the adenosine triphosphate–sensitive potassium channel, ABCC9 (*62*), have been identified as sleep regulators following their identification in human GWAS. In addition, cross-species analysis has found that the epidermal growth factor receptor promotes sleep in *C. elegans*, *Drosophila*, and zebrafish and is associated with variation in human sleep duration and quality (*63*–*65*). Similar approaches used in NPCs can be applied to broader cell types including glia, insulin-regulating cells, and the fat body (adipose tissue), all of which have been found to be regulators of sleep (*22*, *27*). Therefore, the application of model organisms combined with variant-to-gene mapping has potential to identify genetic regulators for many traits that have been studied using GWAS.

Together, this study provides a proof-of-principle application of physical gene-variant mapping and screen-based in vivo validation for a complex behavior. This approach identified PIG family proteins as conserved regulators of sleep and raises the possibility that differences in GPI biosynthesis contribute to naturally occurring variation in sleep. This study also provides a framework for interrogation of the large number of results emerging from other GWAS of sleep and circadian phenotypes (*66*–*68*). While the number of candidate loci has surged in recent years, these results have not yet been translated to biological insights into sleep/wake regulation or the pathophysiology of sleep and circadian disorders. These GWAS data have also demonstrated significant pleiotropy, with loci associated with both sleep phenotypes and mental health traits in particular. Identifying causal genes for sleep-related traits may thus also yield insights into the genetic architecture of psychiatric disorders. With the growth of biobanks in multiple health systems, there are unique opportunities to identify individuals with common and rare genetics variants in *PIG-Q* and other causal genes for in-depth sleep and psychiatric phenotyping to improve our understanding of the potential functional effects of these genes in humans.

METHODS

GWAS of insomnia

Summary statistics from an insomnia GWAS meta-analysis–published combined sample size of 1,331,010 participants were used to identify our initial pool of candidate variants (*8*). The meta-analysis identified ~12,000 genome-wide significant variants ($P < 5 \times 10^{-8}$) located in 202 genomic risk loci in the U.K. Biobank (*8*).

Cell culture

Frozen NPCs derived from iPSC from two healthy individuals (WT10 and WT14) were obtained from CHOP stem cell core and thawed slowly in 37°C water bath. The thawed cells were gently washed in neuronal expansion media: 49% Neurobasal medium (Thermo Fisher Scientific, catalog no. 21103049), 49% advanced Dulbecco′s modified Eagle′s medium/nutrient mixture F-12 (Thermo Fisher Scientific, catalog no. 12634010), and 2% 50× neuronal induction supplement in a 15-ml conical tube, followed by centrifuging at 300*g* for 5 min. Cells were resuspended in 1 ml of prewarmed neuronal expansion media with Rock inhibitor (Y-27632 compound; STEMCELL Technologies, catalog no. 72304) at a final concentration of 10 μM, and a cell count was performed. NPCs were seeded at a density of 150,000 cells/cm^2 onto human embryonic stem cell (hESC)-qualified Matrigel-coated plates (Corning, catalog no. 354277) in neuronal expansion media (2.5 ml per well) and cultured at 37°C in a humidified cell culture incubator with 5% CO_2. The day after, the medium was changed to remove the Y-27632 compound. NPCs were expanded for 6 to 7 days in 2.5 ml of neuronal expansion media exchanged every 48 hours before harvesting.

ATAC-seq library preparation

Five technical replicates of two iPSC-derived NPC lines (CHOPWT10 and CHOPWT14) were harvested using Accutase, followed by a Dulbecco′s phosphate-buffered saline wash, and then counted. Cells (50,000) of each sample were spun down at 550*g* for 5 min at 4°C. The cell pellet was then resuspended in 50 μl of cold lysis buffer [10 mM tris-HCl (pH 7.4), 10 mM NaCl, 3 mM MgCl2, and 0.1% IGEPAL CA-630] and spun down immediately at 550*g* for 10 min at 4°C. The nuclei were resuspended on ice in the transposition reaction mix (2× TD buffer, 2.5 μl of Tn5 transposes, and nuclease-free H_2O) (Illumina, catalog no. FC-121-1030, Nextera) on ice, and the transposition reaction was incubated at 37°C for 45 min. The transposed DNA was then purified using a MinElute kit (Qiagen) adjusted to 10.5 μl of elution buffer. The transposed DNA was converted into libraries using NEBNext High-Fidelity 2× PCR Master Mix (NEB) and the Nextera Index Kit (Illumina) by polymerase chain reaction (PCR) amplification for 12 cycles. The PCR reaction was subsequently cleaned up using AMPureXP beads (Agencourt), checked on a Bioanalyzer 2100 (Agilent) high-sensitivity DNA chip (Aglient), and paired-end–sequenced on the Illumina NovaSeq 6000 platform (51-bp read length) at the Center for Spatial and Functional Genomics at CHOP.

RNA-seq library preparation

RNA was isolated from two iPSC-derived NPC lines (CHOPWT10 and CHOPWT14) in technical triplicates using TRIzol reagent (Invitrogen). RNA was then purified using the Directzol RNA Miniprep Kit (Zymol) and depleted of contaminating genomic DNA using deoxyribonuclease I. Purified RNA was then checked for quality on the Bioanalyzer 2100 using the Nano RNA chip, and samples with an RNA integrity number above 7 were used for RNA-seq library synthesis. RNA samples were depleted of ribosomal RNA (rRNA) using the QIAseq FastSelect RNA Removal Kit and then processed into libraries using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB) according to manufacturer′s instructions. The quality and quantity of the libraries were measured using the Bioanalyzer 2100 DNA 1000 chip and Qubit fluorometer (Life Technologies). Completed libraries were pooled and sequenced on the NovaSeq 6000 platform using paired-end 51-bp reads at the Center for Spatial and Functional Genomics at CHOP.

Promoter focused Capture C library preparation

We used standard methods for generation of 3C libraries (*14*). For each library, 10^7 fixed cells were thawed at room temperature, followed by centrifugation at room temperature for 5 min at 14,000 rpm. The cell pellet was resuspended in 1 ml of distilled H_2O (dH_2O) supplemented with 5 μl of 200× protease inhibitor cocktail, incubated on ice for 10 min, and then centrifuged. The cell pellet was resuspended to a total volume of 650 μl in dH_2O. Cell suspension (50 μl) was set aside for predigestion quality control (QC), and the remaining sample was divided into six tubes. Both predigestion controls and samples underwent a predigestion incubation in a Thermomixer (BenchMark) with the addition of 0.3% SDS, 1× NEB Dpn II restriction buffer, and dH_2O for 1 hour at 37°C with shaking at 1000 rpm. A 1.7% solution of Triton X-100 was added to each tube, and shaking was continued for another hour. After the predigestion incubation, 10 μl of Dpn II (50 U/μl; NEB) was added to each sample tube only, and shaking was continued along with predigestion control until the end of the day. An additional 10 μl of Dpn II was added to each digestion reaction and digested overnight. The next day, a further 10 μl of Dpn II was added, and shaking was continued for another 2 to 3 hours. A total of 100 μl of each digestion reaction was then removed, pooled into two 1.5-ml tube, and set aside for digestion efficiency QC. The remaining samples were heat-inactivated at 1000 rpm in a MultiTherm for 20 min, at 65°C, and cooled on ice for 20 min. Digested samples were ligated with 8 μl of T4 DNA ligase (HC Thermo Fisher Scientific, 30 U/μl) and 1× ligase buffer at 1000 rpm overnight at 16°C in a MultiTherm. The next day, an additional 2 μl of T4 DNA ligase was spiked into each sample and incubated for another few hours. The ligated samples were then decross-linked overnight at 65°C with proteinase K (20 mg/ml; Denville Scientific) along with predigestion and digestion control. The following morning, both controls and ligated samples were incubated for 30 min at 37°C with ribonuclease A (Millipore), followed by phenol/chloroform extraction and ethanol precipitation at −20°C, and then the 3C libraries were centrifuged at 3000 rpm for 45 min at 4°C to pellet the samples. The controls were centrifuged at 14,000 rpm. The pellets were resuspended in 70% ethanol and centrifuged as described above. The pellets of 3C libraries and controls were resuspended in 300 and 20 μl of dH_2O, respectively, and stored at −20°C. Sample concentrations were measured by Qubit. Digestion and ligation efficiencies were assessed by gel electrophoresis on a 0.9% agarose gel and by quantitative PCR (SYBR Green, Thermo Fisher Scientific).

Isolated DNA from 3C libraries was quantified using a Qubit fluorometer (Life technologies), and 10 μg of each library was sheared in dH_2O using a QSonica Q800R to an average fragment size of 350 bp. QSonica settings used were 60% amplitude, 30-s on, 30-s off, and 2-min intervals, for a total of five intervals at 4°C. After shearing, DNA was purified using AMPureXP beads (Agencourt). DNA size was assessed on a Bioanalyzer 2100 using a DNA 1000 chip (Agilent), and DNA concentration was checked via Qubit. SureSelect XT library prep kits (Agilent) were used to repair DNA ends and for adaptor ligation following the manufacturer's protocol. Excess adaptors were removed using AMPureXP beads. Size and concentration were checked again by Bioanalyzer 2100 using a DNA 1000 chip and by Qubit fluorometer before hybridization. One microgram of adaptor-ligated library was used as input for the SureSelect XT Capture Kit using the manufacturer's protocol and our custom-designed 41K promoter Capture C probe set. The quantity and quality of the captured libraries were assessed by Bioanalyzer using a high-sensitivity DNA chip and by Qubit fluorometer. SureSelect XT libraries were then paired-end–sequenced on Illumina NovaSeq 6000 platform (51-bp read length) at the Center for Spatial and Functional Genomics at CHOP.

ATAC-seq peak calling

CNPC ATAC-seq peaks were called using the ENCODE ATAC-seq pipeline (https://encodeproject.org/atac-seq/) with the optimal peak IDR option.

Promoter Capture C preprocessing and interaction calling

Paired-end reads from NPCs were preprocessed using the HICUP pipeline (*69*) (v0.5.9), with bowtie2 as aligner and hg19 as the reference genome. Nonhybrid read count from all baited promoters was used for significant promoter interaction calling. Significant promoter interactions at one Dpn II fragment resolution were called using CHiCAGO (v1.1.8) (*70*) with default parameters except for bin size set to 2500. Significant interactions at four Dpn II fragment resolution were also called using CHiCAGO with artificial baitmap and rmap files, in which Dpn II fragments were concatenated in silico into four consecutive fragments. Interactions with a CHiCAGO score > 5 in at least one cell type in either one-fragment or four-fragment resolution were considered as significant interactions. The significant interactions were lastly converted to ibed format, in which each line represents a physical interaction between fragments.

RNA-seq expression analysis

STAR (v2.5.2b) was used to align each paired-end fastq file for each RNA-seq library independently to reference GRCh37. GencodeV19 was used for gene feature annotation, and the raw read count for gene feature was calculated by htseq-count (v0.6.1) with parameter settings -f bam -r pos -s reverse -t exon -m union (*71*). The gene features localized on chrM or annotated as rRNAs were removed from the final sample-by-gene read count matrix. TPM and percentile expression values were calculated from the raw read counts for each gene with a custom script in R using GencodeV19 annotation for gene lengths.

Variant to gene mapping

Proxy SNPs for each sentinel SNP (*8*) were calculated using online SNP annotator SNiPA (https://snipa.helmholtz-muenchen.de/snipa/) (settings: genome assembly as GRCh37, variant set as 1000 Genome Phase 3 v5, LD r-square cutoff as 0.7) in the European population. Proxy SNPs positions were intersected with the position of the NPC ATAC-seq peaks to identify open "informative" proxies. Capture C chromatin loops to open gene promoters were annotated to each open proxy SNP using custom scripts.

Drosophila husbandry

Flies were grown and maintained on standard *Drosophila* food media (Bloomington Recipe, Genesee Scientific, San Diego, CA) in incubators (Powers Scientific, Warminster, Pennsylvania) at 25°C on a 12:12 light/dark cycle with humidity set to 55 to 65%. The following fly strains were obtained from the Bloomington Stock Center (*72*): w^{1118} (#5905), nsyb-GAL4 (#39171), and UAS-*PIG-Q*–RNAi^{1} (#67955), while the UAS-*PIG-Q*–RNAi^{2} (#107774) was obtained from the Vienna *Drosophila* Resource Center (*73*) or the Bloomington Stock Center (*72*). The stock numbers of all lines used for screeding are described in tables S2 to S4 unless otherwise stated. Mated females aged 3 to 5 days were used for all experiments performed in this study.

Sleep and arousal threshold measurements

Flies were acclimated to experimental conditions for at least 24 hours before the start of all behavioral analysis. Measurements of sleep and arousal threshold were then measured over the course of 3 days starting at ZT0 using the DAM system (Trikinetics, Waltham, MA, USA), as previously described (*29*, *30*, *74*). For each individual fly, the DAM system measures activity by counting the number of infrared beam crossings over time. These activity data were then used to calculate sleep, defined as bouts of immobility of 5 min or more, using the *Drosophila* Sleep Counting Macro (*75*), from which sleep traits were then extracted. Waking activity was quantified as the average number of beam crossings per waking minute, as previously described (*75*).

Arousal threshold was measured using the DART, as previously described (*34*). In brief, individual female flies were loaded into plastic tubes (Trikinectics, Waltham, Massachusetts) and placed onto trays containing vibrating motors. Flies were recorded continuously using a USB webcam (QuickCam Pro 900, Logitech, Lausanne, Switzerland) with a resolution of 960 × 720 at 5 frames/s. The vibrational stimulus, video tracking parameters, and data analysis were performed using the DART interface developed in MATLAB (The MathWorks Inc., Natick, MA). To track fly movement, raw video flies were subsampled to 1 frame/s. Fly movement, or a difference in pixilation from one frame to the next, was detected by subtracting a background image from the current frame. The background image was generated as the average of 20 randomly selected frames from a given video. Fly activity was measured as movement of greater than 3 mm. Sleep was determined by the absolute location of each fly and was measured as bouts of immobility for 5 min or more. Reactivity was assessed using a vibration intensity of 1.2 g, once per hour over 3 days starting at ZT0.

All measurements of sleep and arousal threshold were combined across the 3 days of testing. Statistical analyses were performed in Prism (GraphPad Software 9.3). Unless otherwise noted, a *t* test or one-way ANOVA was used for comparisons between two genotypes or two or more genotypes, respectively. All post hoc analyses were performed using Dunnett's multiple comparisons test. Measurements of arousal threshold were not normally distributed, so

the nonparametric restricted maximum likelihood estimation was used. To characterize the relationship between the change in reactivity and bout length, we performed linear regression analyses. An analysis of covariance was used to compare the elevations of different genotypes.

Sleep deprivation

Sleep deprivation experiments were performed as previously described (*76*). Upon experiment onset, baseline sleep was measured starting at ZT0 for 24 hours. For the following 24 hours, flies were mechanically sleep-deprived, during which sleep was also measured. To assess homeostatic rebound, flies were returned to standard conditions, and sleep was measured during the subsequent day (ZT0 to ZT12). To determine whether there exists a homeostatic rebound in sleep duration, baseline daytime sleep was compared to daytime sleep during recovery.

Generation of zebrafish mutant

PIG-Q KO mutants were generated using CRISPR-Cas9 as previously described (*77*). Single-cell stage embryos were injected with preformed ribonucleoprotein complexes containing Cas9 protein and a single guide RNA (sgRNA), which has a target sequence of 5′-TCTAAAGAGTCGCCAGAGCGAGG-3′. Scrambled sgRNA with sequence 5′-CGTTAATCGCGTATAATACG-3′ was used for negative control injections. Mutant animals were genotyped using headloop PCR as described previously (*41*, *78*). Briefly, the headloop primer design acts by forming a hairpin structure in the DNA. In wild-type larvae, the primer will form a hairpin that suppresses the PCR amplification, and no band will appear in the gel. In CRISPR mutant larvae, the primer will not form the hairpin because the DNA is cut, and the PCR amplification will not be suppressed; therefore, a band will appear in the gel at the appropriate product size. Larvae were included in the sleep assay analysis if they had mutation efficiency greater than 0.9 (90%), as determined by the ratio of headloop to standard PCR product using ImageJ. Primers for the target region included standard primers 5′-GTTGGAGTGACTCACCAGGG-3′ and 5′-TGAGTACTGCAGGGTGGTTTC-3′ and headloop primers 5′- GTTGGAGTGACTCACCAGGG-3′ and 5′-AGAGCGAGGAGAGACCGTAGTGAGTACTG-CAGGGTGGTTTC-3′. gRNA design was performed using CRISPOR tefor (http://crispor.tefor.net/) to optimize sensitivity and specificity >95%, with minimal off-target effects. CRISPR sgRNA is designed for the exonic region with the highest conservation across species using MARRVEL (*42*) and is a component of the *N*-acetylglucosaminyl transferase component that overlaps between both *PIG-Q* transcripts to ensure disruption of all possible transcripts. Target region was also free of in silico–predicted SNPs, and Sanger sequencing was performed to experimentally demonstrate that target region was free of mutations.

Sleep/wake assay in zebrafish

Zebrafish experiments were performed in accordance with University of Pennsylvania Institutional Animal Care and Use Committee guidelines (animal protocol #806646). Zebrafish embryos were collected from AB/TL strain incross breeding pairs in the morning of spawning and injected with CRISPR-Cas9 reagents at the single-cell stage. Embryos were raised on a 14:10-hour light/dark cycle. Animals were housed in petri dishes with approximately 50 per dish in standard embryo medium (E3 medium; 5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM $MgSO_4$, and 10^{-5}% methylene blue) and kept in an incubator at 28.5°C. Dead embryos and shed chorion membranes were removed on days 2 to 4 after fertilization. Scramble gRNA–injected and *PIG-Q* KO larvae were individually placed into each well in alternating rows of a 96-well plate in 650 μl of E3 embryo medium without methylene blue on 5 days after fertilization. Activity was captured using automated video tracking (ViewPoint Life Sciences) for 72 hours. Behavioral phenotyping of larvae at F0 generation, which display high-mutation efficiencies greater than 90%, has been modified from (*41*). The 96-well plate was housed in a Zebrabox (ViewPoint Life Sciences) with customizable light parameters and a Dinion one-third inch monochrome camera (Dragonfly2, Point Grey) fitted with a variable-focus megapixel lens (M5018-MP, Computar) and infrared filter. The plate was fitted in a chamber filled with recirculating water connected to a temperature control unit to maintain a stable temperature of 28.5°C, which is the optimum growth temperature of zebrafish. Activity was captured in quantization mode with the following detection parameters: threshold, 20; burst, 29; freeze, 3; bin size, 60 s, as described previously (*79*).

Zebrafish sleep/wake data analysis

An acclimation period was removed, and data analysis consisted of 2 days and two nights. Data were processed using custom MATLAB (The MathWorks Inc.) scripts as performed in (*79*) with modifications. Movement was captured as seconds per minute, and any 1-min period with less than 0.5 s of total movement was defined as 1 min of sleep [modified from (*80*)]. Sleep bouts were defined as a continuous string of sleep defined in minutes, and sleep bout length was calculated as minutes across the continuous string. The average activity was defined as the average amount of activity using the threshold of 0.5 s to define waking activity and reported as seconds per awake minute per hour. Statistical tests were performed in Prism (Graphpad). Activity was combined across both days and sleep across both nights for analysis. Independent Student′s *t* test was used to compare groups with equal variances, while Welch′s correction was applied when variances were determined to be unequal.

DNA extraction and PCR validation

DNA extraction was performed per the manufacturer′s protocol (Quantabio, Beverly, MA) immediately following completion of the sleep assay. Larvae were euthanized by rapid cooling on a mixture of ice and water between 2° and 4°C for a minimum of 30 min after complete cessation of movement was observed. Larvae were transferred to a new 96-well PCR plate, and excess E3 buffer was removed. DNA extraction buffer (25 μl) was then added, and larvae were completely submerged. The plate was sealed and heated for 30 min at 95°C and then cooled to room temperature. DNA stabilization buffer (25 μl) was then added, and genomic DNA was stored at 4°C. For PCR validation, each well of a PCR plate contained 0.1 μl of Phusion DNA polymerase, 5 μl of 5× Phusion HF buffer, 0.5 μl of deoxynucleotide triphosphate mix, 0.5 μl of 10 μM *PIG-Q* forward primer, 0.5 μl of 10 μM *PIG-Q* reverse primer or headloop antisense primer, 16.4 μl of nuclease-free water, and 2 μl of twofold-diluted genomic DNA for a final volume of 25 μl. PCR plate was sealed and placed into a thermocycler. The PCR reaction conditions were one cycle of 98°C for 90 s; 30 cycles of 98°C for 10 s, 64°C for 10 s, 72°C for 15 s, 72°C for 15 s; and

one cycle of 72°C for 5 min and then stored at 4°C. Samples were run on 2% agarose gel and quantified using ImageJ for the ratio of headloop PCR product to standard PCR product to calculate mutation efficiency, as previously described (*41*).

Supplementary Materials

This PDF file includes:
Figs. S1 to S5

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S4

REFERENCES AND NOTES

1. S. Banks, D. F. Dinges, Behavioral and physiological consequences of sleep restriction. *J. Clin. Sleep Med.* **3**, 519–528 (2007).
2. D. M. Arble, J. Bass, C. D. Behn, M. P. Butler, E. Challet, C. Czeisler, C. M. Depner, J. Elmquist, P. Franken, M. A. Grandner, E. C. Hanlon, A. C. Keene, M. J. Joyner, I. Karatsoreos, P. A. Kern, S. Klein, C. J. Morris, A. I. Pack, S. Panda, L. J. Ptacek, N. M. Punjabi, P. Sassone-Corsi, F. A. Scheer, R. Saxena, E. R. Seaquest, M. S. Thimgan, E. van Cauter, K. P. Wright, Impact of sleep and circadian disruption on energy balance and diabetes: A summary of workshop discussions. *Sleep* **38**, 1849–1860 (2015).
3. M. J. Lind, P. R. Gehrman, Genetic pathways to insomnia. *Brain Sci.* **6**, 64 (2016).
4. M. A. Grandner, F.-X. Fernandez, The translational neuroscience of sleep: A contextual framework. *Science* **374**, 568–573 (2021).
5. S. E. Jones, J. M. Lane, A. R. Wood, V. T. van Hees, J. Tyrrell, R. N. Beaumont, A. R. Jeffries, H. S. Dashti, M. Hillsdon, K. S. Ruth, M. A. Tuke, H. Yaghootkar, S. A. Sharp, Y. Jie, W. D. Thompson, J. W. Harrison, A. Dawes, E. M. Byrne, H. Tiemeier, K. V. Allebrandt, J. Bowden, D. W. Ray, R. M. Freathy, A. Murray, D. R. Mazzotti, P. R. Gehrman, D. A. Lawlor, T. M. Frayling, M. K. Rutter, D. A. Hinds, R. Saxena, M. N. Weedon, Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* **10**, 343 (2019).
6. H. J. Ban, S. C. Kim, J. Seo, H. B. Kang, J. K. Choi, Genetic and metabolic characterization of insomnia. *PLOS ONE* **6**, e18455 (2011).
7. E. M. Byrne, P. R. Gehrman, S. E. Medland, D. R. Nyholt, A. C. Heath, P. A. F. Madden, I. B. Hickie, C. M. van Duijn, A. K. Henders, G. W. Montgomery, N. G. Martin, N. R. Wray; Chronogen Consortium, A genome-wide association study of sleep habits and insomnia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **162B**, 439–451 (2013).
8. P. R. Jansen, K. Watanabe, S. Stringer, N. Skene, J. Bryois, A. R. Hammerschlag, C. A. de Leeuw, J. S. Benjamins, A. B. Muñoz-Manchado, M. Nagel, J. E. Savage, H. Tiemeier, T. White; 23andMe Research Team, J. Y. Tung, D. A. Hinds, V. Vacic, X. Wang, P. F. Sullivan, S. van der Sluis, T. J. C. Polderman, A. B. Smit, J. Hjerling-Leffler, E. J. W. van Someren, D. Posthuma, Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).
9. T. Lappalainen, D. G. MacArthur, From variant to function in human disease genetics. *Science* **373**, 1464–1468 (2021).
10. M. Claussnitzer, J. H. Cho, R. Collins, N. J. Cox, E. T. Dermitzakis, M. E. Hurles, S. Kathiresan, E. E. Kenny, C. M. Lindgren, D. G. MacArthur, K. N. North, S. E. Plon, H. L. Rehm, N. Risch, C. N. Rotimi, J. Shendure, N. Soranzo, M. I. McCarthy, A brief history of human disease genetics. *Nature* **577**, 179–189 (2020).
11. S. Smemo, J. J. Tena, K. H. Kim, E. R. Gamazon, N. J. Sakabe, C. Gómez-Marín, I. Aneas, F. L. Credidio, D. R. Sobreira, N. F. Wasserman, J. H. Lee, V. Puviindran, D. Tam, M. Shen, J. E. Son, N. A. Vakili, H. K. Sung, S. Naranjo, R. D. Acemel, M. Manzanares, A. Nagy, N. J. Cox, C. C. Hui, J. L. Gomez-Skarmeta, M. A. Nóbrega, Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*. *Nature* **507**, 371–375 (2014).
12. D. R. Sobreira, A. C. Joslin, Q. Zhang, I. Williamson, G. T. Hansen, K. M. Farris, N. J. Sakabe, N. Sinnott-Armstrong, G. Bozek, S. O. Jensen-Cody, K. H. Flippo, C. Ober, W. A. Bickmore, M. Potthoff, M. Chen, M. Claussnitzer, I. Aneas, M. A. Nóbrega, Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of *IRX3* and *IRX5*. *Science* **372**, 1085–1091 (2021).
13. M. Claussnitzer, S. N. Dankel, K.-H. Kim, G. Quon, W. Meuleman, C. Haugen, V. Glunk, I. S. Sousa, J. L. Beaudry, V. Puviindran, N. A. Abdennur, J. Liu, P.-A. Svensson, Y.-H. Hsu, D. J. Drucker, G. Mellgren, C.-C. Hui, H. Hauner, M. Kellis, FTO obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015).
14. A. Chesi, Y. Wagley, M. E. Johnson, E. Manduchi, C. Su, S. Lu, M. E. Leonard, K. M. Hodge, J. A. Pippin, K. D. Hankenson, A. D. Wells, S. F. A. Grant, Genome-scale Capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. *Nat. Commun.* **10**, 1260 (2019).
15. D. L. Cousminer, Y. Wagley, J. A. Pippin, A. Elhakeem, G. P. Way, M. C. Pahl, S. E. McCormack, A. Chesi, J. A. Mitchell, J. M. Kindler, D. Baird, A. Hartley, L. Howe, H. J. Kalkwarf, J. M. Lappe, S. Lu, M. E. Leonard, M. E. Johnson, H. Hakonarson, V. Gilsanz, J. A. Shepherd, S. E. Oberfield, C. S. Greene, A. Kelly, D. A. Lawlor, B. F. Voight, A. D. Wells, B. S. Zemel, K. D. Hankenson, S. F. A. Grant, Genome-wide association study implicates novel loci and reveals candidate effector genes for longitudinal pediatric bone accrual. *Genome Biol.* **22**, 1 (2021).
16. J. Pluta, L. C. Pyle, K. T. Nead, R. Wilf, M. Li, N. Mitra, B. Weathers, K. D'Andrea, K. Almstrup, L. Anson-Cartwright, J. Benitez, C. D. Brown, S. Chanock, C. Chen, V. K. Cortessis, A. Ferlin, C. Foresta, M. Gamulin, J. A. Gietema, C. Grasso, M. H. Greene, T. Grotmol, R. J. Hamilton, T. B. Haugen, R. Hauser, M. A. T. Hildebrandt, M. E. Johnson, R. Karlsson, L. A. Kiemeney, D. Lessel, R. A. Lothe, J. T. Loud, C. Loveday, P. Martin-Gimeno, C. Meijer, J. Nsengimana, D. I. Quinn, T. Rafnar, S. Ramdas, L. Richiardi, R. I. Skotheim, K. Stefansson, C. Turnbull, D. J. Vaughn, F. Wiklund, X. Wu, D. Yang, T. Zheng, A. D. Wells, S. F. A. Grant, E. Rajpert-De Meyts, S. M. Schwartz, D. T. Bishop, K. A. Mc Glynn, P. A. Kanetsky, K. L. Nathanson; The Testicular Cancer Consortium, Identification of 22 susceptibility loci associated with testicular germ cell tumors. *Nat. Commun.* **12**, 4487 (2021).
17. C. Su, M. E. Johnson, A. Torres, R. M. Thomas, E. Manduchi, P. Sharma, P. Mehra, C. le Coz, M. E. Leonard, S. Lu, K. M. Hodge, A. Chesi, J. Pippin, N. Romberg, S. F. A. Grant, A. D. Wells, Mapping effector genes at lupus GWAS loci using promoter Capture-C in follicular helper T cells. *Nat. Commun.* **11**, 3294 (2020).
18. M. C. Pahl, C. A. Doege, K. M. Hodge, S. H. Littleton, M. E. Leonard, S. Lu, R. Rausch, J. A. Pippin, M. C. de Rosa, A. Basak, J. P. Bradfield, R. K. Hammond, K. Boehm, R. I. Berkowitz, C. Lasconi, C. Su, A. Chesi, M. E. Johnson, A. D. Wells, B. F. Voight, R. L. Leibel, D. L. Cousminer, S. F. A. Grant, Cis-regulatory architecture of human ESC-derived hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits. *Nat. Commun.* **12**, 6749 (2021).
19. T. E. Scammell, E. Arrigoni, J. O. Lipton, Neural circuitry of wakefulness and sleep. *Neuron* **93**, 747–765 (2017).
20. C. B. Saper, P. M. Fuller, Wake-sleep circuitry: An overview. *Curr. Opin. Neurobiol.* **44**, 186–192 (2017).
21. C. Su, M. Argenziano, S. Lu, J. A. Pippin, M. C. Pahl, M. E. Leonard, D. L. Cousminer, M. E. Johnson, C. Lasconi, A. D. Wells, A. Chesi, S. F. A. Grant, 3D promoter architecture reorganization during iPSC-derived neuronal cell differentiation implicates target genes for neurodevelopmental disorders. *Prog. Neurobiol.* **201**, 102000 (2021).
22. O. T. Shafer, A. C. Keene, The regulation of *Drosophila* sleep. *Curr. Biol.* **31**, R38–R49 (2021).
23. G. Artiushin, A. Sehgal, The *Drosophila* circuitry of sleep–wake regulation. *Curr. Opin. Neurobiol.* **44**, 243–250 (2017).
24. C. Su, M. C. Pahl, S. F. A. Grant, A. D. Wells, Restriction enzyme selection dictates detection range sensitivity in chromatin conformation capture-based variant-to-gene mapping approaches. *Hum. Genet.* **140**, 1441–1448 (2021).
25. A. V. Salminen, B. Schormair, C. Flachskamm, M. Torres, B. Müller-Myhsok, M. Kimura, J. Winkelmann, Sleep disturbance by pramipexole is modified by Meis1 in mice. *J. Sleep Res.* **27**, e12557 (2018).
26. F. Sarayloo, P. A. Dion, G. A. Rouleau, *MEIS1* and restless legs syndrome: A comprehensive review. *Front. Neurol.* **10**, 935 (2019).
27. W. J. Joiner, Unraveling the evolutionary determinants of sleep. *Curr. Biol.* **26**, R1073–R1087 (2016).
28. S. Ly, A. I. Pack, N. Naidoo, The neurobiological basis of sleep: Insights from *Drosophila*. *Neurosci. Biobehav. Rev.* **87**, 67–86 (2018).
29. P. J. Shaw, C. Cirelli, R. J. Greenspan, G. Tononi, Correlates of sleep and waking in *Drosophila* melanogaster. *Science* **287**, 1834–1837 (2000).
30. J. C. Hendricks, S. M. Finn, K. A. Panckeri, J. Chavkin, J. A. Williams, A. Sehgal, A. I. Pack, Rest in *Drosophila* is a sleep-like state. *Neuron* **25**, 129–138 (2000).
31. Y. Hu, I. Flockhart, A. Vinayagam, C. Bergwitz, B. Berger, N. Perrimon, S. E. Mohr, An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics.* **12**, 357 (2011).
32. L. J. Starr, J. W. Spranger, V. K. Rao, R. Lutz, A. T. Yetman, PIGQ glycosylphosphatidylinositol-anchored protein deficiency: Characterizing the phenotype. *Am. J. Med. Genet. A* **179**, 1270–1275 (2019).
33. B. van Alphen, M. H. W. Yap, L. Kirszenblat, B. Kottler, B. van Swinderen, A dynamic deep sleep stage in *Drosophila*. *J. Neurosci.* **33**, 6917–6927 (2013).
34. R. Faville, B. Kottler, G. J. Goodhill, P. J. Shaw, B. van Swinderen, How deeply does your mutant sleep? Probing arousal to better understand sleep defects in *Drosophila*. *Sci. Rep.* **5**, 8454 (2015).
35. S. C. Renn, J. H. Park, M. Rosbash, J. C. Hall, P. H. Taghert, A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in *Drosophila*. *Cell* **99**, 791–802 (1999).

36. D. Park, J. A. Veenstra, J. H. Park, P. H. Taghert, Mapping peptidergic cells in *Drosophila*: Where DIMM fits in. *PLOS ONE* **3**, e1896 (2008).
37. T. Kinoshita, Biosynthesis and biology of mammalian GPI-anchored proteins. *Open Biol.* **10**, 190290 (2020).
38. K. Koh, W. J. Joiner, M. N. Wu, Z. Yue, C. J. Smith, A. Sehgal, Identification of SLEEPLESS, a sleep-promoting factor. *Science* **321**, 372–376 (2008).
39. Y. Hu, A. Comjean, N. Perrimon, S. E. Mohr, The *Drosophila* gene expression tool (DGET) for expression analyses. *BMC Bioinformatics* **18**, 98 (2017).
40. C. N. Chiu, D. A. Prober, Regulation of zebrafish sleep and arousal states: Current and prospective approaches. *Front. Neural Circuits* **7**, 58 (2013).
41. F. Kroll, G. T. Powell, M. Ghosh, G. Gestri, P. Antinucci, T. J. Hearn, H. Tunbak, S. Lim, H. W. Dennis, J. M. Fernandez, D. Whitmore, E. Dreosti, S. W. Wilson, E. J. Hoffman, J. Rihel, A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* **10**, e59683 (2021).
42. J. Wang, R. Al-Ouran, Y. Hu, S.-Y. Kim, Y.-W. Wan, M. F. Wangler, S. Yamamoto, H.-T. Chao, A. Comjean, S. E. Mohr, N. P. UDN, Z. Liu, H. J. Bellen, MARRVEL: Integration of human and model organism genetic resources to facilitate functional annotation of the human genome. *Am. J. Hum. Genet.* **100**, 843–853 (2017).
43. H. Shams-Eldin, N. Azzouz, M. H. Kedees, P. Orlean, T. Kinoshita, R. T. Schwarz, The GPI1 homologue from Plasmodium falciparum complements a Saccharomyces cerevisiae GPI1 anchoring mutant. *Mol. Biochem. Parasitol.* **120**, 73–81 (2002).
44. C. N. Chiu, J. Rihel, D. A. Lee, C. Singh, E. A. Mosser, S. Chen, V. Sapin, U. Pham, J. Engle, B. J. Niles, C. J. Montz, S. Chakravarthy, S. Zimmerman, K. Salehi-Ashtiani, M. Vidal, A. F. Schier, D. A. Prober, A zebrafish genetic screen identifies neuromedin U as a regulator of sleep/wake states. *Neuron* **89**, 842–856 (2016).
45. J. Rihel, D. A. Prober, A. Arvanites, K. Lam, S. Zimmerman, S. Jang, S. J. Haggarty, D. Kokel, L. L. Rubin, R. T. Peterson, A. F. Schier, Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science* **327**, 348–351 (2010).
46. W. Yi, Y. Zhang, Y. Tian, J. Guo, Y. Li, A. Guo, A subset of cholinergic mushroom body neurons requires Go signaling to regulate sleep in *Drosophila*. *Sleep* **36**, 1809–1821 (2013).
47. X. Dai, E. Zhou, W. Yang, R. Mao, W. Zhang, Y. Rao, Molecular resolution of a behavioral paradox: Sleep and arousal are regulated by distinct acetylcholine receptors in different neuronal types in *Drosophila*. *Sleep* **44**, zsab017 (2021).
48. J. M. Donlea, D. Pimentel, G. Miesenböck, Neuronal machinery of sleep homeostasis in *Drosophila*. *Neuron* **81**, 860–872 (2014).
49. K. M. Parisky, J. Agosto, S. R. Pulver, Y. Shang, E. Kuklin, J. J. L. Hodge, K. Kang, X. Liu, P. A. Garrity, M. Rosbash, L. C. Griffith, PDF cells are a GABA-responsive wake-promoting component of the *Drosophila* sleep circuit. *Neuron* **60**, 672–682 (2008).
50. M. N. Wu, W. J. Joiner, T. Dean, Z. Yue, C. J. Smith, D. Chen, T. Hoshi, A. Sehgal, K. Koh, SLEEPLESS, a Ly-6/neurotoxin family member, regulates the levels, localization and activity of Shaker. *Nat. Neurosci.* **13**, 69–75 (2010).
51. T. Kinoshita, Glycosylphosphatidylinositol (GPI) anchors: Biochemistry and cell biology: Introduction to a thematic review series. *J. Lipid Res.* **57**, 4–5 (2016).
52. T. M. Frayling, N. J. Timpson, M. N. Weedon, E. Zeggini, R. M. Freathy, C. M. Lindgren, J. R. B. Perry, K. S. Elliott, H. Lango, N. W. Rayner, B. Shields, L. W. Harries, J. C. Barrett, S. Ellard, C. J. Groves, B. Knight, A. M. Patch, A. R. Ness, S. Ebrahim, D. A. Lawlor, S. M. Ring, Y. Ben-Shlomo, M. R. Jarvelin, U. Sovio, A. J. Bennett, D. Melzer, L. Ferrucci, R. J. F. Loos, I. Barroso, N. J. Wareham, F. Karpe, K. R. Owen, L. R. Cardon, M. Walker, G. A. Hitman, C. N. A. Palmer, A. S. F. Doney, A. D. Morris, G. D. Smith, A. T. Hattersley, M. I. McCarthy, A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 (2007).
53. W. Wen, Y. S. Cho, W. Zheng, R. Dorajoo, N. Kato, L. Qi, C. H. Chen, R. J. Delahanty, Y. Okada, Y. Tabara, D. Gu, D. Zhu, C. A. Haiman, Z. Mo, Y. T. Gao, S. M. Saw, M. J. Go, F. Takeuchi, L. C. Chang, Y. Kokubo, J. Liang, M. Hao, L. le Marchand, Y. Zhang, Y. Hu, T. Y. Wong, J. Long, B. G. Han, M. Kubo, K. Yamamoto, M. H. Su, T. Miki, B. E. Henderson, H. Song, A. Tan, J. He, D. P. K. Ng, Q. Cai, T. Tsunoda, F. J. Tsai, N. Iwai, G. K. Chen, J. Shi, J. Xu, X. Sim, Y. B. Xiang, S. Maeda, R. T. H. Ong, C. Li, Y. Nakamura, T. Aung, N. Kamatani, J. J. Liu, W. Lu, M. Yokota, M. Seielstad, C. S. J. Fann, J. Y. Wu, J. Y. Lee, F. B. Hu, T. Tanaka, E. S. Tai, X. O. Shu, Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.* **44**, 307–311 (2012).
54. Y. Okada, M. Kubo, H. Ohmiya, A. Takahashi, N. Kumasaka, N. Hosono, S. Maeda, W. Wen, R. Dorajoo, M. J. Go, W. Zheng, N. Kato, J. Y. Wu, Q. Lu; GIANT consortium, T. Tsunoda, K. Yamamoto, Y. Nakamura, N. Kamatani, T. Tanaka, Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. *Nat. Genet.* **44**, 302–306 (2012).
55. M. T. Hassanein, H. N. Lyon, T. T. Nguyen, E. L. Akylbekova, K. Waters, G. Lettre, B. Tayo, T. Forrester, D. F. Sarpong, D. O. Stram, J. L. Butler, R. Wilks, J. Liu, L. le Marchand, L. N. Kolonel, X. Zhu, B. Henderson, R. Cooper, C. McKenzie, H. A. Taylor Jr., C. A. Haiman, J. N. Hirschhorn, Fine mapping of the association with obesity at the *FTO* locus in African-derived populations. *Hum. Mol. Genet.* **19**, 2907–2916 (2010).
56. J. P. Bradfield, H. R. Taal, N. J. Timpson, A. Scherag, C. Lecoeur, N. M. Warrington, E. Hypponen, C. Holst, B. Valcarcel, E. Thiering, R. M. Salem, F. R. Schumacher, D. L. Cousminer, P. M. A. Sleiman, J. Zhao, R. I. Berkowitz, K. S. Vimaleswaran, I. Jarick, C. E. Pennell, D. M. Evans, B. S. Pourcain, D. J. Berry, D. O. Mook-Kanamori, A. Hofman, F. Rivadeneira, A. G. Uitterlinde n, C. M. van Duijn, R. J. P. van der Valk, J. C. de Jongste, D. S. Postma, D. I. Boomsma, W. J. Gauderman, M. T. Hassanein, C. M. Lindgren, R. Mägi, C. A. G. Boreham, C. E. Neville, L. A. Moreno, P. Elliott, A. Pouta, A. L. Hartikainen, M. Li, O. Raitakari, T. Lehtimäki, J. G. Eriksson, A. Palotie, J. Dallongeville, S. Das, P. Deloukas, G. Mcmahon, S. M. Ring, J. P. Kemp, J. L. Buxton, A. I. F. Blakemore, M. Bustamante, M. Guxens, J. N. Hirschhorn, M. W. Gillman, E. Kreiner-Müller, H. Bisgaard, F. D. Gilliland, J. Heinrich, E. Wheeler, I. Barroso, S. O'rahilly, A. Meirhaeghe, T. I. A. Sørensen, C. Power, L. J. Palmer, A. Hinney, I. S. Farooqi, M. I. Mccarthy, P. Froguel, D. Meyre, J. Hebebrand, M. R. Jarvelin, G. D. Smith, H. Hakonarson, S. F. A. Grant; Early Growth Genetics Consortium, A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.* **44**, 526–531 (2012).
57. H. Huang, C. T. Zhu, L. L. Skuja, D. J. Hayden, A. C. Hart, Genome-wide screen for genes involved in *Caenorhabditis elegans* developmentally timed sleep. *G3* **7**, 2907–2917 (2017).
58. O. Lenz, J. Xiong, M. D. Nelson, D. M. Raizen, J. A. Williams, FMRFamide signaling promotes stress-induced sleep in *Drosophila*. *Brain Behav. Immun.* **47**, 141–148 (2015).
59. D. Rogulja, M. W. Young, Control of sleep by cyclin A and its regulator. *Science* **335**, 1617–1621 (2012).
60. H. Funato, C. Miyoshi, T. Fujiyama, T. Kanda, M. Sato, Z. Wang, J. Ma, S. Nakane, J. Tomita, A. Ikkyu, M. Kakizaki, N. Hotta-Hirashima, S. Kanno, H. Komiya, F. Asano, T. Honda, S. J. Kim, K. Harano, H. Muramoto, T. Yonezawa, S. Mizuno, S. Miyazaki, L. Connor, V. Kumar, I. Miura, T. Suzuki, A. Watanabe, M. Abe, F. Sugiyama, S. Takahashi, K. Sakimura, Y. Hayashi, Q. Liu, K. Kume, S. Wakana, J. S. Takahashi, M. Yanagisawa, Forward-genetics analysis of sleep in randomly mutagenized mice. *Nature* **539**, 378–383 (2016).
61. S. Hidalgo, J. M. Campusano, J. J. L. Hodge, Assessing olfactory, memory, social and circadian phenotypes associated with schizophrenia in a genetic model based on Rim. *Transl. Psychiatry* **11**, 292 (2021).
62. K. V. Allebrandt, N. Amin, B. Müller-Myhsok, T. Esko, M. Teder-Laving, R. V. D. M. Azevedo, C. Hayward, J. van Mill, N. Vogelzangs, E. W. Green, S. A. Melville, P. Lichtner, H. E. Wichmann, B. A. Oostra, A. C. J. W. Janssens, H. Campbell, J. F. Wilson, A. A. Hicks, P. P. Pramstaller, Z. Dogas, I. Rudan, M. Merrow, B. Penninx, C. P. Kyriacou, A. Metspalu, C. M. van Duijn, T. Meitinger, T. Roenneberg, A K(ATP) channel gene effect on sleep duration: From genome-wide association studies to function in *Drosophila*. *Mol. Psychiatry* **18**, 122–132 (2013).
63. D. A. Lee, J. Liu, Y. Hong, J. M. Lane, A. J. Hill, S. L. Hou, H. Wang, G. Oikonomou, U. Pham, J. Engle, R. Saxena, D. A. Prober, Evolutionarily conserved regulation of sleep by epidermal growth factor receptor signaling. *Sci. Adv.* **5**, eaax4249 (2019).
64. K. Foltenyi, R. J. Greenspan, J. W. Newport, Activation of EGFR and ERK by rhomboid signaling regulates the consolidation and maintenance of sleep in *Drosophila*. *Nat. Neurosci.* **10**, 1160–1167 (2007).
65. J. Konietzka, M. Fritz, S. Spiri, R. McWhirter, A. Leha, S. Palumbos, W. S. Costa, A. Oranth, A. Gottschalk, D. M. Miller III, A. Hajnal, H. Bringmann, Epidermal growth factor signaling promotes sleep through a combined series and parallel neural circuit. *Curr. Biol.* **30**, 1–16.e13 (2020).
66. S. E. Jones, V. T. van Hees, D. R. Mazzotti, P. Marques-Vidal, S. Sabia, A. van der Spek, H. S. Dashti, J. Engmann, D. Kocevska, J. Tyrrell, R. N. Beaumont, M. Hillsdon, K. S. Ruth, M. A. Tuke, H. Yaghootkar, S. A. Sharp, Y. Ji, J. W. Harrison, R. M. Freathy, A. Murray, A. I. Luik, N. Amin, J. M. Lane, R. Saxena, M. K. Rutter, H. Tiemeier, Z. Kutalik, M. Kumari, T. M. Frayling, M. N. Weedon, P. R. Gehrman, A. R. Wood, Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. *Nat. Commun.* **10**, 1585 (2019).
67. H. Wang, J. M. Lane, S. E. Jones, H. S. Dashti, H. M. Ollila, A. R. Wood, V. T. van Hees, B. Brumpton, B. S. Winsvold, K. Kantojärvi, T. Palviainen, B. E. Cade, T. Sofer, Y. Song, K. Patel, S. G. Anderson, D. A. Bechtold, J. Bowden, R. Emsley, S. D. Kyle, M. A. Little, A. S. Loudon, F. A. J. L. Scheer, S. M. Purcell, R. C. Richmond, K. Spiegelhalder, J. Tyrrell, X. Zhu, C. Hublin, J. A. Kaprio, K. Kristiansson, S. Sulkava, T. Paunio, K. Hveem, J. B. Nielsen, C. J. Willer, J. A. Zwart, L. B. Strand, T. M. Frayling, D. Ray, D. A. Lawlor, M. K. Rutter, M. N. Weedon, S. Redline, R. Saxena, Genome-wide association analysis of self-reported daytime sleepiness identifies 42 loci that suggest biological subtypes. *Nat. Commun.* **10**, 3503 (2019).
68. B. E. Cade, J. Lee, T. Sofer, H. Wang, M. Zhang, H. Chen, S. A. Gharib, D. J. Gottlieb, X. Guo, J. M. Lane, J. Liang, X. Lin, H. Mei, S. R. Patel, S. M. Purcell, R. Saxena, N. A. Shah, D. S. Evans, C. L. Hanis, D. R. Hillman, S. Mukherjee, L. J. Palmer, K. L. Stone, G. J. Tranah, N. Abe, G. Abecasis, C. Albert, L. Almasy, A. Alonso, S. Ament, P. Anderson, P. Anugu, D. Applebaum-Bowden, D. Arking, D. K. Arnett, A. Ashley-Koch, S. Aslibekyan, T. Assimes, P. Auer, D. Avramopoulos, J. Barnard, K. Barnes, R. G. Barr, E. Barron-Casella, T. Beaty, D. Becker, L. Becker, R. Beer, F. Begum, A. Beitelshees, E. Benjamin, M. Bezerra, L. Bielak, J. Bis, T. Blackwell, J. Blangero, E. Boerwinkle, I. Borecki, D. W. Bowden, R. Bowler, J. Brody, U. Broeckel, J. Broome, K. Bunting, E. Burchard, B. Cade, J. Cardwell, C. Carty, R. Casaburi,

J. Casella, M. Chaffin, C. Chang, D. Chasman, S. Chavan, B. J. Chen, W. M. Chen, Y. D. I. Chen, M. Cho, S. H. Choi, L. M. Chuang, M. Chung, E. Cornell, C. Crandall, J. Crapo, J. Curran, J. Curtis, B. Custer, C. Damcott, D. Darbar, S. Das, S. David, C. Davis, M. Daya, M. de Andrade, M. DeBaun, R. Deka, D. DeMeo, S. Devine, R. Do, Q. Duan, R. Duggirala, P. Durda, S. Dutcher, C. Eaton, L. Ekunwe, P. Ellinor, L. Emery, C. Farber, L. Farnam, T. Fingerlin, M. Flickinger, M. Fornage, N. Franceschini, M. Fu, S. M. Fullerton, L. Fulton, S. Gabriel, W. Gan, Y. Gao, M. Gass, B. Gelb, X. P. Geng, S. Germer, C. Gignoux, M. Gladwin, D. Glahn, S. Gogarten, D. W. Gong, H. Goring, C. C. Gu, Y. Guan, J. Haessler, M. Hall, D. Harris, N. Hawley, J. He, B. Heavner, S. Heckbert, R. Hernandez, D. Herrington, C. Hersh, B. Hidalgo, J. Hixson, J. Hokanson, E. Hong, K. Hoth, C. A. Hsiung, H. Huston, C. M. Hwu, M. R. Irvin, R. Jackson, D. Jain, C. Jaquish, M. A. Jhun, J. Johnsen, A. Johnson, C. Johnson, R. Johnston, K. Jones, H. M. Kang, R. Kaplan, S. Kardia, S. Kathiresan, L. Kaufman, S. Kelly, E. Kenny, M. Kessler, A. Khan, G. Kinney, B. Konkle, C. Kooperberg, H. Kramer, S. Krauter, C. Lange, E. Lange, L. Lange, C. Laurie, C. Laurie, M. LeBoff, S. S. Lee, W. J. Lee, J. LeFaive, D. Levine, D. Levy, J. Lewis, Y. Li, H. Lin, K. H. Lin, S. Liu, Y. Liu, R. Loos, S. Lubitz, K. Lunetta, J. Luo, M. Mahaney, B. Make, A. Manichaikul, J. A. Manson, L. Margolin, L. Martin, S. Mathai, R. Mathias, P. McArdle, M. L. McDonald, S. McFarland, S. McGarvey, D. A. Meyers, J. Mikulla, N. Min, M. Minear, R. L. Minster, B. D. Mitchell, M. E. Montasser, S. Musani, S. Mwasongwe, J. C. Mychaleckyj, G. Nadkarni, R. Naik, T. Naseri, P. Natarajan, S. Nekhai, D. Nickerson, K. North, J. O'Connell, T. O'Connor, H. Ochs-Balcom, N. Palmer, J. Pankow, G. Papanicolaou, M. Parker, A. Parsa, S. Penchev, J. M. Peralta, M. Perez, J. Perry, U. Peters, P. Peyser, L. S. Phillips, S. Phillips, T. Pollin, W. Post, J. P. Becker, M. P. Boorgula, M. Preuss, D. Prokopenko, B. Psaty, P. Qasba, D. Qiao, Z. Qin, N. Rafaels, L. Raffield, D. C. Rao, L. Rasmussen-Torvik, A. Ratan, R. Reed, E. Regan, A. Reiner, M. S. Reupena, K. Rice, S. Rich, D. Roden, C. Roselli, J. Rotter, I. Ruczinski, P. Russell, S. Ruuska, K. Ryan, P. Sakornsakolpat, S. Salimi, S. Salzberg, K. Sandow, V. Sankaran, C. Scheller, E. Schmidt, K. Schwander, D. Schwartz, F. Sciurba, C. Seidman, J. Seidman, V. Sheehan, A. Shetty, A. Shetty, W. H. H. Sheu, M. B. Shoemaker, B. Silver, E. Silverman, J. Smith, J. Smith, N. Smith, T. Smith, S. Smoller, B. Snively, N. Sotoodehnia, A. Stilp, E. Streeten, J. L. Su, Y. J. Sung, J. Sylvia, A. Szpiro, C. Sztalryd, D. Taliun, H. Tang, M. Taub, K. D. Taylor, S. Taylor, M. Telen, T. A. Thornton, L. Tinker, D. Tirschwell, H. Tiwari, R. Tracy, M. Tsai, D. Vaidya, P. VandeHaar, S. Vrieze, T. Walker, R. Wallace, A. Walts, E. Wan, F. F. Wang, K. Watson, D. E. Weeks, B. Weir, S. Weiss, L. C. Weng, C. Willer, K. Williams, L. K. Williams, C. Wilson, J. Wilson, Q. Wong, H. Xu, L. Yanek, I. Yang, R. Yang, N. Zaghloul, M. Zekavat, Y. Zhang, S. X. Zhao, W. Zhao, X. Zheng, D. Zhi, X. Zhou, M. Zody, S. Zoellner, B. Cade, S. Gharib, M. Goodman, D. Gottlieb, L. Hale, K. Knutson, D. Lauderdale, J. Lane, Y. Liu, B. Mitchell, D. Ngo, J. O'Connell, H. Ochs-Balcom, Patel, S. Purcell, J. Rhodes, N. Shah, J. H. Sul, S. Sunyaev, J. Wilson, H. Zhou, R. P. Tracy, R. S. Vasan, J. G. Wilson, Whole-genome association analyses of sleep-disordered breathing phenotypes in the NHLBI TOPMed program. *Genome Med.* **13**, 1–17 (2021).

69. S. Wingett, P. Ewels, M. Furlan-Magaril, T. Nagano, S. Schoenfelder, P. Fraser, S. Andrews, HiCUP: pipeline for mapping and processing Hi-C data. *F1000Research* **4**, 1310–NaN (2015).
70. J. Cairns, P. Freire-Pritchett, S. W. Wingett, C. Várnai, A. Dimond, V. Plagnol, D. Zerbino, S. Schoenfelder, B. M. Javierre, C. Osborne, P. Fraser, M. Spivakov, CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome biology* **17**, 127–NaN (2016).
71. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015); 10.1093/bioinformatics/btu638.
72. J. Q. Ni, L. P. Liu, R. Binari, R. Hardy, H. S. Shim, A. Cavallaro, M. Booker, B. D. Pfeiffer, M. Markstein, H. Wang, C. Villalta, T. R. Laverty, L. A. Perkins, N. Perrimon, A *Drosophila* resource of transgenic RNAi lines for neurogenetics. *Genetics* **182**, 1089–1100 (2009).
73. G. Dietzl, D. Chen, F. Schnorrer, K.-C. Su, Y. Barinova, M. Fellner, B. Gasser, K. Kinsey, S. Oppel, S. Scheiblauer, A. Couto, V. Marra, K. Keleman, B. J. Dickson, A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
74. D. D. S. Garbe, W. W. L. Bollinger, A. Vigderman, P. Masek, J. Gertowski, A. Sehgal, A. C. A. Keene, Context-specific comparison of sleep acquisition systems in *Drosophila*. *Biol Open.* **4**, 1558–1568 (2015).
75. C. Pfeiffenberger, B. C. Lear, K. P. Keegan, R. Allada, Locomotor activity level monitoring using the *Drosophila* activity monitoring (DAM) system. *Cold Spring Harb Protoc.* **2010**, pdb.prot5518 (2010).
76. K. Murakami, M. E. M. E. Yurgel, B. A. B. A. Stahl, P. Masek, A. Mehta, R. Heidker, W. Bollinger, R. M. R. M. Gingras, Y. J. Y.-J. Kim, W. W. W. W. Ja, B. Suter, J. R. J. R. Diangelo, A. C. A. C. Keene, Translin is required for metabolic regulation of sleep. *Curr. Biol.* **26**, 972–980 (2016).
77. W. Y. Hwang, Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J.-R. R. J. Yeh, J. K. Joung, Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
78. K. N. Rand, T. Ho, W. Qu, S. M. Mitchell, R. White, S. J. Clark, P. L. Molloy, Headloop suppression PCR and its application to selective amplification of methylated DNA sequences. *Nucleic Acids Res.* **33**, e127 (2005).
79. S. Chen, S. Reichert, C. Singh, G. Oikonomou, J. Rihel, D. A. Prober, Light-dependent regulation of sleep and wake states by prokineticin 2 in zebrafish. *Neuron* **95**, 153–168.e6 (2017).
80. D. A. Prober, J. Rihel, A. A. Onah, R.-J. Sung, A. F. Schier, Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish. *J. Neurosci.* **26**, 13400–13410 (2006).
81. A. Nott, I. R. Holtman, N. G. Coufal, J. C. M. Schlachetzki, M. Yu, R. Hu, C. Z. Han, M. Pena, J. Xiao, Y. Wu, Z. Keulen, M. P. Pasillas, C. O'Connor, C. K. Nickl, S. T. Schafer, Z. Shen, R. A. Rissman, J. B. Brewer, D. Gosselin, D. D. Gonda, M. L. Levy, M. G. Rosenfeld, G. McVicker, F. H. Gage, B. Ren, C. K. Glass, Brain cell type-specific enhancer-promoter interactome maps and disease - risk association. *Science* **366**, 1134–1139 (2019).

Acknowledgments

Funding: This work was supported by NIH grants 5R01HL143790 to A.C.K., S.F.A.G., and P.G.; 1R21NS122166 to A.C.K.; and K99AB071833 to E.B.B. S.F.A.G. is also supported by NIH grants R01 HD056465, R01 AG057516, and the Daniel B. Burke Endowed Chair for Diabetes Research. A.C. was supported by NIH grant R35 HG011959. A.I.P. was supported by P01 HL094307 and T32 HL07953. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 18 March 2022
Accepted 5 December 2022
Published 6 January 2023
10.1126/sciadv.abq0844