Zebrafish screen of high-confidence targets at insomnia GWAS loci implicates MEIS1, SKIV2L, and ARFGAP2 as conserved regulators of sleep-wake behaviors

Amber J. Zimmerman[1,4], Fusun Doldur-Balli[4], Brendan T. Keenan[4], Zoe Y. Shetty[4], Justin Palermo[2], Alessandra Chesi[1,3], Shilpa Sonti[1], Matthew C. Pahl[1], Elizabeth B. Brown[2], James A. Pippin[1], Andrew D. Wells[1,3,4], Olivia J. Veatch[5], Diego R. Mazzotti[6,7], Anitra Krishnan[4], Phillip R. Gehrman[8], and Alex C. Keene[2], Struan F.A. Grant[1,9,10,11#], Allan I. Pack[4#]

[#]Denotes equal contribution

[1]Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA

[2]Department of Biology, Texas A&M University, College Station, TX, United States.

[3]Department of Pathology and Laboratory Medicine University of Pennsylvania Perelman School of Medicine Philadelphia PA USA.

[4]Division of Sleep Medicine/Department of Medicine, Perelman School of Medicine, University of Pennsylvania, PA, United States

[5]Department of Psychiatry and Behavioral Sciences, University of Kansas Medical Center, Kansas City, KS, United States

[6]Division of Medical Informatics, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS, United States

[7]Division of Pulmonary Critical Care and Sleep Medicine, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS, United States

[8]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, PA, United States

[9]Department of Pediatrics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104, USA

[10]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA

[11]Department of Genetics, University of Pennsylvania, Philadelphia, PA, 19104, USA

Abstract

Study objectives: Insomnia is a pervasive sleep disorder with heritability estimates ranging from 22-59 percent. Human genome-wide association studies (GWAS) have identified more than 200 genome-wide significant loci associated with the pathogenesis of insomnia; however, few of the implicated effector genes at these loci have undergone functional characterization, and little is known of their relevance to sleep dysfunction in insomnia. High-throughput phenotyping in model organisms, especially for conserved traits like sleep, provides an opportunity for enhanced gene prioritization to improve mechanistic understanding of this complex polygenic disorder and to yield strong genetic candidates for therapeutic intervention.

Methods: Leveraging existing datasets, we performed a functional screen of candidate insomnia-associated genes in zebrafish. Our previous 3D-genomics approach, which integrated ATAC-seq RNA-seq and high-resolution promoter-focused Capture C in neural progenitor cells, revealed putatively causal variants and corresponding effector genes at insomnia GWAS loci. We went on to report first-pass neuron-specific RNA interference screening in *Drosophila* to implicate high-confidence effector genes at these genetic signals. This current study builds on those initial observations by leveraging CRISPR/Cas9 in a diurnal vertebrate model, namely larval zebrafish (5-7 days post fertilization), to perform functional validation of six top candidates to determine conserved regulators of sleep behaviors relevant to insomnia.

Results: CRISPR mutation of three high-confidence insomnia-implicated effector genes (*MEIS1*, *SKIV2L*, and *ARFGAP2*) resulted in robust changes to sleep and activity patterns in larval zebrafish. Specifically, we observed phenotypes that reflect insomnia-like behaviors, including altered sleep duration, poor sleep consolidation, and impaired latency to sleep onset after lights off. Moreover, mutation of *ARFGAP2* also impacted development resulting in a significant motor phenotype.

Conclusions: Our results reveal the utility of CRISPR/Cas9 in F0 larval zebrafish for high-throughput screening of GWAS-implicated candidate effector genes. We validated three of the six strongest candidate effector genes identified in our previous combined 3D Genomic and *D*rosophila screen, two of which were 3D-mapped to different genes rather than the nearest gene reported in the original insomnia GWAS study. Together, we demonstrate the necessity for functional validation of GWAS-implicated effector genes, and we provide evidence for two novel genes producing sleep-wake phenotypes not previously suggested through positional mapping alone.

Introduction

Chronic sleep disruption is linked to a variety of negative health sequelae from metabolic to cognitive dysfunction. Nearly one-third of the adult population reports chronic sleep disturbance and symptoms of insomnia[1]. Insomnia is characterized as (1) difficulty initiating sleep (increased sleep latency), (2) difficulty maintaining sleep with frequent awakenings, (3) excessive daytime sleepiness, and (4) reduced sleep duration given ample opportunity[2, 3]. Insomnia, along with other sleep traits including sleep duration, napping, and daytime dozing, are highly heritable and extremely polygenic[4, 5]. Large-scale genome-wide association studies (GWAS) have been conducted for insomnia and other sleep traits revealing hundreds of genomic loci[4-8]. Publicly available datasets for chromatin accessibility and gene expression have improved gene mapping in these GWAS studies; however, the majority of loci are still positionally mapped to the nearest gene or mapped using *in silico* prediction on aggregate data[7]. These approaches may misidentify the true causal gene at a GWAS locus[9-14], which in turn can lead to mischaracterization of mechanisms underlying insomnia. Detailed fine-mapping of genome-wide significant loci should ideally be carried out in a specific cell-type yielding disease-relevant information to prioritize candidate effector genes[15-19], and should be subsequently validated through functional phenotyping to identify those with the greatest impact on disease pathogenesis.

Model organisms can be leveraged to study sleep traits by assessing changes to sleep architecture relevant to insomnia. High-throughput phenotyping in model organisms can greatly speed gene prioritization for drug discovery and therapeutic development. Zebrafish are well positioned as a model organism to deploy efficient CRISPR/Cas9 mutagenesis paired with high-throughput sleep phenotyping given both their genetic tractability and rapid developmental

timeline[20]. Additionally, zebrafish are vertebrates whose sleep is more representative of human sleep than other model organisms. Zebrafish sleep is diurnal, with the majority of sleep consolidated to the dark phase. It is circadian-regulated, reversable, and has a heightened arousal threshold, making zebrafish an appropriate model for assaying behavior relevant to sleep dysfunction[20, 21]. Moreover, genetic conservation between zebrafish and human is relatively high[22], as they are both vertebrates, and many of the genes already identified to regulate sleep are particularly highly conserved[8].

Our previous work used 3D genomic mapping of GWAS loci to identify effector genes at given loci for insomnia[23]. This approach integrates RNA-seq, assay for transposase-accessible chromatin using sequencing (ATAC-seq) data, and high-resolution promoter-focused Capture C in order to identify physical contacts between putatively causal variants and open promoters at candidate effector genes[16, 17, 19]. This 3D-mapping was previously performed using induced pluripotent stem cell- (iPSC) derived neural progenitor cells (NPCs), as insomnia GWAS signals were significantly enriched in the genomic features for this cell type[23]. In a subsequent high-throughput screen, we functionally phenotyped these candidate effector genes using RNA interference (RNAi) in *Drosophila* to identify those that produced a strong sleep phenotype. This work produced a refined candidate gene list identifying highly conserved regulators of sleep function that are relevant to human insomnia, including *SKIV2L, GNB3, CBX1, MEIS1, TCF12 and ARFGAP2*. Of these genes, only *MEIS1* has received a degree of attention and has been functionally connected to a behavioral phenotype reminiscent of insomnia [5, 6, 24](ADD).

The current study leverages these data to further refine our list of candidate effector genes to best prioritize those with robust changes to sleep behaviors in a vertebrate model. We elected to focus on these six candidates as they produced the strongest sleep phenotypes in

Drosophila, are highly conserved, and have multiple lines of evidence supporting each as a causal effector gene at the respective GWAS locus. This multi-pronged approach has yielded a refined list of high-confidence effector genes with strong functional relevance to insomnia that warrant prioritization for future modeling and therapeutic discovery.

Methods

Animal use

All experiments with zebrafish were conducted in accordance with the University of Pennsylvania Institutional Animal Care and Use Committee guidelines. Breeding pairs consisted of wild-type AB and TL (Tupfel Long-fin) strains. Fish were housed in standard conditions with 14:10 hr light:dark cycle at 28.5°C, with lights on at 9 a.m. and off at 11 p.m.

CRISPR/Cas9 mutagenesis

Single guide RNAs (sgRNAs) were designed using the online tool Crispor (http://crispor.tefor.net/) with the reference genome set to “NCBI GRCz11” and the protospacer adjacent motif (PAM) set to “20bp-NGG-Sp Cas9, SpCas9-HF, eSpCas9 1.1.” sgRNAs were prioritized by specificity score (>95%) with 0 predicted off-targets with up to 3 mismatches. The zebrafish sequence was obtained using Ensembl (https://useast.ensembl.org/) with GRCz11 as the reference genome. Sequence was aligned to the human amino acid sequence using MARRVEL (http://marrvel.org/) to identify the region with highest conservation, and each sgRNA was designed around this exonic region. AB/TL breeding pairs were mated overnight and embryos collected the following morning at lights-on. Pre-formed ribonuclear protein (RNP) complexes containing the sgRNA and Cas9 enzyme were injected at the single-cell stage alternating between the gene group and scramble-injected negative control group. Embryos were left unperturbed for one day before being transported to embryonic growth media ((E3 medium; 5mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4) in petri dishes (approximately 50 per dish). All embryos and larvae were housed in an incubator at 28.5°C, with lights on at 9 a.m. and off at 11 p.m. Dead embryos and chorion membranes were removed daily until day 5 post fertilization (5dpf). On day 5, CRISPR mutants and scramble-injected controls were pipetted into individual wells of a 96-well plate and placed into a zebrabox

(Viewpoint Life Sciences) for automated video monitoring. Genotypes were placed into alternating rows to minimize location bias within the plate. Each zebrabox is sound-attenuating and contains circulating water held at a temperature of 28.5ºC with automated lights cycling on the same 14:10hr light/dark cycle. Sleep-wake behaviors were measured through automated video-tracking as described previously[23, 25].

DNA extraction and PCR for genotyping

DNA extraction was performed per the manufacturer's protocol (Quanta bio, Beverly, MA) 610 immediately following completion of the sleep assay as described previously[23]. Larvae were euthanized by rapid cooling on a mixture of ice and water between 2-4°C for a minimum of 30 minutes after complete cessation of movement was observed. Genotyping was performed on individual fish at the conclusion of each sleep assay. Either restriction digest or headloop PCR methods were used to validate mutations[23, 25, 26]. Primers for genotyping are listed in Supplementary table 1. All primers were run on a 2% agarose gel and sequence verified using Sanger sequencing to verify the target region.

Data collection and analysis for sleep phenotyping

Activity data were captured using automated video tracking (Viewpoint Life Sciences) software in quantization mode[23]. Threshold for detection was set as following: detection threshold: 20; burst: 29; freeze: 3; bin size: 60sec as described previously[27]. Data were processed using custom MATLAB scripts[28] to calculate the following parameters for both day and night separately: sleep duration (min/hour), activity duration (seconds/hour), waking activity (seconds/awake minute/hour), sleep bout length (min/bout), sleep bout number (number/hour) and nighttime sleep latency (min). All animals were allowed to acclimate to the box for

approximately 24 hours before beginning continuous data collection for 48 hours beginning at lights-on.

Arousal Threshold Assay

A mechano-acoustic stimulus was used to determine arousal threshold using a protocol adapted from previous work[29, 30]. Individual fish were placed into alternating columns of a 96-well plate (**Supplementary Fig. 3A**) to avoid location bias. Ten different vibration frequencies were applied that consistently produced a step-wise increase in arousability. Frequencies were pseudorandomly assigned to prevent acclimation to any given stimulus frequency throughout the trials. Frequency steps of 40Hz were ordered as follows: 560Hz, 400Hz, 520Hz, 720Hz, 440Hz, 680Hz, 480Hz, 760Hz, 600Hz, 640Hz. These ten frequencies were each presented ten times for 5 seconds every 3 minutes (5 seconds on, 2 mins 55 seconds off) beginning at 1 a.m. and ending at 6 a.m (**Supplementary Fig. 3B**). Lower frequencies produced larger changes in movement and an increase in the fraction of responsive larvae. Therefore, analyses were presented as highest-to-lowest frequency representing lowest-to-highest intensity of stimulation (i.e. 760Hz, 720Hz, 680Hz, 640Hz, 600Hz, 560Hz, 520Hz, 480Hz, 440Hz, 400Hz). The response to stimuli was measured by automated video tracking and analyzed using Matlab (Mathworks) and Excel (Microsoft). The response fraction represents the percentage of larvae considered to be asleep during the 5 second pre-stimulus baseline (activity <0.01s/5sec, determined empirically) and whose activity increased over baseline during the stimulus presentation. Curve-fit and statistical analyses were carried out in Prism V9 (Graphpad) using non-linear regression (Agonist vs Response--Variable slope (four parameters) with extra sum-of-squares F-test used to compare EC50 (half-maximal response) and top (maximal response).

Statistical Analyses

Statistical analysis and control for multiple comparisons across sleep traits

Continuous data are summarized using means and standard deviations. Phenotypes of interest included a total of 11 measurements related to sleep and activity, including day and night sleep duration, activity, waking activity, sleep bout length, and sleep bout number, as well as nocturnal sleep latency. Analyses compared these traits between scramble-injected and CRIPSR mutant fish for six different candidate genes – *skiv2l, meis1b*, *gnb3a*, *cbx1b*, *tcf12*, and *arfgap2.* Primary comparisons between mutant and scramble-injected fish were performed using non-parametric Wilcoxon rank-sum tests to mitigate any potential impact of non-normality of the endpoints. A Hochberg step-up procedure[31, 32] was applied independently to analysis of each gene to maintain gene-specific type I error at the desired level of 0.05 across the tested hypotheses. Briefly, to implement the Hochberg method the p-values for the set of eleven null hypotheses are ordered from largest to smallest, and each p-value is compared to a sequentially decreasing alpha-level to determine whether the null hypothesis (and subsequent hypotheses) should be rejected. Symbolically, for the set of p-values $\{p_1, \ldots p_{11}\}$, ordered from largest to smallest and testing the corresponding set of null hypotheses $\{H_{o1}, \ldots H_{o11}\}$, the procedure is implemented as:

Step 1: Evaluate whether $p_1 < 0.05$. If yes, reject H_{o1} and all subsequent null hypotheses $\{H_{o2}, \ldots H_{o11}\}$. Else, do not reject H_{o1} and go to Step 2.

Step 2: Evaluate whether $p_2 < 0.05/2$. If yes, reject H_{o2} and all subsequent null hypotheses $\{H_{o3}, \ldots H_{o11}\}$. Else, do not reject H_{o2} and go to Step 3.

[…]

Step 11: Evaluate whether $p_k < 0.05/11$. If yes, reject H_{o11}. Else, none of the null hypotheses $\{H_{o1}, \ldots H_{o11}\}$ are rejected and stop.

Results

Identification of high-confidence insomnia effector genes

Genes assessed in this study were previously implicated through our established variant-to-gene mapping approach[15-19, 23], which integrates chromatin conformation and accessibility datasets in relevant cell types to infer causal effector genes at GWAS loci. Promoters of the candidate effector genes used in this study reside in open chromatin and display high resolution chromatin contacts with putative insomnia causal variants associated with significant GWAS loci (**Fig. 1A**) in neural progenitor cells (NPCs)[23]. These genes are highly expressed in NPCs and demonstrate high conservation across human, zebrafish, and *Drosophila,* making them high-priority candidates for functional assessment. Our previous work performed high-throughput screening of candidate effector genes using *Drosophila* RNAi, revealing a subset of genes producing exceptionally strong sleep phenotypes[23] (**Fig. 1A, right panel**), of which we chose to perform functional follow-up using a zebrafish model. These genes were *SKIV2L*, *GNB3*, *CBX1*, *MEIS1*, *TCF12*, and *ARFGAP2*.

To test the hypothesis that these high-confidence insomnia effector genes produce sleep-specific phenotypes in a vertebrate model, we employed CRISPR/Cas9 mutagenesis in F0 embryos at the single-cell stage followed by rapid behavioral screening of larval zebrafish harboring mutations 5 to 7 days post fertilization (dpf) (**Fig. 1B**). CRISPR mutagenesis at the single-cell stage produced highly-efficient disruption of the target gene in greater than 90 percent of cells as previously demonstrated[23, 25]. Single guide RNAs (sgRNAs) were designed to target the exonic sequence with highest conservation between zebrafish and human. If multiple transcripts existed, exons overlapping across transcripts were prioritized. Given several of these genes are duplicated in the zebrafish genome, the best predicted ortholog[33] was targeted for subsequent experiments.

Larval zebrafish carrying mutations in the target genes were monitored using automated video tracking for two consecutive days and nights to capture sleep and activity phenotypes across multiple circadian cycles[23, 25]. Multiple sleep and activity traits relevant to insomnia were assessed to identify subtle changes in sleep architecture as well as sleep duration. We measured eleven traits in total for each larval zebrafish group; these included: average hourly night and day sleep duration, average hourly activity across night and day, average waking activity (activity only during minutes scored as "awake"), average sleep bout number during day and night, average sleep bout length during day and night, and average nighttime sleep latency. Non-parametric Wilcoxon rank sum tests were performed to assess genotype differences relative to control across all eleven traits followed by a Hochberg step-up procedure[31, 32] to determine threshold for significance across all tested hypotheses. Of the six genotypes examined, mutations within *meis1b*, *skiv2l*, and *arfgap2* resulted in significant phenotypes across several sleep traits – described in detail below.

Total sleep duration is perturbed in CRISPR mutants.

We measured eleven sleep-specific metrics that characterize several domains of sleep architecture. Shortened sleep duration given ample sleep opportunity is one measure relevant to human insomnia; therefore, we assessed total sleep duration across the fourteen-hour day and ten-hour night. Diurnal sleep-wake patterns were intact in all mutants (**Fig. 2A-C**). Total sleep duration during the day was significantly reduced in larvae harboring mutations in *skiv2l* (mean diff. = 71.5 mins per day, $P = 0.0001$) (**Fig. 2D**), with a modest reduction in night sleep duration (mean diff. = 26.5 mins per night, $P < 0.1$) (**Fig. 2E**). Both day (mean diff. = 365.3 min per day, $P < 0.0001$) (**Fig. 2H**) and night (mean diff. = 83.7 min per night, $P < 0.0001$) (**Fig. 2I**) sleep duration were dramatically increased in *arfgap2* mutants, while mutations in *meis1b*, *gnb3a*, *tcf12*, and *cbx1b* mutants showed modest-to-little changes in total sleep duration ($P > 0.05$) (**Fig. 2F-G** and **Supplementary Fig. 1A-F**).

Another key feature of insomnia is difficulty initiating sleep with an increase in latency to sleep onset. We measured sleep latency at night, which is calculated as the time from lights-off to the first sleep bout. Mutations in *skiv2l* produced an increase in sleep latency (mean diff. = 1.43 min, P = 0.02) (**Fig. 2J**); however, this difference did not strictly meet the threshold for significance following correction. Consistent with its implied role in insomnia, *meis1b* mutants showed a significantly increased sleep latency (mean diff. = 1.12 min, P = 0.0004) (**Fig. 2K**). Fitting the hypersomnia phenotype, *arfgap2* mutants showed a trend towards reduced sleep latency relative to controls (mean diff. = 1.4, P = 0.08), but was not statistically significant (**Fig. 2L**). Sleep latency was not disrupted in *gnb3a*, *tcf12*, or *cbx1b* mutants (**Supplementary Fig. 2A-C**).

Diurnal activity patterns are impacted by insomnia-associated genes.

Given that insomnia is often attributed to hypervigilance[34] and hyperarousal[35, 36], we measured a variety of characteristics that capture changes in general activity patterns and nighttime arousal. Activity patterns showed robust entrainment by light-dark cycles as expected across mutants (**Fig. 3A-C**). Consistent with findings of reduced sleep duration, we observed significantly increased daytime activity duration in *skiv2l* mutants (mean diff. = 20.6 s/h, P = 0.005) (**Fig. 3D**) with no change to night activity (**Fig. 3E**). Additionally, *arfgap2* mutants revealed significantly reduced day (mean diff. = 64.9 s/h, P < 0.0001) (**Fig. 3H**) and nighttime activity (mean diff. = 5.6 s/h, P = 0.004) (**Fig. 3I**). Similar to the lack of sleep duration phenotype in *meis1b* mutants, no global change in activity was observed during the day (**Fig. 3F**) or night (**Fig. 3G**) (P > 0.05) in these fish. No significant activity phenotypes were observed for *gnb3*, *tcf12*, or *cbx1b* mutants (**Supplementary Fig. 1G-L**). We chose to measure arousal threshold using mechano-acoustic stimuli of increasing intensity (**Supplementary Fig. 3A-B**), a metric of sleep depth, in *skiv2l*, *meis1*, and *arfgap2* mutants since they showed robust sleep and activity changes. We observed

no significant changes to arousal threshold in these animals, measured as the stimulus intensity required to elicit the half maximal response or in their maximal response to stimuli of increasing intensity (**Supplementary Fig. 3C-E**).

Insomnia-associated genes contribute to changes in sleep architecture.

Total sleep duration is not always impacted in patients with insomnia, rather sleep is fragmented or considered not restorative leading to excessive daytime sleepiness[3]. We measured sleep bout length (the average length of each sleep period) during night and day to observe changes to sleep depth across mutant lines. Furthermore, we measured sleep bout number (total number of sleep episodes) during day and night as a representation of sleep fragmentation. Consistent with reduced daytime sleep in *skiv2l* mutants, these larvae also demonstrated a reduction in sleep bout number during the day (mean diff. = 1.75 bouts/h, $P = 0.0001$) (**Fig. 4A**), with no change in bout number at night (**Fig. 4B**). Although *meis1b* mutants did not present with a generalized sleep duration abnormality, we observed a significant increase in sleep bout number in these fish at night (mean diff. = 1.33 bouts/h, $P < 0.001$) (**Fig. 4D**), with no change during the day (**Fig. 4C**), indicating sleep fragmentation caused by a gene that is commonly associated with restless legs syndrome (RLS)[37, 38]. *Arfgap2* mutants, which showed drastic increases in total sleep duration, also displayed a significant increase in the number of daytime sleep bouts (mean diff. = 4.11 bouts/h, $P < 0.0001$).) (**Fig. 4E**). No change was observed for nighttime sleep bout number in *arfgap2* mutants (**Fig. 4F**).

While no significant change in sleep bout length was observed for *skiv2l* mutants (**Fig. 4G and H**), nighttime sleep bout length was significantly shortened (mean diff. = 1.49 min/bout, $P = 0.002$) in *meis1b* mutants (**Fig. 4J**), with no changes observed during the day (**Fig 4I**), suggesting depth of sleep is also impacted by this gene. *Arfgap2* mutants showed highly consolidated sleep as marked by an increase in daytime sleep bout length (mean diff. = 2.85

min/bout, $P < 0.0001$) (**Fig. 4K**) and an increase in sleep bout length at night (mean diff. = 3.68 min/bout, $P = 0.002$) (**Fig. 4L**). No significant changes were observed for these sleep characteristics with the *gnb3a*, *tcf12* or *cbx1b* mutants (**Supplementary Fig. 4A-L**). Together, these data support a role for *meis1b*, *skiv2l*, and *arfgap2* in promoting sleep disruption primarily through altering sleep fragmentation and consolidation.

Mutation in insomnia-associated gene, *ARFGAP2*, leads to significant developmental abnormalities in zebrafish.

Many genes influencing insomnia are also involved in neuronal development and can differentially influence behavior across developmental stages. Larval zebrafish develop an intact nervous system within the first few days of development and gene mutations that disrupt neurodevelopment lead to apparent changes in body morphology[20]. We observed a clear developmental phenotype in *arfgap2* mutants with gross changes to waking activity (**Fig. 5A**) and morphological abnormalities, whereby nearly all mutants presented with a strong curvature in the tail beginning on approximately days 3-4 post fertilization (**Fig. 5B**). We calculated waking activity as movement only during minutes that were scored as “awake” (movement duration greater than 0.5sec/min). Waking activity serves as a proxy for general movement disruptions and may indicate gross motor changes following genetic manipulation. While these mutants had robust “sleep” phenotypes, they also had significant reductions in total activity and, specifically, daytime waking activity (mean diff. = 64.9 s/h, $P < 0.0001$) (**Fig. 5C**), indicating the sleep phenotypes observed are predominantly driven by motor defects manifesting as dramatically reduced movement. No significant differences were observed for night waking activity (**Fig. 5D**) for *arfgap2* mutants, and no significant changes were found in waking activity of all other mutants (**Supplementary Fig. 5A-L**), suggesting that the sleep phenotypes are not driven by gross motor abnormalities in these fish.

Discussion

There is a wealth of genomic data available through public repositories generated from GWAS and other sequencing approaches; however, functional characterization lags in validating the actual underlying genomic factors contributing to different phenotypes. Here we demonstrate a proof-of-principle approach for moving from GWAS-implicated effector genes to validation in a vertebrate model organism within a high-throughput setting.

We elected to perform functional validation of our top candidate genes identified via our established variant-to-gene mapping pipeline[23]. The candidate effector genes we screened were previously mapped to insomnia GWAS-associated loci using ATAC-seq and promoter-focused Capture C protocols. Through a large-scale neuron-specific RNAi screen in adult *Drosophila melanogaster*, loss of function of some of these genes was shown to produce robust sleep phenotypes demonstrating their high conservation and potential regulatory function in sleep. We chose to focus on six genes which are highly conserved across species and produced strong sleep phenotypes in *Drosophila* (*MEIS1*, *CBX1*, *TCF12, ARFGAP2, SKIV2L*, *and GNB3)*.

Increasingly, studies have shown that total sleep duration is less predictive of negative health outcomes than altered sleep architecture and poor sleep quality[39, 40]. This includes day-to-day variability in sleep duration, sleep onset and waking time, as well as sleep fragmentation and excessive daytime sleepiness. Zebrafish provide a model system to observe these nuanced behaviors absent of external influences. For example, we observed more robust alterations in daytime sleep characteristics in several mutant lines including increased daytime sleep duration. This behavior may reflect excessive daytime sleepiness reported in humans and could implicate a role for these genes in a more complex sleep architecture. Moreover, it may be difficult to

capture changes in total sleep duration at night due to a ceiling effect whereby fish spend the majority of the dark cycle in a quiescent, sleep-like state.

We observed that CRISPR mutation of *meis1b*, the strongest ortholog for human *MEIS1*, results in severely fragmented nighttime sleep as well as increased sleep latency. *MEIS1* is highly conserved with 95 percent amino acid identity between zebrafish and humans[33]. This gene encodes the Myeloid Ecotrpoic Viral Integration Site 1 protein, which acts as a transcription factor[41]. The *MEIS1* locus is one of the strongest association signals from previous insomnia GWAS[5, 7, 8]. Our variant 3D-mapping approach identified a putative causal variant in strong linkage disequilibrium (LD) with the sentinel GWAS SNP, rs1519102, at this locus, which mapped to the *MEIS1* promoter residing in open chromatin within NPCs[23]. rs1519102 resides within an intronic region, which suggests it is harbored in a cis-regulatory element acting as a transcriptional enhancer in a cell-specific manner[42]. Work by Lam and colleagues[42] further identified expression quantitative trait loci (eQTL) residing within this region specific to brain cell types, including within the cerebellum, one of the few brain regions where *MEIS1* is highly expressed. It has repeatedly been identified as a candidate gene for insomnia [5, 6, 8]; however, there is debate as to whether its role in sleep disturbance is primarily due to its association with RLS[7, 37]. *MEIS1* knockouts have shown hyperactive phenotypes in animal models[43, 44], but it is unclear if this phenotype translates to sleep. Our model reveals a nocturnal-specific phenotype suggestive of fragmented sleep-in line with the potential role in RLS, which predominantly occurs at night, and may demonstrate that MEIS1 is acting in a circadian pattern to alter arousal and sleep consolidation.

Our mutation of *skiv2l* in zebrafish significantly reduced sleep duration and modestly increased sleep latency. In addition, these mutants showed daytime hyperactivity, suggesting a state of hyperarousal consistent with an insomnia-like condition. *SKIV2L* encodes the Superkiller Viralicilic Activity 2-Like RNA helicase and was 3D-mapped to the *MICB* GWAS locus (sentinel SNP rs3131638) on chromosome 6[23]. It is located within a conserved region of the major histocompatibility complex[45, 46] and is believed to play a role in antiviral activity by blocking translation of poly(A) deficient mRNAs[46]. *SKIV2L* is fairly highly conserved between zebrafish and humans with 60 percent overall amino acid conservation and greater than 90 percent conservation of the specific sequence encoding the DEAD box helicase[33] suggesting the role in RNA metabolism is particularly crucial. The mechanism by which *SKIV2L* acts to influence sleep is unknown; however, RNA metabolism is circadian-regulated and disturbed sleep has been shown to be associated with alterations in RNA metabolism[47]. Loss-of-function in *SKIV2L* produced extremely robust sleep phenotypes in both flies[23] and fish, suggesting strong conservation of function related to this gene.

ARFGAP2 encodes the ADP Ribosylation Factor GTPase Activating Protein 2 and was 3D-mapped to the *NDUFS3* GWAS locus (sentinel SNP rs11605348) on chromosome 11[23]. It is 66 percent conserved at the amino acid level between human and zebrafish[33] and is highly expressed early in development[48], likely owing to the developmental phenotype observed in our larval model. *ARFGAP2* has not been extensively studied in the context of behavioral characteristics; however, polymorphisms within *ARFGAP2* have been associated with other neurocognitive disorders including depression[49] and Alzheimer's disease[50], which have a sleep component. The fish harboring mutations in *arfgap2* appeared normal through early development (1-2dpf); however, by the third day, approximately half of the larvae began to appear abnormal with a curved tail. By 5dpf, when the sleep assay began, nearly all mutant

larvae appeared developmentally abnormal; however, very few died. *Arfgap2* expression peaks during this time frame and likely serves an important role during a critical period of development.

Limitations

Several of these genes exhibit differential expression across development. In these experiments, we assayed larval zebrafish, while in contrast our previous work tested sleep in adult *Drosophila*. Knockdown of the *arfgap2* ortholog in *Drosophila* produced a significant reduction in sleep duration, yet the CRISPR mutation in zebrafish produced abnormal wake and sleep patterns paired with developmental abnormalities marked by abnormal tail curvature. The differences in phenotypes observed between larval zebrafish and adult *Drosophila* with perturbed *arfgap2* expression is likely due to developmental expression differences, demonstrating the benefit of cross-species and cross-development observation. Moreover, our experiments in *Drosophila*[23] used neuron-specific RNA-interference to knockdown gene expression in neuronal cells, sparing expression elsewhere. This, too, likely contributed to the phenotypic differences observed.

Using F0 larvae for screening is an incredibly rapid and efficient approach for assaying sleep behavior to narrow long lists of candidate genes; however, it has inherent limitations. While we found that extremely little mosaicism was apparent in our larvae harboring the CRISPR mutations, there was a high degree of variability in behavior of both mutants and controls. The variability is possibly due to trauma induced by the injections, which is why we used scramble-sgRNA-injected controls for comparison. This variability likely diminishes true sleep-activity phenotypes and may result in false-negative results. Given we previously observed robust sleep phenotypes in *Drosophila* RNAi lines for these genes[23], we cannot rule out those candidate

genes that demonstrated minimal phenotypes. Rather, future studies assessing F1 mutant lines represent a promising tool to assess these genes. Additionally, several of these genes are duplicated in zebrafish and one isoform may offer compensatory action over the mutated isoform. We chose to focus on the isoform with strongest conservation across species as we hypothesized that this would be most consistent with the *Drosophila* screen; however, double knockouts are warranted in the future to assess for stronger phenotypes.

Acknowledgements

The work was supported by NIH grant R01 HL143790. Dr. Grant is supported by NIH awards R01 AG057516 and R01 HD056465 and the Daniel B. Burke Endowed Chair.

Disclosure Statement

The authors have nothing to disclose.

References

1. Stranges S, Tigbe W, Gomez-Olive FX, Thorogood M, Kandala NB. 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. 2012;35(8):1173-81. Epub 20120801. doi: 10.5665/sleep.2012. PubMed PMID: 22851813; PMCID: PMC3397790.
2. Morin CM, Drake CL, Harvey AG, Krystal AD, Manber R, Riemann D, Spiegelhalder K. Insomnia disorder. Nat Rev Dis Primers. 2015;1:15026. Epub 2015/01/01. doi: 10.1038/nrdp.2015.26. PubMed PMID: 27189779.
3. Association AP. Diagnostic and statistical manual of mental disorders 5th ed2013.
4. Dashti HS, Jones SE, Wood AR, Lane JM, van Hees VT, Wang H, Rhodes JA, Song Y, Patel K, Anderson SG, Beaumont RN, Bechtold DA, Bowden J, Cade BE, Garaulet M, Kyle SD, Little MA, Loudon AS, Luik AI, Scheer F, Spiegelhalder K, Tyrrell J, Gottlieb DJ, Tiemeier H, Ray DW, Purcell SM, Frayling TM, Redline S, Lawlor DA, Rutter MK, Weedon MN, Saxena R. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. Nat Commun. 2019;10(1):1100. Epub 2019/03/09. doi: 10.1038/s41467-019-08917-4. PubMed PMID: 30846698; PMCID: PMC6405943.
5. Lane JM, Liang J, Vlasac I, Anderson SG, Bechtold DA, Bowden J, Emsley R, Gill S, Little MA, Luik AI, Loudon A, Scheer FA, Purcell SM, Kyle SD, Lawlor DA, Zhu X, Redline S, Ray DW, Rutter MK, Saxena R. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. Nat Genet. 2017;49(2):274-81. Epub 2016/12/20. doi: 10.1038/ng.3749. PubMed PMID: 27992416; PMCID: PMC5491693.
6. Hammerschlag AR, Stringer S, de Leeuw CA, Sniekers S, Taskesen E, Watanabe K, Blanken TF, Dekker K, Te Lindert BHW, Wassing R, Jonsdottir I, Thorleifsson G, Stefansson H, Gislason T, Berger K, Schormair B, Wellmann J, Winkelmann J, Stefansson K, Oexle K, Van Someren EJW, Posthuma D. Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. Nat Genet. 2017;49(11):1584-92. Epub 2017/06/13. doi: 10.1038/ng.3888. PubMed PMID: 28604731; PMCID: PMC5600256.
7. Watanabe K, Jansen PR, Savage JE, Nandakumar P, Wang X, andMe Research T, Hinds DA, Gelernter J, Levey DF, Polimanti R, Stein MB, Van Someren EJW, Smit AB, Posthuma D. Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and psychiatric pathways. Nat Genet. 2022;54(8):1125-32. Epub 2022/07/15. doi: 10.1038/s41588-022-01124-w. PubMed PMID: 35835914.

8. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, de Leeuw CA, Benjamins JS, Munoz-Manchado AB, Nagel M, Savage JE, Tiemeier H, White T, andMe Research T, Tung JY, Hinds DA, Vacic V, Wang X, Sullivan PF, van der Sluis S, Polderman TJC, Smit AB, Hjerling-Leffler J, Van Someren EJW, Posthuma D. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. Nat Genet. 2019;51(3):394-403. Epub 2019/02/26. doi: 10.1038/s41588-018-0333-3. PubMed PMID: 30804565.

9. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467-84. Epub 2019/05/10. doi: 10.1038/s41576-019-0127-1. PubMed PMID: 31068683.

10. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, Kathiresan S, Kenny EE, Lindgren CM, MacArthur DG, North KN, Plon SE, Rehm HL, Risch N, Rotimi CN, Shendure J, Soranzo N, McCarthy MI. A brief history of human disease genetics. Nature. 2020;577(7789):179-89. Epub 2020/01/10. doi: 10.1038/s41586-019-1879-7. PubMed PMID: 31915397; PMCID: PMC7405896.

11. Lappalainen T, MacArthur DG. From variant to function in human disease genetics. Science. 2021;373(6562):1464-8. Epub 2021/09/24. doi: 10.1126/science.abi8207. PubMed PMID: 34554789.

12. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gomez-Marin C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviindran V, Tam D, Shen M, Son JE, Vakili NA, Sung HK, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui CC, Gomez-Skarmeta JL, Nobrega MA. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014;507(7492):371-5. Epub 2014/03/22. doi: 10.1038/nature13138. PubMed PMID: 24646999; PMCID: PMC4113484.

13. Fulco CP, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR, Anyoha R, Doughty BR, Patwardhan TA, Nguyen TH, Kane M, Perez EM, Durand NC, Lareau CA, Stamenova EK, Aiden EL, Lander ES, Engreitz JM. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. Nat Genet. 2019;51(12):1664-9. Epub 20191129. doi: 10.1038/s41588-019-0538-0. PubMed PMID: 31784727; PMCID: PMC6886585.

14. Forgetta V, Jiang L, Vulpescu NA, Hogan MS, Chen S, Morris JA, Grinek S, Benner C, Jang DK, Hoang Q, Burtt N, Flannick JA, McCarthy MI, Fauman E, Greenwood CMT, Maurano MT, Richards JB. An effector index to predict target genes at GWAS loci. Hum Genet. 2022;141(8):1431-47. Epub 20220211. doi: 10.1007/s00439-022-02434-z. PubMed PMID: 35147782.

15. Lasconi C, Pahl MC, Pippin JA, Su C, Johnson ME, Chesi A, Boehm K, Manduchi E, Ou K, Golson ML, Wells AD, Kaestner KH, Grant SFA. Variant-to-gene-mapping analyses reveal a role for pancreatic islet cells in conferring genetic susceptibility to sleep-related traits. Sleep. 2022;45(8). Epub 2022/05/11. doi: 10.1093/sleep/zsac109. PubMed PMID: 35537191; PMCID: PMC9366645.

16. Pahl MC, Doege CA, Hodge KM, Littleton SH, Leonard ME, Lu S, Rausch R, Pippin JA, De Rosa MC, Basak A, Bradfield JP, Hammond RK, Boehm K, Berkowitz RI, Lasconi C, Su C, Chesi A, Johnson ME, Wells AD, Voight BF, Leibel RL, Cousminer DL, Grant SFA. Cis-regulatory architecture of human ESC-derived hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits. Nat Commun. 2021;12(1):6749. Epub

2021/11/21. doi: 10.1038/s41467-021-27001-4. PubMed PMID: 34799566; PMCID: PMC8604959.

17. Chesi A, Wagley Y, Johnson ME, Manduchi E, Su C, Lu S, Leonard ME, Hodge KM, Pippin JA, Hankenson KD, Wells AD, Grant SFA. Genome-scale Capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. Nat Commun. 2019;10(1):1260. Epub 2019/03/21. doi: 10.1038/s41467-019-09302-x. PubMed PMID: 30890710; PMCID: PMC6425012.

18. Lasconi C, Pahl MC, Cousminer DL, Doege CA, Chesi A, Hodge KM, Leonard ME, Lu S, Johnson ME, Su C, Hammond RK, Pippin JA, Terry NA, Ghanem LR, Leibel RL, Wells AD, Grant SFA. Variant-to-Gene-Mapping Analyses Reveal a Role for the Hypothalamus in Genetic Susceptibility to Inflammatory Bowel Disease. Cell Mol Gastroenterol Hepatol. 2021;11(3):667-82. Epub 2020/10/19. doi: 10.1016/j.jcmgh.2020.10.004. PubMed PMID: 33069917; PMCID: PMC7843407.

19. Su C, Johnson ME, Torres A, Thomas RM, Manduchi E, Sharma P, Mehra P, Le Coz C, Leonard ME, Lu S, Hodge KM, Chesi A, Pippin J, Romberg N, Grant SFA, Wells AD. Mapping effector genes at lupus GWAS loci using promoter Capture-C in follicular helper T cells. Nat Commun. 2020;11(1):3294. Epub 2020/07/06. doi: 10.1038/s41467-020-17089-5. PubMed PMID: 32620744; PMCID: PMC7335045.

20. Tran S, Prober DA. Validation of Candidate Sleep Disorder Risk Genes Using Zebrafish. Front Mol Neurosci. 2022;15:873520. Epub 2022/04/26. doi: 10.3389/fnmol.2022.873520. PubMed PMID: 35465097; PMCID: PMC9021570.

21. Chiu CN, Prober DA. Regulation of zebrafish sleep and arousal states: current and prospective approaches. Front Neural Circuits. 2013;7:58. Epub 2013/04/12. doi: 10.3389/fncir.2013.00058. PubMed PMID: 23576957; PMCID: PMC3620505.

22. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W, Kilian B, Quintais LT, Guerra-Assuncao JA, Zhou Y, Gu Y, Yen J, Vogel JH, Eyre T, Redmond S, Banerjee R, Chi J, Fu B, Langley E, Maguire SF, Laird GK, Lloyd D, Kenyon E, Donaldson S, Sehra H, Almeida-King J, Loveland J, Trevanion S, Jones M, Quail M, Willey D, Hunt A, Burton J, Sims S, McLay K, Plumb B, Davis J, Clee C, Oliver K, Clark R, Riddle C, Elliot D, Threadgold G, Harden G, Ware D, Begum S, Mortimore B, Kerry G, Heath P, Phillimore B, Tracey A, Corby N, Dunn M, Johnson C, Wood J, Clark S, Pelan S, Griffiths G, Smith M, Glithero R, Howden P, Barker N, Lloyd C, Stevens C, Harley J, Holt K, Panagiotidis G, Lovell J, Beasley H, Henderson C, Gordon D, Auger K, Wright D, Collins J, Raisen C, Dyer L, Leung K, Robertson L, Ambridge K, Leongamornlert D, McGuire S, Gilderthorp R, Griffiths C, Manthravadi D, Nichol S, Barker G, Whitehead S, Kay M, Brown J, Murnane C, Gray E, Humphries M, Sycamore N, Barker D, Saunders D, Wallis J, Babbage A, Hammond S, Mashreghi-Mohammadi M, Barr L, Martin S, Wray P, Ellington A, Matthews N, Ellwood M, Woodmansey R, Clark G, Cooper J, Tromans A, Grafham D, Skuce C, Pandian R, Andrews R, Harrison E, Kimberley A, Garnett J, Fosker N, Hall R, Garner P, Kelly D, Bird C, Palmer S, Gehring I, Berger A, Dooley CM, Ersan-Urun Z, Eser C, Geiger H, Geisler M, Karotki L, Kirn A, Konantz J, Konantz M, Oberlander M, Rudolph-Geiger S, Teucke M, Lanz C, Raddatz G, Osoegawa K, Zhu B, Rapp A, Widaa S, Langford C, Yang F, Schuster SC, Carter NP, Harrow J, Ning Z, Herrero J, Searle SM, Enright A, Geisler R, Plasterk RH, Lee C, Westerfield M, de Jong PJ, Zon LI, Postlethwait JH, Nusslein-Volhard C, Hubbard TJ, Roest Crollius H, Rogers J, Stemple DL. The zebrafish reference genome sequence and its relationship

to the human genome. Nature. 2013;496(7446):498-503. Epub 2013/04/19. doi: 10.1038/nature12111. PubMed PMID: 23594743; PMCID: PMC3703927.
23. Palermo J, Chesi A, Zimmerman A, Sonti S, Lasconi C, Brown EB, Pippin JA, Wells AD, Doldur-Balli F, Mazzotti DR, Pack AI, Gehrman PR, Grant SFA, Keene AC. Variant-to-gene-mapping followed by cross-species genetic screening identifies GPI-anchor biosynthesis as novel regulator of sleep. bioRxiv. 2021:2021.12.19.472248. doi: 10.1101/2021.12.19.472248.
24. Thireau J, Farah C, Molinari N, Bouilloux F, Torreilles L, Winkelmann J, Scholz S, Richard S, Dauvilliers Y, Marmigere F. MEIS1 variant as a determinant of autonomic imbalance in Restless Legs Syndrome. Sci Rep. 2017;7:46620. Epub 2017/04/21. doi: 10.1038/srep46620. PubMed PMID: 28425489; PMCID: PMC5397858.
25. Kroll F, Powell GT, Ghosh M, Gestri G, Antinucci P, Hearn TJ, Tunbak H, Lim S, Dennis HW, Fernandez JM, Whitmore D, Dreosti E, Wilson SW, Hoffman EJ, Rihel J. A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. Elife. 2021;10. Epub 20210108. doi: 10.7554/eLife.59683. PubMed PMID: 33416493; PMCID: PMC7793621.
26. Rand KN, Ho T, Qu W, Mitchell SM, White R, Clark SJ, Molloy PL. Headloop suppression PCR and its application to selective amplification of methylated DNA sequences. Nucleic Acids Res. 2005;33(14):e127. Epub 20050809. doi: 10.1093/nar/gni120. PubMed PMID: 16091627; PMCID: PMC1184225.
27. Chen S, Reichert S, Singh C, Oikonomou G, Rihel J, Prober DA. Light-Dependent Regulation of Sleep and Wake States by Prokineticin 2 in Zebrafish. Neuron. 2017;95(1):153-68 e6. Epub 2017/06/27. doi: 10.1016/j.neuron.2017.06.001. PubMed PMID: 28648499; PMCID: PMC5653285.
28. Lee DA, Oikonomou G, Prober DA. Large-scale Analysis of Sleep in Zebrafish. Bio Protoc. 2022;12(3):e4313. Epub 2022/03/15. doi: 10.21769/BioProtoc.4313. PubMed PMID: 35284597; PMCID: PMC8855086.
29. Singh C, Oikonomou G, Prober DA. Norepinephrine is required to promote wakefulness and for hypocretin-induced arousal in zebrafish. Elife. 2015;4:e07000. Epub 2015/09/17. doi: 10.7554/eLife.07000. PubMed PMID: 26374985; PMCID: PMC4606453.
30. Reichert S, Pavon Arocas O, Rihel J. The Neuropeptide Galanin Is Required for Homeostatic Rebound Sleep following Increased Neuronal Activity. Neuron. 2019;104(2):370-84 e5. Epub 20190916. doi: 10.1016/j.neuron.2019.08.010. PubMed PMID: 31537465.
31. Hochberg Y. A sharper bonferroni procedure for multiple tests of significance: New York University Graduate School of Business Administration.; 1988.
32. Huang Y, Hsu JC. Hochberg's Step-up Method: Cutting Corners off Holm's Step-down Method. Biometrika. 2007;94(4):965-75.
33. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, Mohr SE. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011;12:357. Epub 20110831. doi: 10.1186/1471-2105-12-357. PubMed PMID: 21880147; PMCID: PMC3179972.
34. Chen MC, Chang C, Glover GH, Gotlib IH. Increased insula coactivation with salience networks in insomnia. Biol Psychol. 2014;97:1-8. Epub 20140108. doi: 10.1016/j.biopsycho.2013.12.016. PubMed PMID: 24412227; PMCID: PMC3961550.
35. Fernandez-Mendoza J, Li Y, Vgontzas AN, Fang J, Gaines J, Calhoun SL, Liao D, Bixler EO. Insomnia is Associated with Cortical Hyperarousal as Early as Adolescence. Sleep.

2016;39(5):1029-36. Epub 20160501. doi: 10.5665/sleep.5746. PubMed PMID: 26951400; PMCID: PMC4835300.

36. Kalmbach DA, Cuamatzi-Castelan AS, Tonnu CV, Tran KM, Anderson JR, Roth T, Drake CL. Hyperarousal and sleep reactivity in insomnia: current insights. Nat Sci Sleep. 2018;10:193-201. Epub 20180717. doi: 10.2147/NSS.S138823. PubMed PMID: 30046255; PMCID: PMC6054324.

37. El Gewely M, Welman M, Xiong L, Yin S, Catoire H, Rouleau G, Montplaisir JY, Desautels A, Warby SC. Reassessing GWAS findings for the shared genetic basis of insomnia and restless legs syndrome. Sleep. 2018;41(11). Epub 2018/09/15. doi: 10.1093/sleep/zsy164. PubMed PMID: 30215811.

38. Schulte EC, Kousi M, Tan PL, Tilch E, Knauf F, Lichtner P, Trenkwalder C, Hogl B, Frauscher B, Berger K, Fietze I, Hornyak M, Oertel WH, Bachmann CG, Zimprich A, Peters A, Gieger C, Meitinger T, Muller-Myhsok B, Katsanis N, Winkelmann J. Targeted resequencing and systematic in vivo functional testing identifies rare variants in MEIS1 as significant contributors to restless legs syndrome. Am J Hum Genet. 2014;95(1):85-95. doi: 10.1016/j.ajhg.2014.06.005. PubMed PMID: 24995868; PMCID: PMC4085638.

39. Wallace ML, Stone K, Smagula SF, Hall MH, Simsek B, Kado DM, Redline S, Vo TN, Buysse DJ, Osteoporotic Fractures in Men Study Research G. Which Sleep Health Characteristics Predict All-Cause Mortality in Older Men? An Application of Flexible Multivariable Approaches. Sleep. 2018;41(1). doi: 10.1093/sleep/zsx189. PubMed PMID: 29165696; PMCID: PMC5806578.

40. Martin JL, Fiorentino L, Jouldjian S, Mitchell M, Josephson KR, Alessi CA. Poor self-reported sleep quality predicts mortality within one year of inpatient post-acute rehabilitation among older adults. Sleep. 2011;34(12):1715-21. Epub 20111201. doi: 10.5665/sleep.1444. PubMed PMID: 22131610; PMCID: PMC3208850.

41. Moskow JJ, Bullrich F, Huebner K, Daar IO, Buchberg AM. Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. Mol Cell Biol. 1995;15(10):5434-43. doi: 10.1128/MCB.15.10.5434. PubMed PMID: 7565694; PMCID: PMC230793.

42. Lam DD, Antic Nikolic A, Zhao C, Mirza-Schreiber N, Krezel W, Oexle K, Winkelmann J. Intronic elements associated with insomnia and restless legs syndrome exhibit cell-type-specific epigenetic features contributing to MEIS1 regulation. Hum Mol Genet. 2022;31(11):1733-46. doi: 10.1093/hmg/ddab355. PubMed PMID: 34888668.

43. Salminen AV, Garrett L, Schormair B, Rozman J, Giesert F, Niedermeier KM, Becker L, Rathkolb B, Racz I, German Mouse Clinic C, Klingenspor M, Klopstock T, Wolf E, Zimmer A, Gailus-Durner V, Torres M, Fuchs H, Hrabe de Angelis M, Wurst W, Holter SM, Winkelmann J. Meis1: effects on motor phenotypes and the sensorimotor system in mice. Dis Model Mech. 2017;10(8):981-91. Epub 20170623. doi: 10.1242/dmm.030080. PubMed PMID: 28645892; PMCID: PMC5560065.

44. Spieler D, Kaffe M, Knauf F, Bessa J, Tena JJ, Giesert F, Schormair B, Tilch E, Lee H, Horsch M, Czamara D, Karbalai N, von Toerne C, Waldenberger M, Gieger C, Lichtner P, Claussnitzer M, Naumann R, Muller-Myhsok B, Torres M, Garrett L, Rozman J, Klingenspor M, Gailus-Durner V, Fuchs H, Hrabe de Angelis M, Beckers J, Holter SM, Meitinger T, Hauck SM, Laumen H, Wurst W, Casares F, Gomez-Skarmeta JL, Winkelmann J. Restless legs syndrome-associated intronic common variant in Meis1 alters enhancer function in the developing

telencephalon. Genome Res. 2014;24(4):592-603. Epub 20140318. doi: 10.1101/gr.166751.113. PubMed PMID: 24642863; PMCID: PMC3975059.

45. Sultmann H, Sato A, Murray BW, Takezaki N, Geisler R, Rauch GJ, Klein J. Conservation of Mhc class III region synteny between zebrafish and human as determined by radiation hybrid mapping. J Immunol. 2000;165(12):6984-93. doi: 10.4049/jimmunol.165.12.6984. PubMed PMID: 11120825.

46. Dangel AW, Shen L, Mendoza AR, Wu LC, Yu CY. Human helicase gene SKI2W in the HLA class III region exhibits striking structural similarities to the yeast antiviral gene SKI2 and to the human gene KIAA0052: emergence of a new gene family. Nucleic Acids Res. 1995;23(12):2120-6. doi: 10.1093/nar/23.12.2120. PubMed PMID: 7610041; PMCID: PMC306998.

47. Moller-Levet CS, Archer SN, Bucca G, Laing EE, Slak A, Kabiljo R, Lo JC, Santhi N, von Schantz M, Smith CP, Dijk DJ. Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome. Proc Natl Acad Sci U S A. 2013;110(12):E1132-41. Epub 20130225. doi: 10.1073/pnas.1217154110. PubMed PMID: 23440187; PMCID: PMC3607048.

48. Alliance of Genome Resources C. Alliance of Genome Resources Portal: unified model organism research platform. Nucleic Acids Res. 2020;48(D1):D650-D8. doi: 10.1093/nar/gkz813. PubMed PMID: 31552413; PMCID: PMC6943066.

49. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, Savage JE, Hammerschlag AR, Skene NG, Munoz-Manchado AB, andMe Research T, White T, Tiemeier H, Linnarsson S, Hjerling-Leffler J, Polderman TJC, Sullivan PF, van der Sluis S, Posthuma D. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet. 2018;50(7):920-7. Epub 20180625. doi: 10.1038/s41588-018-0151-7. PubMed PMID: 29942085.

50. Gouveia C, Gibbons E, Dehghani N, Eapen J, Guerreiro R, Bras J. Genome-wide association of polygenic risk extremes for Alzheimer's disease in the UK Biobank. Sci Rep. 2022;12(1):8404. Epub 20220519. doi: 10.1038/s41598-022-12391-2. PubMed PMID: 35589863; PMCID: PMC9120074.

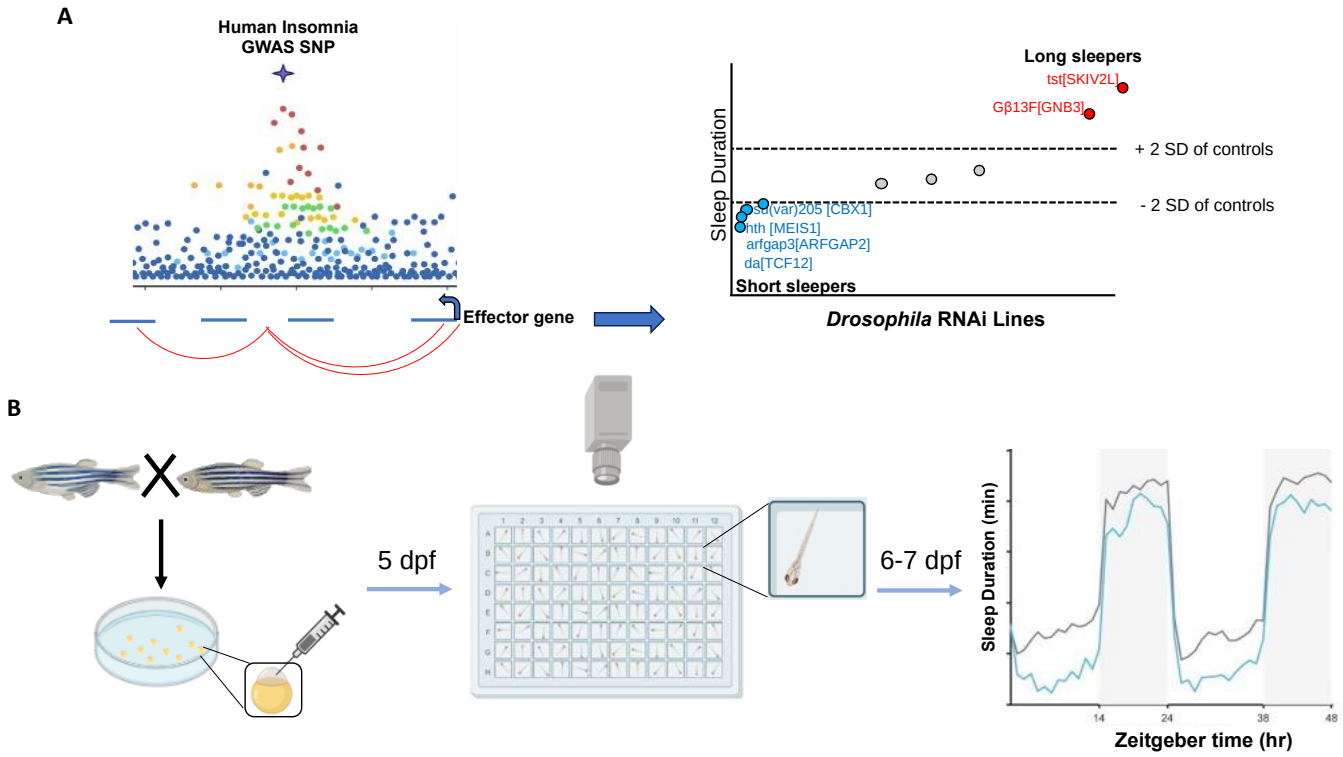

Fig. 1. Identification of high-confidence insomnia candidate genes from GWAS signals. A. Insomnia GWAS variants were mapped to their target effector gene in neural progenitor cells previously, then candidate effector genes were screened using *Drosophila* RNA interference to identify those producing significant sleep phenotypes. **B**. This work used CRISPR/Cas9 to create mutations in single cell embryos from zebrafish for effector genes that produced strong sleep phenotypes in *Drosophila*. F0 larvae were tested for sleep characteristics beginning on 5 days post fertilization (dpf) and phenotyped from day 6-7 using automated video tracking.

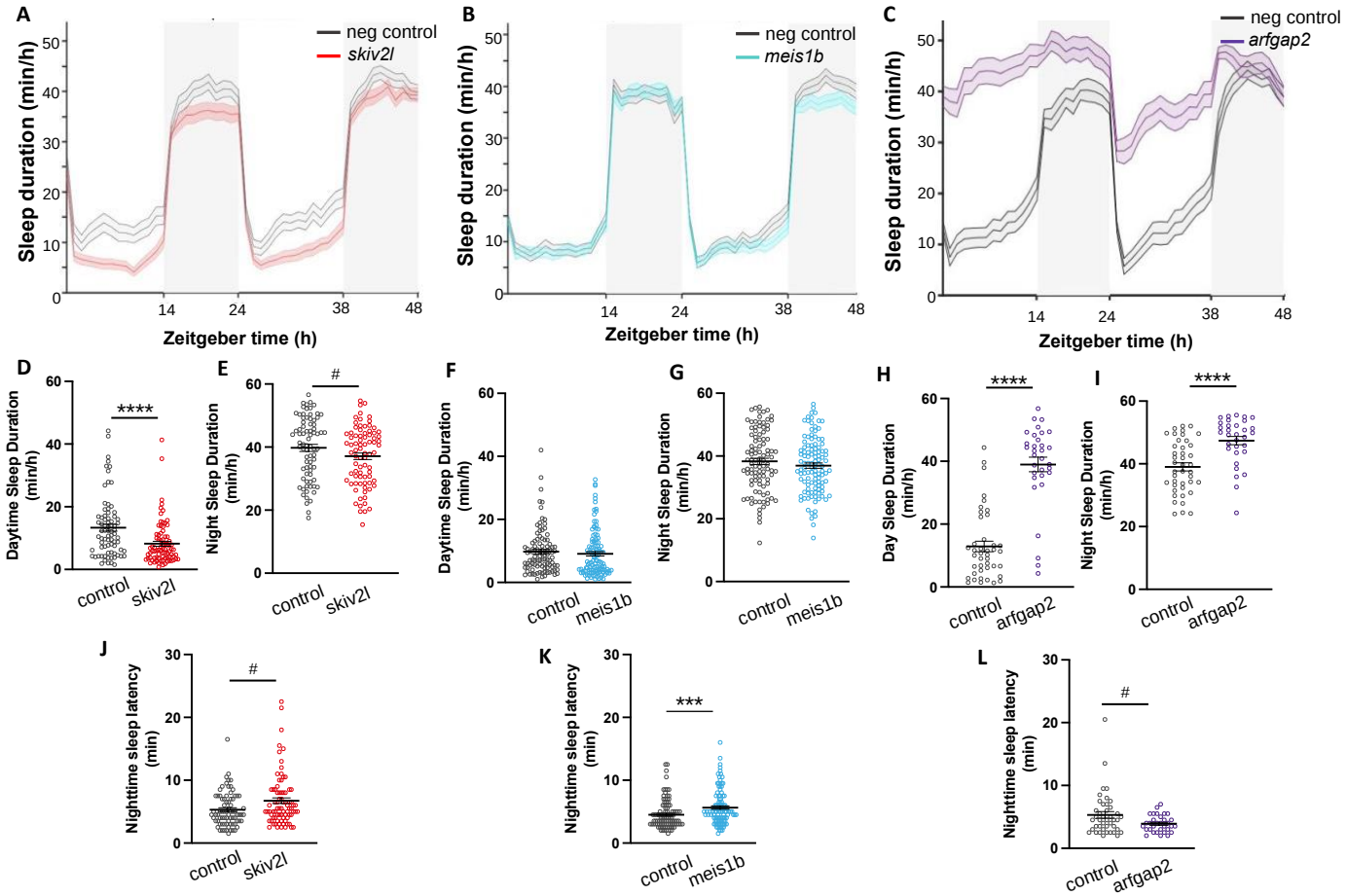

Fig. 2. Total sleep duration is perturbed by CRISPR mutation of skiv2l and arfgap2. **A-C**. Rest-activity graphs for CRISPR mutants represented as mean ± SEM for sleep duration (min/h). **D-E**. Mean ± SEM for day (**D**) and night (**E**) sleep duration in *skiv2l* mutants (n = 77 neg. control, 80 *skiv2l*). **F-G**. Mean ± SEM for day (**F**) and night (**G**) sleep duration in *meis1b* mutants (n = 100 neg. control, 111 *meis1b*). **H-I**. Mean ± SEM for day (**H**) and night (**I**) sleep duration in *arfgap2* mutants (n = 42 neg. control, 32 *arfgap2*). Mean ± SEM for nighttime sleep latency in *skiv2l* (**J**), *meis1b* (**K**), and *arfgap2* (**L**) mutants. Significance determined by Wilcoxon rank sum test followed by Hochberg step-up procedure for each gene relative to its own negative control group (see **Methods**). # p<0.1, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

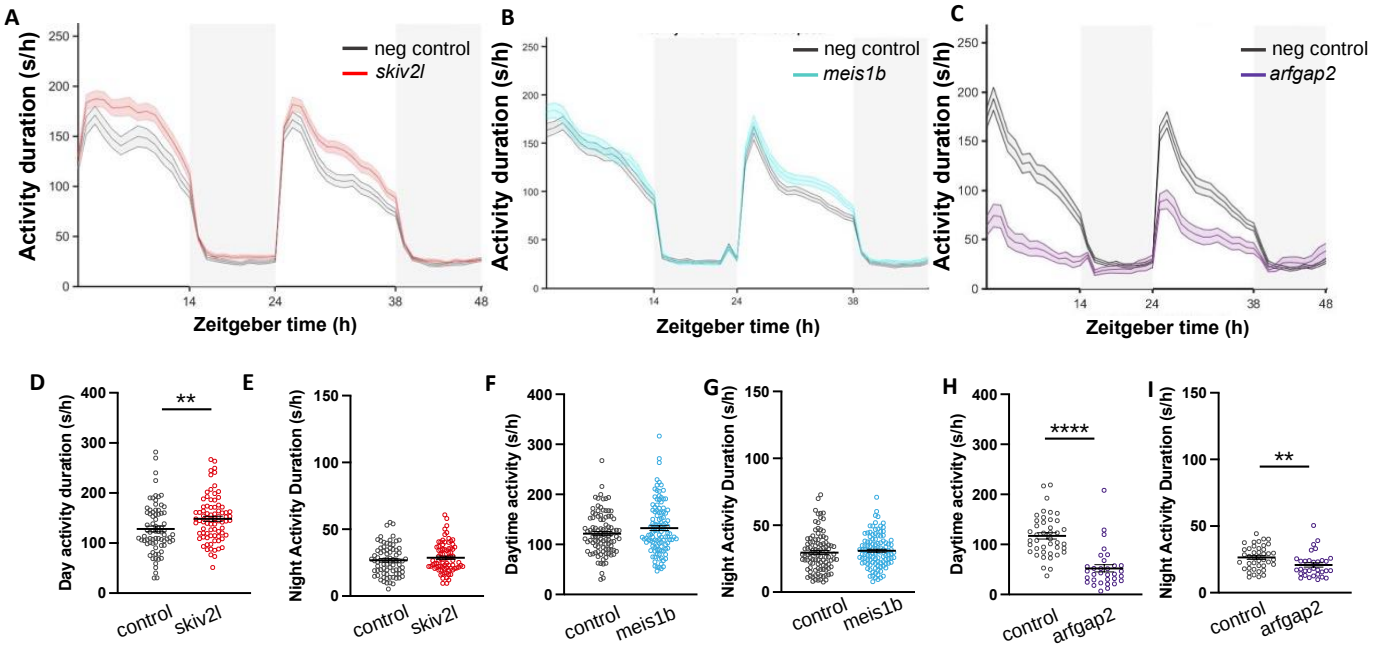

Fig 3. Activity duration is altered by skiv2l and arfgap2 mutations. A-C. Rest-activity graphs for CRISPR mutants represented as mean ± SEM for activity duration (s/h). **D-E**. Mean ± SEM for day (**D**) and night (**E**) activity duration in *skiv2l* mutants (n = 77 neg control, 80 *skiv2l*). **F-G**. Mean ± SEM for day (F) and night (G) activity duration in *meis1b* mutants ((n = 100 control, 111 meis1b). **H-I.** Mean ± SEM for day (**H**) and night (**I**) activity duration in arfgap2 mutants (n = 42 control, 32 arfgap2). Significance determined by Wilcoxon rank sum test followed by Hochberg step-up procedure for each gene relative to its own negative control group (see **Methods**). # p<0.1, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

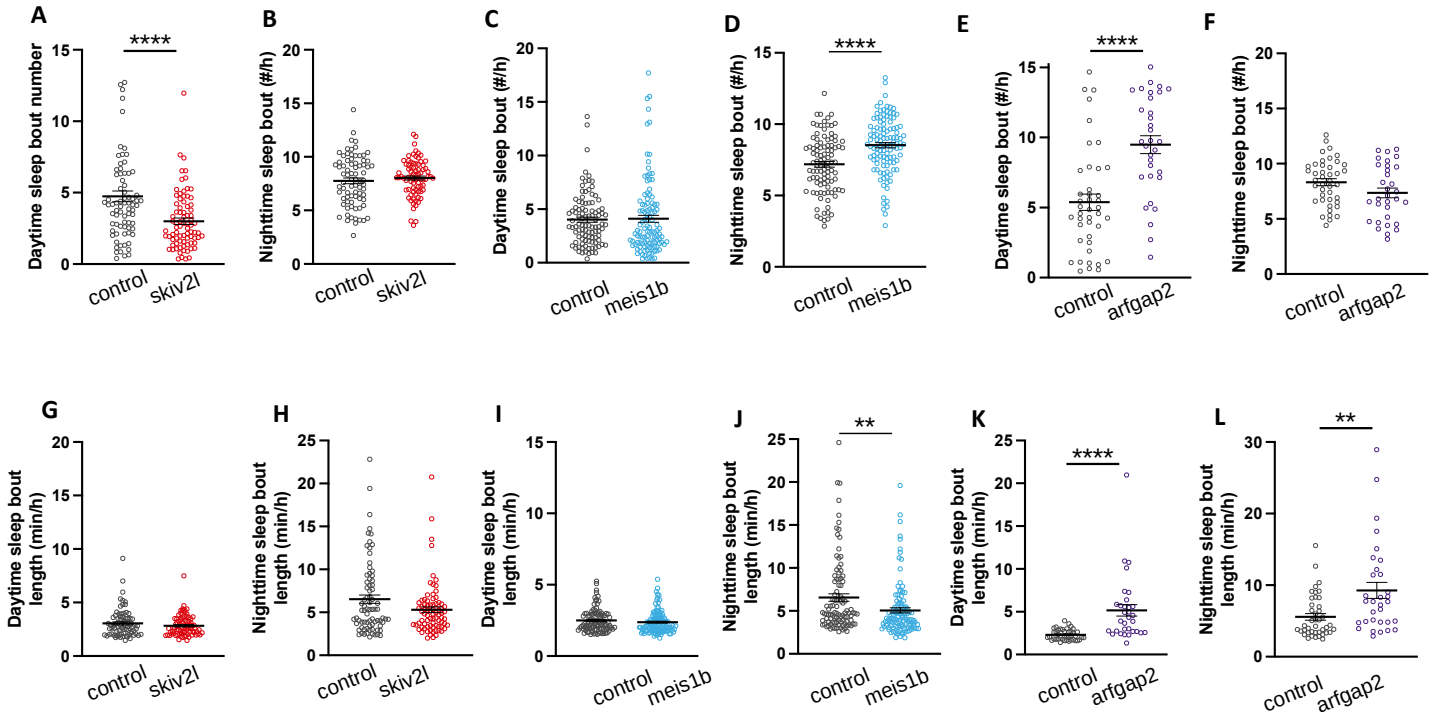

Fig. 4. Sleep architecture is significantly altered by CRIPSR mutation of *skiv2l*, *meis1b*, and *arfgap2*. Mean ± SEM for day and night sleep bout number in *skiv2l* (**A-B**), *meis1b* (**C-D**), and *arfgap2* (**E-F**). Mean ± SEM for day and night sleep bout length for *skiv2l* (**G-H**), *meis1b* (**I-J**) and *arfgap2* (**K-L**) mutants. Significance determined by Wilcoxon rank sum test followed by Hochberg step-up procedure for each gene relative to its own negative control group (see **Methods**). # p<0.1, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

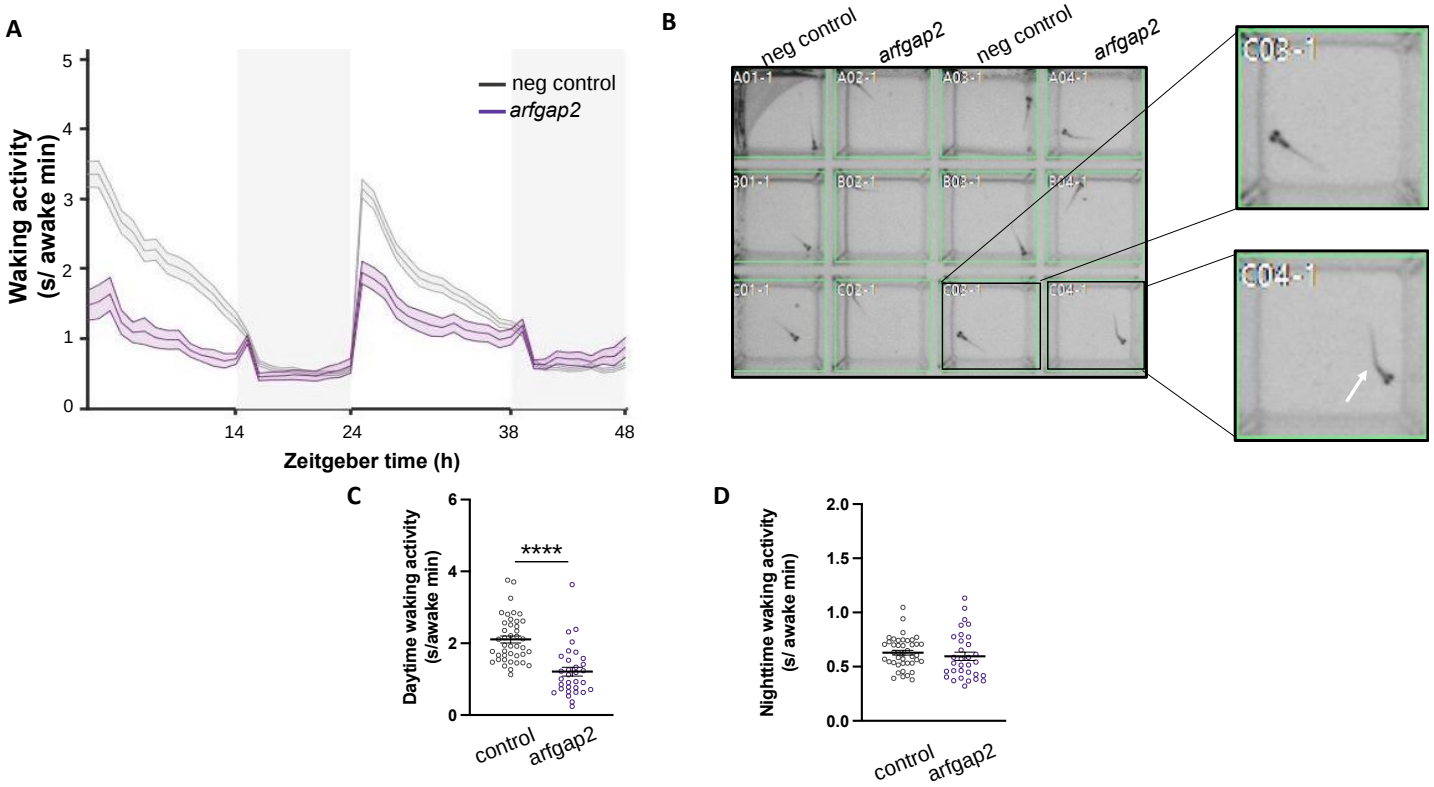

Fig. 5. CRISPR mutation of *arfgap2* produces developmental and movement abnormalities. A. Waking activity graph for *arfgap2* mutants represented as mean ± SEM. **B.** Representative image of *arfgap2* mutants and controls. Notable tail curvature (white arrow) is apparent in nearly all *arfgap2* mutants. **C-D.** Mean ± SEM for day (**C**) and night (**D**) waking activity for control and *arfgap2* mutants. n = 42 neg. control and 32 *arfgap2* mutants. ****p<0.0001 determined by Wilcoxon rank sum test and Hochberg step-up procedure for each gene relative to its own negative control group (see **Methods**).