

## GENETICS

# Evolutionarily conserved regulation of sleep by epidermal growth factor receptor signaling

Daniel A. Lee<sup>1</sup>, Justin Liu<sup>1</sup>, Young Hong<sup>1</sup>, Jacqueline M. Lane<sup>2,3</sup>, Andrew J. Hill<sup>1</sup>, Sarah L. Hou<sup>1</sup>, Heming Wang<sup>2,3</sup>, Grigorios Oikonomou<sup>1</sup>, Uyen Pham<sup>1</sup>, Jae Engle<sup>1</sup>, Richa Saxena<sup>2,3,4</sup>, David A. Prober<sup>1\*</sup>

The genetic bases for most human sleep disorders and for variation in human sleep quantity and quality are largely unknown. Using the zebrafish, a diurnal vertebrate, to investigate the genetic regulation of sleep, we found that epidermal growth factor receptor (EGFR) signaling is necessary and sufficient for normal sleep levels and is required for the normal homeostatic response to sleep deprivation. We observed that EGFR signaling promotes sleep via mitogen-activated protein kinase/extracellular signal-regulated kinase and RFamide neuropeptide signaling and that it regulates RFamide neuropeptide expression and neuronal activity. Consistent with these findings, analysis of a large cohort of human genetic data from participants of European ancestry revealed that common variants in genes within the EGFR signaling pathway are associated with variation in human sleep quantity and quality. These results indicate that EGFR signaling and its downstream pathways play a central and ancient role in regulating sleep and provide new therapeutic targets for sleep disorders.

## INTRODUCTION

Identifying how sleep is regulated is a critical health priority. Sleep loss and sleep disorders are among the most common, yet frequently overlooked, human health problems. An estimated 50 to 70 million Americans suffer from a chronic sleep disorder (1), but the genetic and neural mechanisms that regulate human sleep remain poorly understood. Large-scale human genome-wide association studies (GWAS) have recently identified genomic regions associated with variation in human sleep and several sleep disorders (2–8), but in most cases, the causative gene(s) remains unclear, and most genes identified in these studies have not been validated in vivo. With the exception of narcolepsy, which results from loss of hypocretin (Hcrt) neurons (1), and circadian disorders that affect sleep timing (1), the genetic basis for variation in human sleep and sleep disorders is largely unknown.

Sleep is an evolutionarily conserved behavioral state defined as a period of rapidly reversible immobility and increased arousal threshold that is under homeostatic control, such that prolonged wakefulness leads to a compensatory increase in sleep duration and intensity (9). This definition is valid for complex organisms such as humans, for simpler vertebrate animal models such as rodents and zebrafish, and for invertebrates such as *Caenorhabditis elegans* and *Drosophila melanogaster* (9, 10). However, while recent studies have begun to reveal mechanisms that regulate sleep in either invertebrate or vertebrate model organisms, there is relatively little evidence for conserved genetic or neuronal mechanisms that regulate sleep in both vertebrates and invertebrates. While neuronal circuits that regulate sleep in *Drosophila* and *C. elegans* have been described in detail (11), there are no clear vertebrate homologs for these invertebrate neuro-

nal systems. Similarly, the few signaling pathways implicated in regulating vertebrate sleep, such as Hcrt- and melanin-concentrating hormone signaling, lack unambiguous invertebrate orthologs (10). As a result, it is unclear whether mechanisms that regulate sleep evolved independently in invertebrate and vertebrate species and whether mechanisms identified in invertebrates are relevant to sleep in humans. Identifying conserved mechanisms, if they exist, would be valuable because they would likely have central roles in regulating sleep and may provide novel therapeutic targets for sleep disorders.

Among the peptidergic signaling pathways known to regulate invertebrate sleep, the pathway with perhaps the clearest homology in vertebrates is that of the epidermal growth factor (EGF) receptor (EGFR) and its ligands. EGFR signaling has primarily been studied for its roles in development and physiology, but more recently it has also been implicated in regulating several behaviors (12–20). Genetic gain- and loss-of-function studies in *Drosophila* and *C. elegans* have demonstrated that EGFR signaling is both necessary and sufficient for normal sleep levels in these invertebrate animals (12, 13, 16, 19, 21, 22). Furthermore, a GWAS using a panel of inbred wild-caught *Drosophila* lines identified an association between common variants in several EGFR pathway genes and sleep duration (17), as did an artificial selection screen for long and short sleep duration in an outbred *Drosophila* population (18). These GWAS observations suggest that EGFR signaling plays a role in natural sleep variation in *Drosophila*. However, mechanisms that link EGFR signaling to downstream sleep regulators are poorly understood. It is also unclear whether EGFR signaling regulates sleep in vertebrate model organisms and in humans.

Here, we address the role of EGFR signaling and downstream mechanisms in regulating sleep using the zebrafish, a vertebrate with established, behaviorally defined diurnal sleep (23). This model organism has several advantages over more commonly used rodent species for modeling aspects of human sleep, including a simpler, yet anatomically and molecularly conserved, vertebrate brain; rapid development with complex behaviors such as sleep present in 5-day-old animals; amenability to high-throughput behavioral assays; and amenability to genetic and pharmacological manipulations (10).

<sup>1</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. <sup>2</sup>Center for Genomic Medicine and Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA. <sup>3</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142, USA. <sup>4</sup>Division of Sleep and Circadian Disorders, Department of Medicine, Brigham and Women's Hospital and Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA.

\*Corresponding author. Email: dprober@caltech.edu

Furthermore, zebrafish sleep has genetic, pharmacological, and neuroanatomical similarities to mammalian sleep (10), suggesting that findings in zebrafish will translate to mammals.

Using genetics and pharmacology, we show that EGFR signaling is both necessary and sufficient for normal sleep duration, promotes consolidation and maintenance of sleep, and is required for the normal sleep rebound response to sleep deprivation (SD) in zebrafish. In addition, we show that EGFR signaling promotes sleep in zebrafish via mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK) and RFamide neuropeptide signaling, and describe a mechanism linking the EGFR and RFamide neuropeptide systems. Last, by analyzing a large cohort of human genetic data (24), we identify associations between sleep duration and quality measures with common variants in genes that participate in EGFR signaling. These results establish EGFR signaling as an evolutionarily conserved sleep-promoting system with implications for human sleep disorders.

## RESULTS

### Transforming growth factor, alpha overexpression increases sleep

To explore the role of EGFR signaling in vertebrate sleep, we first characterized the expression pattern of each zebrafish ortholog of EGFR and its ligands in the brain at 5 days postfertilization (dpf) using in situ hybridization (ISH). We found that one EGFR paralog, *egfra*, is expressed in juxtaventricular cells that colocalize with *sox2* (Fig. 1, A, A', and B), a marker of glial and ependymal cells (25), as well as in the hypothalamus, hindbrain, tectum, and cerebellum (Fig. 1, A' and A''). The EGFR ligands transforming growth factor, alpha (TGFA) and EGF are implicated in regulating mammalian behavior (14, 20, 26), and each gene has a single zebrafish ortholog. Zebrafish *egf* is expressed just dorsal to *egfra*-expressing cells in the diencephalon (fig. S1A). Similar to mammals and zebrafish *egf* (27), zebrafish *tgfa* is also expressed in juxtaventricular cells in the diencephalon, just dorsal to *egfra*-expressing cells (Fig. 1, C and C'), that also express *sox2* (Fig. 1, D and D'). We observed a significant day/night oscillation in *tgfa* mRNA level by reverse transcription quantitative polymerase chain reaction (qPCR) [ $P < 0.05$ , one-way analysis of variance (ANOVA)], with peak *tgfa* expression at 12 a.m. 150% higher than the trough at 12 p.m. in animals entrained under 14-hour light/10-hour dark conditions until 6 dpf ( $P < 0.05$ , one-way ANOVA, Holm-Sidak test; fig. S1, B and C), consistent with a potential role for *tgfa* in regulating sleep. On the basis of these results, we tested the roles of *tgfa* and *egfra* in zebrafish sleep.

To test whether EGFR signaling is sufficient to promote sleep, we used a genetic strategy to conditionally overexpress TGFA while monitoring behavior using a video tracking assay (28). As previously established using locomotor activity and arousal threshold assays, a period of 1 min (or longer) of inactivity corresponds to a sleep bout in zebrafish larvae (28). We generated *Tg(hs:tgfa)* zebrafish, in which a heat shock (HS)-inducible promoter regulates expression of zebrafish *tgfa*. Before HS, we observed no difference in locomotor activity or sleep between transgenic animals and their wild-type (WT) siblings (Fig. 1, E to H). In contrast, after HS-induced TGFA overexpression, transgenic animals were less active and slept more than their WT siblings during both the night and the following day (night activity, -53%; day activity, -52%; night sleep, +54%; day sleep, +400%;  $P < 0.005$ , two-way ANOVA, Holm-Sidak test) (Fig. 1, E to H). The increase in sleep was most pronounced during the day, when zebrafish

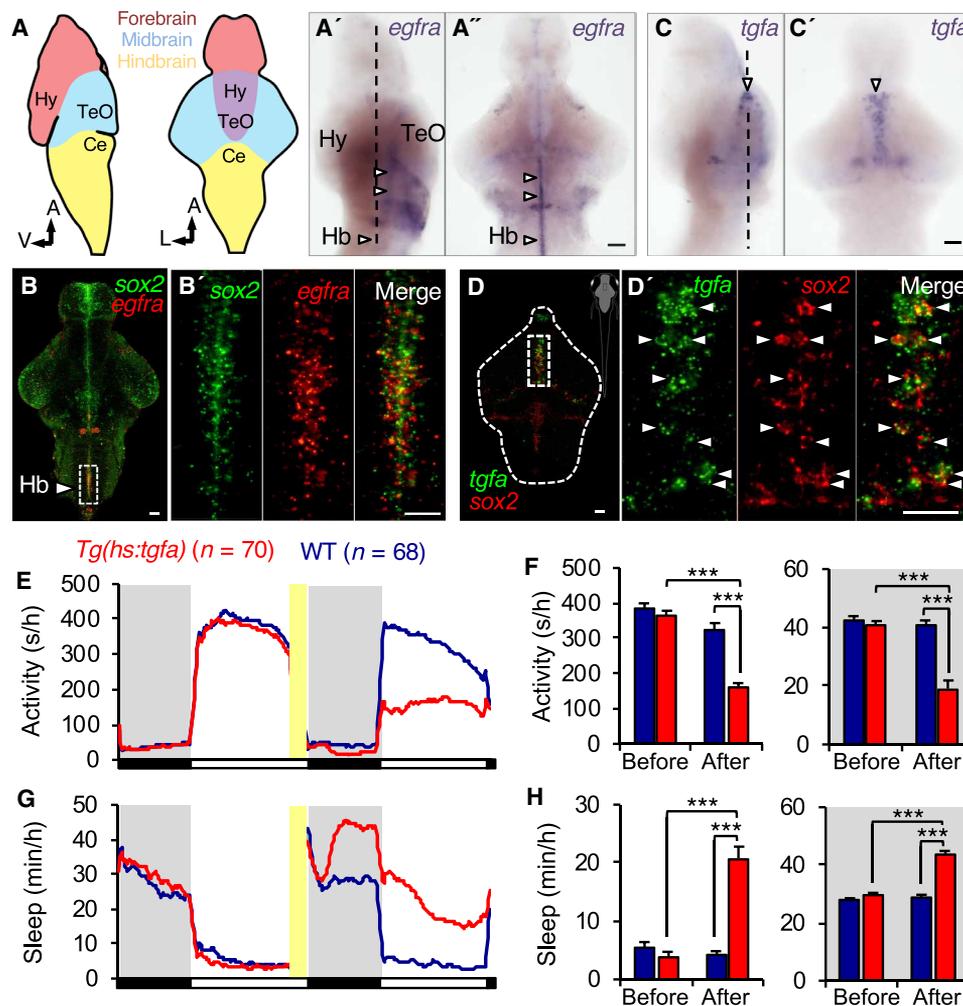
are normally most awake. This phenotype was due to an increase in sleep bout length during both the night and day, an increase in the number of sleep bouts during the day, and a decrease in wake bout length during the day (fig. S1, D to F). TGFA-overexpressing animals also exhibited a decrease in sleep latency (time to first sleep bout following lights on or off) during the day and a decrease in daytime and nighttime waking activity (locomotor activity while awake) (fig. S1, G and H). Thus, activation of EGFR signaling results in increased sleep due to consolidation of the sleep state and decreased sleep latency.

### Genetic loss of EGFR signaling decreases sleep

We next tested the hypothesis that EGFR signaling is required for normal sleep levels by creating zebrafish with mutations in *egfra* and its ligands. Mice containing a null mutation in *Tgfa* are viable and fertile (29), but sleep studies using this mutant have not been reported. Using the transcription activator-like effector nuclease (TALEN) method (30), we generated a zebrafish *tgfa* mutant that is predicted to lack all known functional domains and should thus be a null mutation (fig. S2A). *tgfa*<sup>-/-</sup> zebrafish are viable and fertile, with no apparent morphological abnormalities. Consistent with the TGFA overexpression phenotype, *tgfa*<sup>-/-</sup> animals were 15% more active and slept 28% less than their *tgfa*<sup>+/+</sup> siblings during the day (activity,  $P < 0.005$ ; sleep,  $P < 0.01$ ; one-way ANOVA, Holm-Sidak test; fig. S1, I to L). These effects were primarily due to fewer sleep bouts and higher levels of waking activity (fig. S1, M to Q).

To ask whether EGF also contributes to sleep regulation, we used CRISPR-Cas9 (31) to isolate an *egf* mutant that is predicted to lack all known functional domains after the initial three EGF-like domains and should thus be a null mutation (fig. S2B). Similar to *tgfa*<sup>-/-</sup> animals, *egf*<sup>-/-</sup> animals were 26% more active and showed a trend of less sleep than their *egf*<sup>+/+</sup> siblings during the day (activity,  $P < 0.01$ ; sleep,  $P = 0.10$ ; one-way ANOVA, Holm-Sidak test; fig. S1, R to U), with sleep architecture phenotypes more modest than those observed in *tgfa* mutants (fig. S1, V to Z). To ask whether *egf* and *tgfa* act redundantly, we compared the behavior of *egf*<sup>-/-</sup>; *tgfa*<sup>-/-</sup> animals to their *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> and *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> siblings. We found that *tgfa*<sup>-/-</sup>; *egf*<sup>-/-</sup> animals are viable and fertile, with no apparent morphological abnormalities. *egf*<sup>-/-</sup>; *tgfa*<sup>-/-</sup> animals were more active and slept less during the day (activity, +20%; sleep, -45%;  $P < 0.005$ , one-way ANOVA, Holm-Sidak test) and slept less at night (-13%;  $P < 0.05$ , one-way ANOVA, Holm-Sidak test) compared to their *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> siblings (Fig. 2, A to D). The decrease in sleep was due to fewer sleep bouts and longer wake bouts during the day and shorter sleep bouts at night (fig. S1, AA to AE), suggesting that loss of both *egf* and *tgfa* results in consolidation of wakefulness during the day and sleep fragmentation at night. These results suggest that *tgfa* and *egf* have partially redundant roles in EGFR-mediated sleep since loss of both genes resulted in a stronger sleep defect (day sleep, -45%;  $P < 0.005$ ; night sleep, -13%;  $P < 0.05$ ; one-way ANOVA, Holm-Sidak test; Fig. 2, C and D) than loss of *tgfa* alone (day sleep, -28%;  $P < 0.01$ ; night sleep, -4.7%,  $P > 0.05$ ; one-way ANOVA, Holm-Sidak test; fig. S1, I to Q).

We next tested whether EGFR is required for normal sleep levels in zebrafish. *Egfr* knockout mice are embryonic lethal, likely because of defects in placental development (32), thus preventing genetic loss-of-function sleep studies. However, since zebrafish develop ex utero without a placenta and contain at least two *egfr* paralogs (*egfra* and *egfrb*) (33), we reasoned that we could analyze the requirement for EGFR in zebrafish sleep. Since we found that *egfra* is expressed



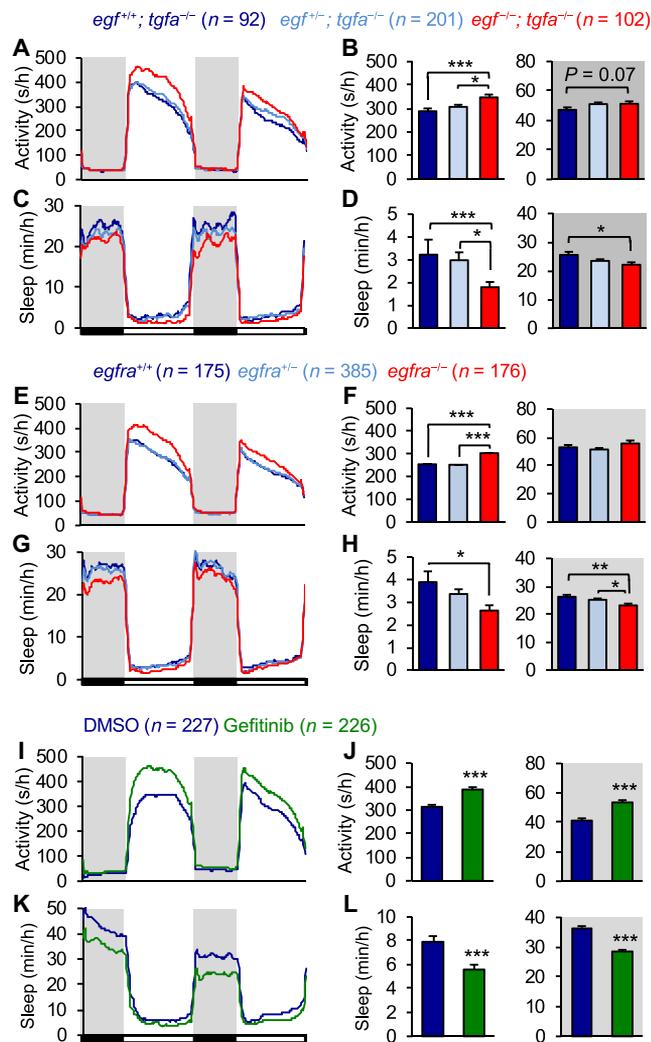
**Fig. 1. TGFA overexpression increases sleep.** (A to D) ISH of *egfra* and *tgfa* in the 5-dpf zebrafish brain (schematic) (A). A, anterior; L, lateral; V, ventral; Ce, cerebellum; Hy, hypothalamus; TeO, tectum. (A' and A'') Sagittal (A') and dorsal (A'') views of *egfra* expression in juxtaventricular cells (white arrowheads). (B and B') *egfra* coexpression with *sox2* in these cells. (C and C') Sagittal (C) and dorsal (C') views of *tgfa* expression in cells just dorsal to juxtaventricular cells in the diencephalon (white arrowheads). (D and D') *tgfa* coexpression with *sox2* in these cells (white arrowheads). Dashed lines in (A) and (C) indicate the horizontal planes shown in (A'') and (C'). Boxed regions in (B) and (D) are magnified in (B') and (D'). Dashed line in (D) shows outline of brain. Scale bars, 30  $\mu\text{m}$  (B, B', D, and D') and 50  $\mu\text{m}$  (A'' and C'). (E to H) Following a 1-hour HS (yellow bars), *Tg(hs:tgfa)* animals were less active (E and F) and slept more (G and H) than their WT siblings. Pre-HS and post-HS are calculated for the day or night before, and the day or night after, HS, respectively. White and black bars indicate day (14 hours) and night (10 hours). Data are obtained from two pooled experiments. Bar graphs show mean  $\pm$  SEM.  $n$  = number of animals. m/h, minutes/hour; s/h, second/hour. \*\*\* $P$  < 0.0001 by two-way ANOVA with Holm-Sidak test.

in juxtaventricular cells adjacent to *tgfa*- and *egf*-expressing cells (Fig. 1, A', A'', C, and C', and fig. S1A), we used CRISPR-Cas9 to generate a zebrafish *egfra* mutant that is predicted to lack the transmembrane and intracellular domains (fig. S2C). *egfra*<sup>-/-</sup> animals are viable and lack obvious morphological abnormalities. Similar to EGFR ligand double mutants, *egfra*<sup>-/-</sup> animals were more active and showed a trend of less sleep during the day (activity, +17%;  $P$  < 0.005; sleep, -22%;  $P$  = 0.08; one-way ANOVA, Holm-Sidak test) and also slept less at night (-11%;  $P$  < 0.005, one-way ANOVA, Holm-Sidak test) compared to their *egfra*<sup>+/+</sup> siblings (Fig. 2, E to H). Similar to EGFR ligand double mutants, the *egfra* mutant phenotype was primarily due to fewer sleep bouts during the day and shorter sleep bouts at night (fig. S1, AF to AJ), suggesting that loss of *egfra* results in sleep fragmentation at night. Together, these results suggest that *tgfa*, *egf*, and *egfra* are required for normal sleep

levels in zebrafish. More broadly, the zebrafish genome contains several additional EGFR ligands, a second *egfr* paralog (33), and several ERBB family co-receptors that can form heterodimers with EGFR that can be activated by EGFR ligands; some of these proteins may also play roles in EGFR signaling-mediated sleep.

### Pharmacological inhibition of EGFR signaling decreases sleep

To address the possibility that EGFR family members, in addition to *egfra*, promote sleep, we acutely and selectively inhibited the function of EGFRs by treating WT animals with the EGFR small-molecule inhibitors gefitinib and erlotinib (34). These compounds inhibit receptor function by binding to the adenosine triphosphate-binding pocket of the tyrosine kinase domain (34), which has >90% amino acid similarity between humans and zebrafish (33). We found



**Fig. 2. Loss of EGFR signaling decreases sleep.** (A to D) *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> animals were more active during the day and slept less during the day and night than *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> siblings. Larvae were generated by mating *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> to *egf*<sup>+/+</sup>; *tgfa*<sup>-/-</sup> animals. For clarity, data for other genotypes are not shown. (E to H) *egfra*<sup>+/+</sup> animals were more active during the day and slept less during the day and night than *egfra*<sup>+/+</sup> siblings. (I to L) Gefitinib-treated WT animals were more active and slept less during the day and night than their dimethyl sulfoxide (DMSO)-treated siblings. Pooled data from nine (A to D), eight (E to H), and six (I to L) experiments are shown. Bar graphs show mean  $\pm$  SEM. *n* = number of animals. \**P*  $\leq$  0.05, \*\**P*  $<$  0.01, and \*\*\**P*  $<$  0.005 for indicated comparisons by one-way ANOVA and Holm-Sidak test (A to H) or Student's *t* test (I to L).

that gefitinib-treated WT animals were more active (day, +23%; night, +30%) and slept less (day, -29%; night, -22%) than vehicle-treated siblings during both the day and night (*P*  $<$  0.0001, Student's *t* test; Fig. 2, I to L). This phenotype was due to fewer sleep bouts during the day, shorter sleep bouts at night, and longer wake bouts during the day and night (fig. S3, C to E), suggesting that inhibition of EGFR signaling results in fragmented sleep at night and consolidated wakefulness during the day and night. Gefitinib treatment also increased waking activity and increased sleep latency during the day and night (fig. S3, F and G), suggesting that EGFR signaling is required for the normal initiation of sleep. Erlotinib treatment induced a similar

behavioral phenotype (fig. S3, H to P). Gefitinib treatment did not further decrease sleep in *egfra*<sup>-/-</sup> animals (fig. S3, A and B), suggesting that pharmacological inhibitors of EGFR promote sleep largely by inhibiting *egfra*, that gefitinib-induced sleep is not due to nonspecific effects on other proteins, and that EGFR ligands promote sleep primarily via *egfra*. Together, these pharmacological and genetic results demonstrate that EGFR signaling is required for normal sleep levels.

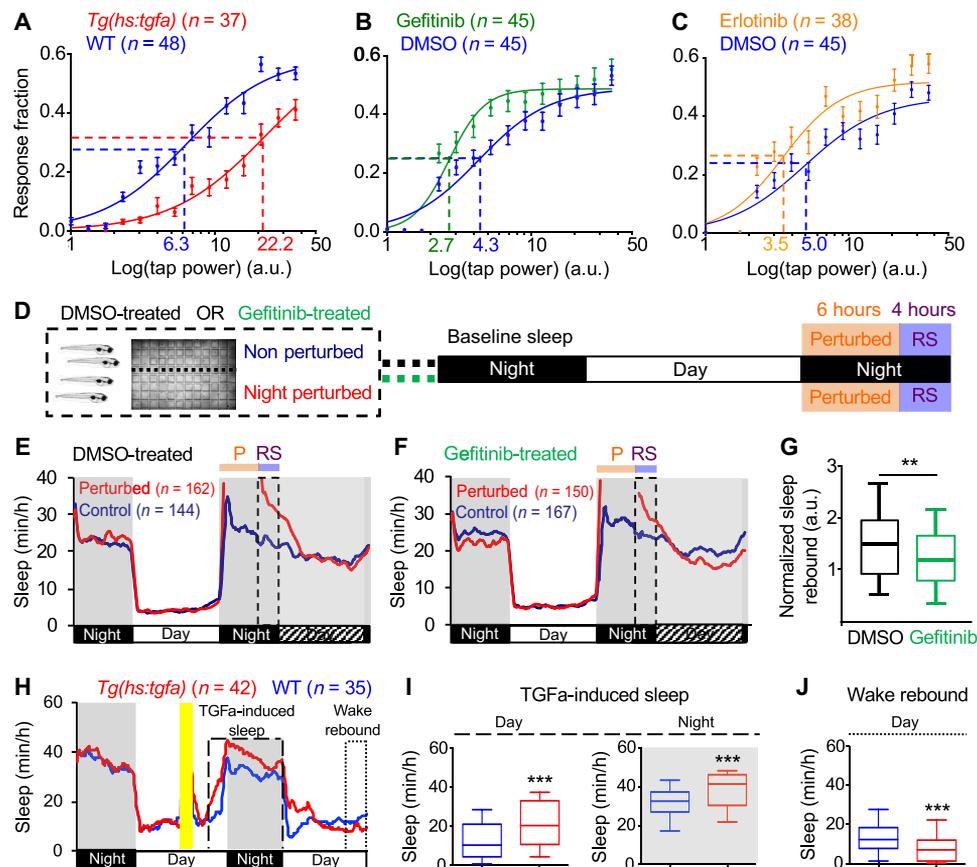
### EGFR signaling modulates arousal threshold

Sleep is distinguished from quiet wakefulness by an increased arousal threshold, which we assayed by monitoring the behavioral response of zebrafish to a mechano-acoustic stimulus (35). We quantified the fraction of animals that responded to the stimulus at 14 different intensities and used these data to construct stimulus-response curves. We observed that the tapping intensity at which the half-maximal response occurred [effective tap power 50 (ETP<sub>50</sub>)] for TGF $\alpha$ -overexpressing larvae was 252% higher than that of their WT siblings ( $F_{1,834} = 20.95$ , *P*  $<$  0.0001 by extra sum-of-squares *F* test) (Fig. 3A). Conversely, the ETP<sub>50</sub> for gefitinib-treated animals was 37% lower than that of their vehicle-treated siblings ( $F_{1,834} = 18.16$ , *P*  $<$  0.0001 by extra sum-of-squares *F* test) (Fig. 3B). Similarly, the ETP<sub>50</sub> for erlotinib-treated animals was 30% lower than that of their vehicle-treated siblings ( $F_{1,834} = 4.9$ , *P*  $<$  0.05 by extra sum-of-squares *F* test) (Fig. 3C). These results demonstrate that TGF $\alpha$  overexpression increases arousal threshold, whereas inhibition of EGFR signaling decreases arousal threshold, consistent with modulation of a sleep state.

### EGFR signaling is required for homeostatic regulation of sleep

Sleep is postulated to be controlled by two processes: a homeostatic process that increases during wakefulness and dissipates during sleep and a circadian process that controls when sleep occurs during the 24-hour day/night cycle (9). While invertebrate studies have demonstrated a role for EGFR signaling in promoting sleep (12, 13), it is unknown whether it does so as part of the homeostatic or circadian process. To distinguish between these possibilities, we first tested whether EGFR signaling is required for circadian regulation of sleep by entraining WT animals in a 14-hour light/10-hour dark cycle until 5 dpf and then treating animals with either gefitinib or dimethyl sulfoxide (DMSO) vehicle control while monitoring their behavior in either constant light or constant dark (fig. S4, A to D). While gefitinib treatment decreased sleep compared to controls, the circadian pattern of sleep was maintained with no obvious effect on circadian period length or phase, suggesting that EGFR signaling is not necessary for circadian regulation of sleep.

To explore the role of EGFR signaling in homeostatic regulation of sleep, we developed a large-scale sleep deprivation (SD) assay. We monitored the behavior of at least 60 animals from 5 to 8 dpf in a 96-well plate that was cut in half. On the beginning of the third night of the assay (7 dpf), we disrupted the sleep of half the animals using acoustic stimuli (see Materials and Methods) in the dark for 6 hours while leaving the other half of the 96-well plate in the video tracker as a nonperturbed control. Light was kept off during the remaining 4 hours of night and the following day to avoid masking effects of light on any recovery sleep (RS) response. We observed a significant increase in sleep during the remaining 4 hours of night for perturbed animals compared to nonperturbed sibling controls (fig. S5, A to C), similar to the RS observed in other species in response to SD. This increase in sleep



**Fig. 3. EGFR signaling regulates arousal threshold and sleep homeostasis.** Representative stimulus response curves for *Tg(hs:tgfa)* animals and their WT siblings following HS (A), gefitinib and DMSO vehicle-treated WT siblings (B) and erlotinib and DMSO vehicle-treated WT siblings (C). Data points indicate mean  $\pm$  SEM fraction of animals that responded to the stimulus. Dashed lines mark ETP<sub>50</sub> values. (A) *Tg(hs:tgfa)* animals had an ETP<sub>50</sub> value of 22.2 versus 6.3 for WT siblings (252% increase;  $F_{1,834} = 20.95$ ,  $P < 0.0001$  by extra sum-of-squares  $F$  test). (B) Gefitinib-treated animals had an ETP<sub>50</sub> of 2.7 versus 4.3 for DMSO-treated siblings (37% decrease;  $F_{1,834} = 18.16$ ,  $P < 0.0001$  by extra sum-of-squares  $F$  test). (C) Erlotinib-treated animals had an ETP<sub>50</sub> of 3.5 versus 5.0 for DMSO-treated siblings (30% decrease;  $F_{1,834} = 4.9$ ,  $P < 0.05$  by extra sum-of-squares  $F$  test). a.u., arbitrary units.  $n$  = number of animals. (D) SD paradigm. (E and F) Sleep behavioral traces for animals treated with DMSO (E) or gefitinib (F) starting at 5 dpf, subjected to SD during the first 6 hours of the night at 7 dpf (P; orange), and monitored thereafter in the dark. The 4-hour periods of RS (purple) are indicated with dashed boxes. (G) Normalized sleep rebound in DMSO- or gefitinib-treated siblings following SD. Normalized sleep rebound = amount of sleep for each perturbed animal during the first 4 hours of RS divided by the average amount of sleep of all nonperturbed controls during this time period. (H to J) After a 1-hour HS in the middle of the day (yellow bar), *Tg(hs:tgfa)* animals showed increased daytime and nighttime sleep for ~24 hours (H and I), followed by decreased sleep (H and J). Sleep in (I) and (J) is quantified for the boxed regions in (H), which includes the entire night and the last 3 hours of each day. Pooled data from five (D to G) and two (H to J) experiments are shown. Black, white, and hatched bars under behavioral traces indicate night (10 hours), day (14 hours), and subjective day (14 hours), respectively.  $n$  = number of animals. \*\* $P < 0.01$  and \*\*\* $P < 0.001$  by Mann-Whitney test.

persisted during the following subjective day (fig. S5A) but returned to normal levels by the following night (fig. S5D). To verify that the increased RS resulted from SD and not simply fatigue or some other response to the stimulus, we applied the same acoustic stimulation to developmentally identical siblings during the day, when the animals spend most of their time awake (fig. S5E). We observed no increase in sleep in perturbed animals compared to nonperturbed sibling controls following daytime perturbation (fig. S5, E to H), consistent with the increased sleep observed following perturbation at night being due to loss of sleep. We calculated a normalized sleep rebound value for each animal as the amount of sleep during the first 4 hours of RS divided by the average amount of sleep of all nonperturbed controls during this time period. Normalized sleep rebound was significantly higher in animals perturbed during the night, when zebrafish spend most of the time asleep, compared to those perturbed during the day, when zebrafish spend most of the

time awake (fig. S5I), as expected for a perturbation that specifically disrupts sleep. These results indicate that this SD assay is both specific and robust in its ability to detect a homeostatic response to SD.

We next used this assay to test whether EGFR signaling is required for homeostatic regulation of sleep. To do so, we sleep-deprived WT siblings treated with either the EGFR inhibitor gefitinib or DMSO vehicle control in neighboring video trackers for 6 hours at night and quantified sleep rebound during the remaining 4 hours of night (Fig. 3, D to G). We observed a 28% reduction in normalized sleep rebound for gefitinib-treated animals compared to DMSO-treated controls (median response = 1.15 for gefitinib versus 1.59 for DMSO,  $P < 0.01$ , Mann-Whitney test; Fig. 3G and fig. S5, J to O). This result indicates that the amount of rebound sleep following SD is reduced in gefitinib-treated animals compared to DMSO-treated controls and, thus, that EGFR signaling participates in the homeostatic regulation of sleep.

If EGFR signaling regulates sleep homeostasis, then sleep induced by activation of EGFR signaling might be followed by a period of decreased sleep. When TGF $\alpha$  overexpression was induced in the middle of the day, we observed an initial increase in sleep that persisted for ~24 hours (day,  $P < 0.001$ ; night,  $P < 0.001$ , two-way ANOVA with Holm-Sidak test; Fig. 3, H and I), followed by a decrease in sleep ( $P < 0.001$ , two-way ANOVA with Holm-Sidak test; Fig. 3, H and J) compared to sibling controls. These data are consistent with the hypothesis that EGFR signaling regulates sleep homeostasis.

### TGF $\alpha$ overexpression–induced sleep is primarily mediated by EGFR $\alpha$

To investigate the cellular mechanism through which TGF $\alpha$  promotes sleep, we first asked which cells in the brain are activated by TGF $\alpha$  overexpression by performing ISH using a probe for *c-fos*, a marker of neural activity (36). We used *Tg(hs:tgfa)* animals and their WT siblings fixed 4 hours after a midday HS, when *Tg(hs:tgfa)* animals sleep ~100% more than their WT siblings (Fig. 4A). We observed that *tgfa* overexpression resulted in strong *c-fos* expression in juxtaventricular cells (Fig. 4C), similar to the expression pattern of *egfra* (Fig. 1A"). In contrast, only low levels of *c-fos* expression were present in identically treated WT siblings (Fig. 4B), indicating that the strong *c-fos* expression in *Tg(hs:tgfa)* animals was not due to the HS. Double fluorescence ISH (FISH) confirmed that TGF $\alpha$  overexpression induced *c-fos* expression in *egfra*-expressing juxtaventricular cells (Fig. 4, D and D'). This result is consistent with the hypothesis that TGF $\alpha$  overexpression–induced sleep is mediated, at least in part, by *egfra*-expressing juxtaventricular cells in the brain.

To test this hypothesis, we compared the behavioral effects of TGF $\alpha$  overexpression on sleep in *egfra*<sup>-/-</sup> and *egfra*<sup>+/+</sup> siblings. We found that *Tg(hs:tgfa)*; *egfra*<sup>+/+</sup> animals slept 270% more than their *egfra*<sup>+/+</sup> siblings after HS-induced TGF $\alpha$  overexpression ( $P < 0.0001$ , two-way ANOVA, Holm-Sidak test; Fig. 4, E and G). These effects were strongly suppressed in *Tg(hs:tgfa)*; *egfra*<sup>-/-</sup> animals compared to their *egfra*<sup>-/-</sup> siblings ( $P = 0.42$ , two-way ANOVA with Holm-Sidak test; Fig. 4, F and G). This result indicates that TGF $\alpha$  overexpression–induced sleep is largely mediated by EGFR $\alpha$ . Consistent with this result, TGF $\alpha$  overexpression–induced *c-fos* expression was significantly reduced in *egfra*<sup>-/-</sup> animals (Fig. 4, H to L). Together, these data suggest that EGFR ligands promote sleep primarily by activating *egfra*-expressing juxtaventricular cells in the brain.

### Inhibition of MAPK/ERK signaling suppresses EGFR-induced sleep

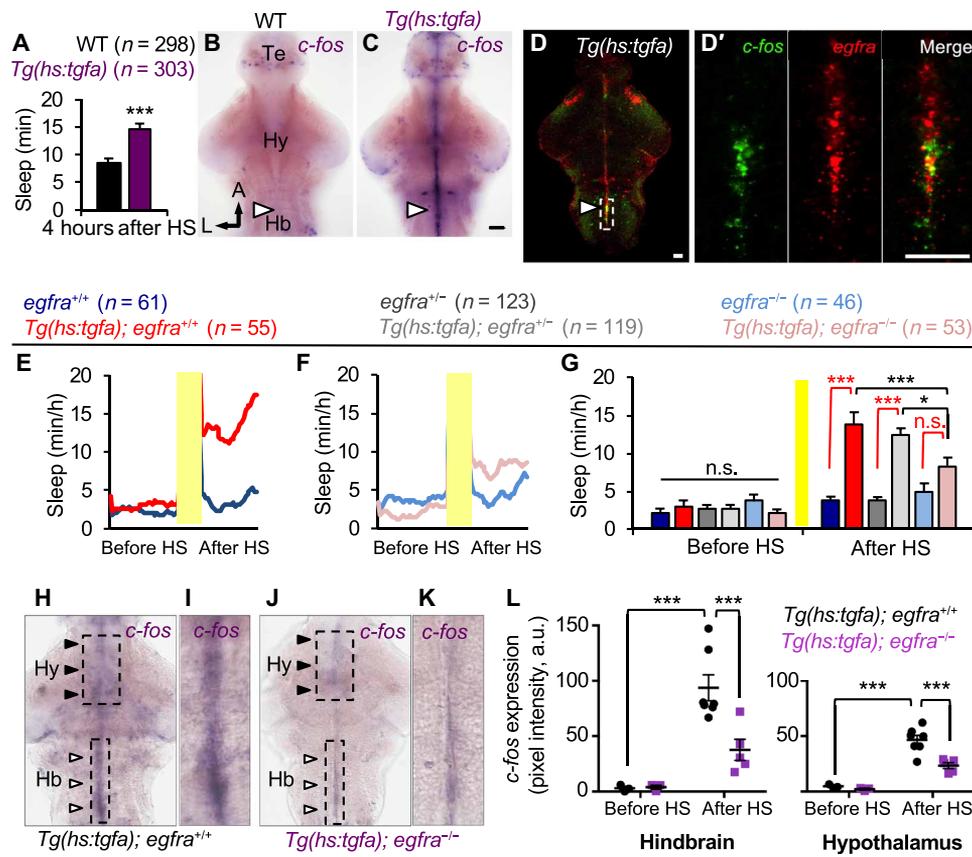
EGFR can act via multiple signal transduction pathways, but genetic (13) and pharmacological (37) studies in *Drosophila* suggest that EGFR signaling promotes sleep, at least in part, via MAPK/ERK signaling. To test whether the MAPK/ERK pathway is required for EGFR-mediated sleep in zebrafish, we treated *Tg(hs:tgfa)* larvae immediately after HS with either of two MEK1/2 (MAPK kinase 1/2) antagonists previously demonstrated to be effective in zebrafish (38). While these antagonists had no effect on sleep or activity in WT animals at the concentrations tested, *Tg(hs:tgfa)* animals treated with either 3  $\mu$ M SL327 (fig. S6, A to F) or 15  $\mu$ M U0126 (fig. S6, G to L) were significantly more active and slept less after HS than vehicle-treated *Tg(hs:tgfa)* siblings ( $P < 0.001$ , two-way ANOVA, Holm-Sidak test), although they were still less active and slept more than their drug-treated WT siblings ( $P < 0.001$ , two-way ANOVA, Holm-Sidak test). SL327 treatment reduced both the number and

length of sleep bouts in *Tg(hs:tgfa)* animals, while U0126 treatment only reduced the number of sleep bouts (fig. S6, C, F, I, and L), possibly because SL327 had a more potent effect on behavior than U0126. Thus, both MEK1/2 antagonists suppressed TGF $\alpha$  overexpression–induced sleep, suggesting that EGFR signaling regulates sleep, in part, via the MAPK/ERK pathway in zebrafish. These results link previous independent observations describing changes in ERK signaling following EGFR-induced sleep in *Drosophila* (13) and a role for ERK in regulating invertebrate and vertebrate sleep (37, 39).

### TGF $\alpha$ overexpression–induced sleep is suppressed by mutation of *npvf*

In *C. elegans*, EGF overexpression and cellular stress induce sleep by activating EGFR in the ALA interneuron (12). Several RFamide family neuropeptides are expressed at high levels in the ALA neuron, including *flp-13* and *flp-24*, and overexpression of these peptides is sufficient to induce sleep (21). It has been reported that EGF-induced sleep is attenuated by ~50% in *flp-13* mutants (22) or in *flp-13*; *flp-24* double mutants (21) and that heat stress–induced sleep in *Drosophila* is suppressed in mutants that lack the RFamide family neuropeptide *fmrifamide* or its receptor (40). These observations suggest that RFamide neuropeptides mediate the sedating effects of EGFR signaling in these invertebrate species. We recently showed that the vertebrate RFamide neuropeptide VF (NPVF), a possible homolog of *flp-13* and *fmrifamide*, is necessary and sufficient to promote sleep in zebrafish (35), similar to our present findings for EGFR signaling. On the basis of these observations, we hypothesized that EGFR signaling promotes sleep, in part, via NPVF signaling in zebrafish. We tested this hypothesis by overexpressing TGF $\alpha$  in *npvf*-mutant animals (35). We found that TGF $\alpha$ -overexpressing animals slept 450% more than their nontransgenic siblings in control *npvf*<sup>+/-</sup> animals ( $P < 0.0001$ , two-way ANOVA, Holm-Sidak test; Fig. 5E and fig. S7K), similar to the effect of TGF $\alpha$  overexpression in WT animals (Fig. 1, E to H). However, TGF $\alpha$  overexpression–induced sleep was suppressed by 50% in *npvf*<sup>-/-</sup> animals compared to TGF $\alpha$  overexpression–induced sleep in their *npvf*<sup>+/-</sup> siblings ( $P < 0.0001$ , two-way ANOVA, Tukey's post hoc test; Fig. 5E and fig. S7K). This result suggests that NPVF signaling plays a key role in mediating EGFR signaling–induced sleep in zebrafish, consistent with observations in *Drosophila* (40) and *C. elegans* (21, 22).

Because loss of *npvf* suppressed EGFR-induced sleep, we hypothesized that NPVF signaling acts downstream of EGFR to regulate normal levels of sleep. To test this hypothesis, we assayed sleep in *egfra* mutants in the presence or absence of RF9, a small-molecule inhibitor of NPFRR1 and NPFRR2, receptors for the mature RFRP1 and RFRP3 peptides derived from the NPVF prepropeptide (41). If EGFR signaling and NPVF signaling act independently to promote sleep, then inhibition of both pathways should result in an additive effect on sleep. Alternatively, if NPVF signaling acts downstream of EGFR, inhibiting both pathways should not have an additive effect. We found that *egfra*<sup>-/-</sup> animals treated with DMSO vehicle control slept significantly less than their *egfra*<sup>+/+</sup> and *egfra*<sup>+/-</sup> siblings during the day (Fig. 5F), as expected. RF9 treatment also decreased sleep in *egfra*<sup>+/+</sup> and *egfra*<sup>+/-</sup> animals (Fig. 5F), as previously reported (35). However, RF9 treatment did not further decrease sleep in *egfra*<sup>-/-</sup> animals compared to *egfra*<sup>-/-</sup> animals treated with DMSO vehicle control (Fig. 5F). These results are consistent with the hypothesis that NPVF signaling acts downstream of EGFR signaling to promote normal sleep levels.



**Fig. 4. TGf $\alpha$  overexpression–induced sleep requires EGFR $\alpha$ .** (A) Total sleep induced following TGf $\alpha$  overexpression quantified during the fourth hour after HS. (B and C) *c-fos* expression in a *Tg(hs:tgfa)* brain at 4 hours after HS shows activation of juxtaventricular cells in the telencephalon, hypothalamus, and hindbrain (arrowhead) (C), but not in the brain of an identically treated WT sibling (B). (D and D') Double FISH of a *Tg(hs:tgfa)* brain fixed at 4 hours after HS shows coexpression of *egfra* and *c-fos* in juxtaventricular cells. Boxed region in (D) is magnified in (D'). Images are 2.9- $\mu$ m-thick (D) and 0.7- $\mu$ m-thick (D') confocal sections. (E to G) TGf $\alpha$  overexpression increased sleep in transgenic animals compared to nontransgenic siblings in *egfra*<sup>+/+</sup>, but not in *egfra*<sup>-/-</sup> animals (red comparisons) (G). TGf $\alpha$  overexpression–induced sleep was significantly decreased in *egfra*<sup>-/-</sup> animals compared to *egfra*<sup>+/-</sup> and *egfra*<sup>+/+</sup> (black comparisons) (G). Yellow bars indicate HS. Pre- and post-HS data are calculated for the day of HS. Data are obtained from six pooled experiments. Bar graphs show mean  $\pm$  SEM. n.s., not significant. (H to L) Juxtaventricular *c-fos* expression that is induced by TGf $\alpha$  overexpression (white arrowheads) is suppressed in *egfra*<sup>-/-</sup> animals (J and K) compared to *egfra*<sup>+/+</sup> animals (H and I) at 4 hours after HS. Hindbrain (Hb)–boxed regions in (H) and (J) are magnified in (I) and (K). (L) Quantification of *c-fos* expression in hindbrain and hypothalamus *egfra*-expressing cells [boxed regions in (H) and (J)]. Data points represent single brains. *n* = number of animals. \**P* < 0.05 and \*\*\**P* < 0.005 by two-way ANOVA with Holm-Sidak test. Te, telencephalon; Hb, hindbrain. Scale bars, 30  $\mu$ m.

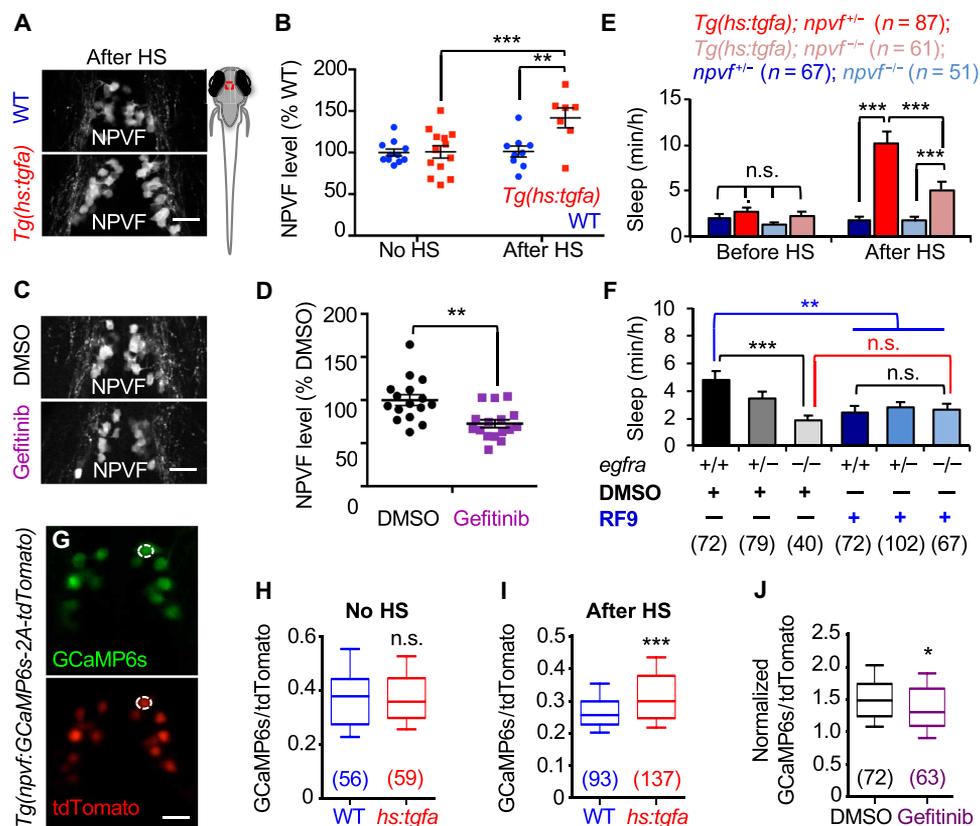
### EGFR signaling regulates NPVF expression

While genetic experiments in *C. elegans* (21, 22) and *Drosophila* (40) and our experiments in zebrafish suggest that EGFR signaling promotes sleep via RFamide neuropeptides, it is unknown how the EGFR and RFamide neuropeptide systems interact. One possibility is that EGFR signaling regulates expression of RFamide neuropeptides. We tested this hypothesis by assaying NPVF mRNA and protein levels in response to gain or loss of EGFR signaling. Using ISH, we found that overexpression of TGf $\alpha$  increased *npvf* mRNA level in endogenous *npvf*-expressing neurons by 68% compared to WT siblings (*P* < 0.05, two-way ANOVA, Holm-Sidak test; fig. S7, A and B). Conversely, treatment with the EGFR antagonist gefitinib decreased *npvf* mRNA level in these neurons by 19% (*P* < 0.05, Student's *t* test; fig. S7, C and D). Similarly, using an NPVF-specific antibody (42), we found that TGf $\alpha$  overexpression increased NPVF protein level in the hypothalamus by 57% compared to WT siblings (*P* < 0.005, two-way ANOVA, Holm-Sidak test; Fig. 5, A and B), while gefitinib treatment decreased NPVF protein level by 17% compared to DMSO-

treated siblings (*P* < 0.01, Student's *t* test; Fig. 5, C and D). In contrast, overexpression of the sleep-promoting neuropeptide Y (NPY) (43) or the wake-promoting neuropeptide Hcrt (28) did not affect NPVF protein levels (fig. S7, E to H). In addition, TGf $\alpha$  overexpression did not affect Hcrt protein levels (fig. S7, I and J). These results suggest that EGFR signaling regulates sleep by specifically regulating expression of NPVF.

### EGFR signaling modulates the activity of *npvf*-expressing neurons

We previously showed that optogenetic and chemogenetic stimulation of *npvf*-expressing neurons is sufficient to induce sleep (35), suggesting that, in addition to affecting NPVF expression, EGFR signaling may also promote sleep by stimulating *npvf*-expressing neurons. To test this hypothesis, we generated *Tg(npvf:GCaMP6s-T2A-tdTomato)* animals, in which *npvf*-expressing neurons express stoichiometric levels of GCaMP6s and tdTomato (44). We used GCaMP6s to monitor calcium levels, a proxy for neuronal



**Fig. 5. EGFR signaling promotes sleep via NPVF.** (A to D) Larval zebrafish schematic. Red box indicates hypothalamic region shown in (A), (C), and (G). Immunostaining with an NPVF-specific antibody reveals higher NPVF protein levels at 2 hours after HS for *Tg(hs:tgfa)* animals compared to WT siblings (A and B) and lower NPVF protein levels in animals fixed at 3 hours after treatment with gefitinib compared to DMSO-treated controls (C and D). Representative images are shown in (A) and (C), and average pixel intensity of NPVF labeling is quantified in (B) and (D); each dot represents one animal. (E) TGFA-induced sleep was partially suppressed in *npvf*<sup>-/-</sup> animals compared to *npvf*<sup>+/-</sup> siblings. Animals were heat-shocked at 3 p.m. Pre- and post-HS quantify sleep for the day of HS. (F) *egfra*<sup>-/-</sup> animals slept less than their *egfra*<sup>+/-</sup> siblings (black comparison), and RF9 suppressed sleep in *egfra*<sup>+/-</sup> animals compared to DMSO-treated *egfra*<sup>+/-</sup> siblings (blue comparison), but RF9 treatment did not further decrease sleep in *egfra*<sup>-/-</sup> animals compared to DMSO-treated *egfra*<sup>-/-</sup> siblings (red comparison). RF9 was added on the afternoon of 5 dpf, and sleep was quantified during the day at 6 and 7 dpf. (G) Representative images from a *Tg(npvf:GCaMP6s-p2A-tdTomato)* animal. An individual NPVF neuron is indicated with a white circle. (H) No difference in GCaMP6s/tdTomato values was observed between non-heat-shocked *Tg(hs:tgfa)* and WT siblings. (I) GCaMP6s/tdTomato values were significantly increased in *Tg(hs:tgfa)* animals compared to WT siblings at 2 hours after HS. (J) Normalized GCaMP6s/tdTomato values were significantly decreased in WT gefitinib-treated animals compared to DMSO-treated siblings. Normalized GCaMP6s/tdTomato = GCaMP6s/tdTomato value in each neuron at 2 hours after the addition of either gefitinib or DMSO divided by the value in the same neuron before treatment is shown. Bar graphs (E and F) and lines (B and D) indicate mean  $\pm$  SEM. Numbers in parentheses indicate number of animals (E and F) or neurons (H to J). \* $P < 0.05$ , \*\* $P < 0.01$ , and \*\*\* $P < 0.001$  by two-way ANOVA with Holm-Sidak test (E and F) or by two-tailed Wilcoxon matched-pairs signed-rank test (H to J). Scale bars, 20  $\mu$ m.

activity, in *npvf*-expressing neurons and normalized GCaMP6s fluorescence intensity values to tdTomato fluorescence to correct for any changes in transgene expression or movement artifacts during live imaging. To test the hypothesis that EGFR signaling is sufficient to stimulate *npvf*-expressing neurons, we mated *Tg(npvf:GCaMP6s-T2A-tdTomato)* to *Tg(hs:tgfa)*<sup>+</sup> fish, imaged GCaMP6s and tdTomato in 5-dpf animals (Fig. 5G), and then used PCR genotyping to identify *Tg(hs:tgfa)* and WT siblings. Control animals not subjected to a HS displayed no difference in GCaMP6s/tdTomato fluorescence between *Tg(hs:tgfa)* and WT siblings (Fig. 5H). However, at 2 hours after HS, *Tg(hs:tgfa)* animals displayed significantly higher levels of GCaMP6s/tdTomato fluorescence compared to WT siblings ( $P < 0.005$ , two-tailed Wilcoxon matched-pairs signed-rank test; Fig. 5I), suggesting that TGFA overexpression stimulates *npvf*-expressing neurons.

We next tested whether EGFR signaling is necessary for normal activity levels of *npvf*-expressing neurons by inhibiting EGFR sig-

naling. We first quantified baseline GCaMP6s/tdTomato fluorescence levels in *Tg(npvf:GCaMP6s-T2A-tdTomato)* animals. We then treated animals with either gefitinib or DMSO vehicle control and reimaged the same neurons 90 min later. Last, we calculated normalized GCaMP6s/tdTomato fluorescence values by dividing posttreatment values by baseline values. Consistent with the TGFA overexpression result, we found that inhibition of EGFR signaling using gefitinib treatment decreased normalized GCaMP6s/tdTomato fluorescence values compared to DMSO-treated siblings ( $P < 0.05$ , two-tailed Wilcoxon matched-pairs signed-rank test; Fig. 5J). This result suggests that EGFR signaling is required to maintain normal activity levels of *npvf*-expressing neurons. Together, these observations suggest that EGFR signaling promotes sleep, at least in part, by stimulating *npvf*-expressing neurons. More broadly, our results indicate that EGFR signaling promotes sleep via RFamide neuropeptides in both invertebrates and vertebrates.

### Common genetic variants at EGFR signaling–associated genes affect human sleep

Having shown that regulation of sleep by EGFR signaling is conserved between invertebrate and vertebrate animal models, we next asked whether this conservation extends to humans. We tested for an association between self-reported sleep traits and genetic variation in EGFR signaling–associated gene regions in participants of European ancestry from the U.K. Biobank (tables S1 and S2) (24). *ERBB4* is a member of the *ERBB* family of receptor tyrosine kinases that can heterodimerize with *ERBB1/EGFR* and can be activated by EGF family ligands (45) and thus participates in EGFR signaling. We found that genetic variation (rs7607363 G allele) at the *ERBB4* genomic locus is associated with increased frequency of excessive daytime sleepiness (EDS) ( $P = 8.00 \times 10^{-9}$ , linear regression analysis; Fig. 6, A and B, and table S1). We also found that genetic variation (rs1846644 C allele) at the genomic locus of kinase suppressor of ras 2 (*KSR2*), a scaffolding protein with established roles in EGFR/MAPK signaling (46), is associated with increased sleep duration ( $P = 5.30 \times 10^{-9}$ , linear regression analysis), increased frequency of daytime napping ( $P = 2.00 \times 10^{-41}$ , linear regression analysis), and increased frequency of EDS ( $P = 2.50 \times 10^{-27}$ , linear regression analysis) (Fig. 6, A and C to E, and table S1). In addition, pathway analysis for EDS revealed associations at EGFR signaling pathways enriched in the sleep propensity subtype of daytime sleepiness (4). Together, these results suggest that the identified genetic variants in the *ERBB4* and *KSR2* loci associate with measures of increased sleep.

Expression quantitative trait loci (eQTL) analysis of the *ERBB4* variant rs7607363 showed increased *ERBB4* expression associated with the minor allele that is associated with increased sleep (fig. S8A) and no association with expression levels of neighboring genes, suggesting a gain of function associated with this single-nucleotide polymorphism (SNP). The lead SNPs in the *ERBB4* and *KSR2* loci, as well as several SNPs highly linked to the lead SNPs, are aligned within each gene, with recombination hotspots separating them from neighboring genes (Fig. 6, B to E). These observations provide supporting evidence that the SNPs are linked to *ERBB4* and *KSR2*. However, since SNPs may influence distant genes, we cannot be absolutely certain that *ERBB4* or *KSR2* is the causal gene at each locus. Despite this caveat, these results suggest that common variants in EGFR signaling pathway components underlie some of the natural variation observed in human sleep.

To validate the association of *KSR2* and *ERBB4* with sleep traits in vivo, we treated zebrafish with small-molecule inhibitors of these proteins (fig. S8). First, we treated WT zebrafish with APS-2-79 (fig. S8, J to Q), a small molecule that stabilizes *KSR2* in an inactive state (46). Treatment with this drug resulted in decreased sleep during both the day (–38%) and night (–14%) compared to DMSO vehicle-treated controls (day,  $P < 0.05$ ; night,  $P < 0.005$ , Student's *t* test; fig. S8, M and N). These effects were primarily due to shorter sleep bouts at night and a trend of higher activity and waking activity during the day (fig. S8, J to L and P). This phenotype is consistent with the finding that *Ksr2* knockout mice exhibit increased locomotor activity (47). Next, we treated WT zebrafish with spironolactone (fig. S8, B to I), an *ERBB4* small-molecule inhibitor (48). Treatment with this drug resulted in a 7% increase in locomotor activity and a 34% decrease in sleep during the day (activity,  $P < 0.05$ ; sleep,  $P < 0.01$ , Student's *t* test), with no significant effect at night, compared to DMSO vehicle-treated siblings (fig. S8, B, C, E, and F). These phenotypes were associated with higher waking activity and

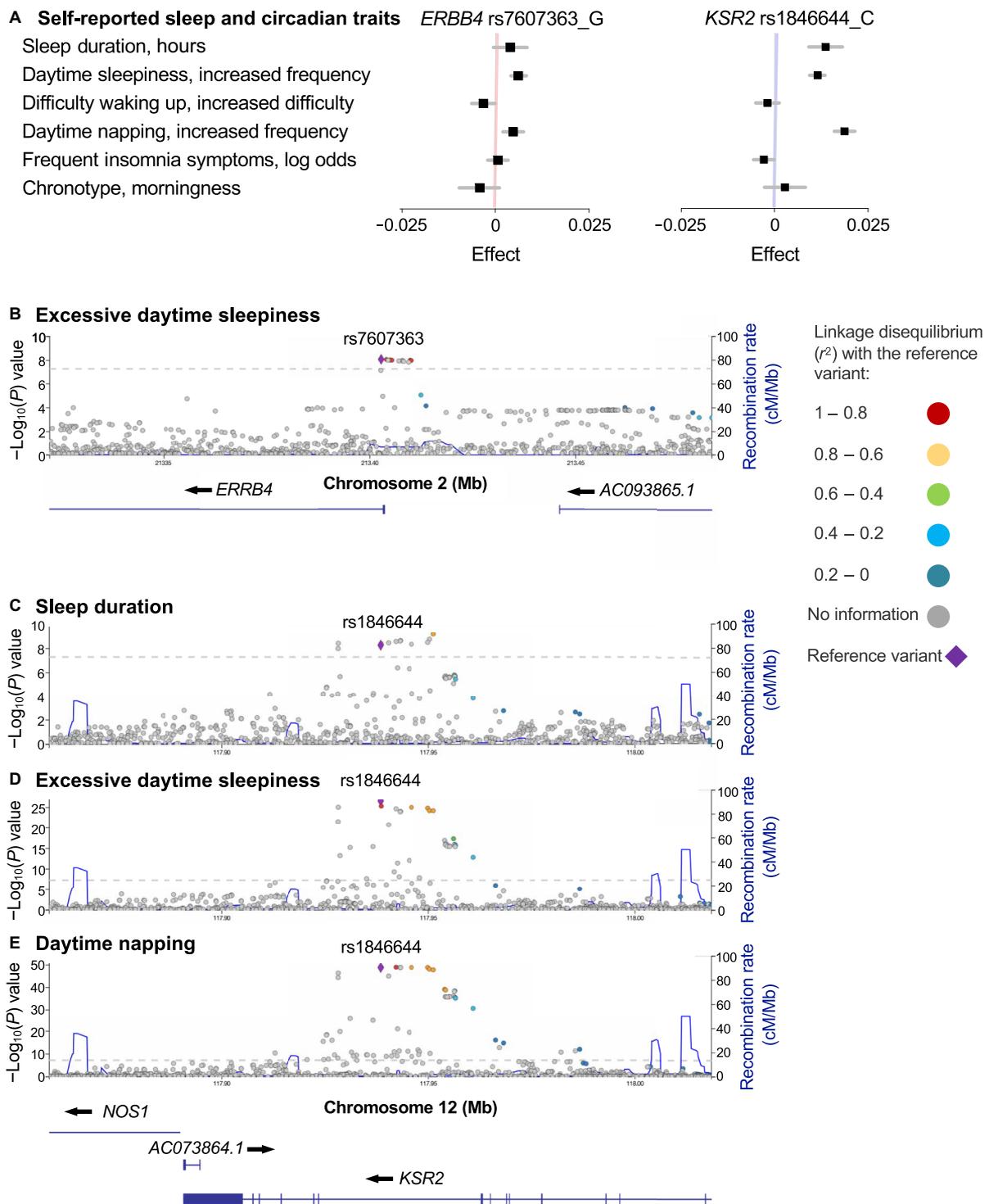
fewer sleep bouts during the day (fig. S8, D and G). Thus, inhibition of *KSR2* or *ERBB4* phenocopies loss of EGFR signaling, although the spironolactone result should be interpreted with caution because this drug has been shown to also target other proteins (48). These results validate a genetic pathway that underlies some of the observed variation in human sleep quality and quantity.

### DISCUSSION

There is abundant evidence that sleep is behaviorally conserved throughout evolution (9), and genetic mechanisms that regulate sleep have been described in both vertebrate and invertebrate species (9). However, while it has been shown that genes that regulate circadian rhythms and thus affect sleep timing are conserved in invertebrates and humans (1, 9), there is relatively little clear and unambiguous evidence for conserved mechanisms that regulate sleep in both vertebrates and invertebrates, including humans. As a result, the extent to which mechanisms that underlie sleep evolved independently in vertebrates and invertebrates remains unclear. Identifying such mechanisms is important not only from an evolutionary standpoint but also because these mechanisms would likely play a central role in regulating sleep in all organisms. Here, we describe a genetic mechanism that regulates sleep in both invertebrates and vertebrates and provide evidence that these findings may extend to humans.

We report that EGFR signaling regulates sleep in the zebrafish, a diurnal vertebrate. We found that TGF $\alpha$  overexpression increased sleep, whereas mutation of EGFR or its ligands, or pharmacological inhibition of EGFR signaling, decreased sleep. The reduced sleep resulting from inhibition of EGFR signaling was due to defects in both initiation and maintenance of sleep, consistent with the human sleep disorder insomnia. While gain- and loss-of-function studies have shown that EGFR signaling is both necessary and sufficient for normal sleep levels in *C. elegans* and *Drosophila* (12, 13, 16), our results extend these invertebrate studies by demonstrating that this mechanism is conserved in a vertebrate and identify a previously undescribed role of EGFR signaling in regulating sleep homeostasis. In both zebrafish and *Drosophila* (13), animals with elevated EGFR signaling had more and longer sleep bouts, whereas pharmacological inhibition of EGFR signaling resulted in shorter sleep bouts, indicating that EGFR signaling promotes consolidation of the sleep state. Combined with our discovery that common variants in two genes that participate in EGFR signaling are associated with variation in human sleep, this study provides one of the first clear examples of a mechanism that regulates sleep in both invertebrate and vertebrate species and suggests that these findings may extend to humans. These results reveal a clear evolutionary link between invertebrate and vertebrate sleep, suggest that EGFR signaling participates in an ancient and central sleep regulatory mechanism, and provide a basis to explore central and ancient mechanisms that regulate vertebrate sleep.

In contrast to our findings using zebrafish and previous invertebrate studies (12, 13, 16, 19), the role of EGFR signaling in mammalian sleep is unclear. While EGFR signaling has been shown to regulate various rodent behaviors (14, 15), attempts to explore its requirement for sleep have been confounded by the embryonic lethality of murine *Egfr*-null mutants (32). Consistent with a role for EGFR signaling in rodent sleep, one study using mice homozygous for the *Egfr*<sup>wa2</sup> hypomorphic allele, which produces a receptor with reduced tyrosine kinase activity, observed increased daytime locomotor activity (14).



**Fig. 6. Genetic variants at *KSR2* and *ERBB4* associate with human sleep quality and quantity.** (A) Forest plot of associations between genetic variant rs7607363 in *ERBB4* and rs1846644 in *KSR2* with sleep quality and quantity measures. Black boxes indicate effect estimate, and gray lines represent 95% confidence interval. Data are based on 453,964 human subjects of European ancestry in the U.K. Biobank. Full results for each single-nucleotide polymorphism (SNP) are shown in table S1. (B) Regional association plot for genome-wide significant association between *ERBB4* rs7607363 and EDS. (C to E). Regional association plots for genome-wide significant association between *KSR2* rs1846644 and sleep duration (C), EDS (D), and daytime napping (E). Genes within each region are shown in the lower panels. Blue lines indicate recombination rate. Filled circles show the  $-\log_{10}P$  value for each SNP, with the named SNP shown in purple. Additional SNPs in the locus are colored according to correlation ( $r^2$ ) with the lead SNP, estimated by LocusZoom based on the CEU HapMap haplotypes. Dashed lines represent the genome-wide significance threshold ( $P < 5 \times 10^{-8}$ ).

However, this allele contains a linked hypomorphic mutation in *Wnt3a* (14) that may confound sleep studies. Wnt signaling is required for brain development (49) and has been implicated in regulating sleep duration in *Drosophila* (18). Mice lacking TGF $\alpha$  or EGF are viable and fertile, but sleep studies using these mutants have not been described. Reports of EGFR gain-of-function manipulations on mammalian sleep are also inconclusive. Acute intracerebroventricular injection of EGF into rabbits was sufficient to increase sleep (20), although the effect required the use of high EGF levels that are unlikely to be physiologically relevant. In contrast, infusion of TGF $\alpha$  had no effect on sleep duration in hamsters, although it did suppress wheel running (14, 15). These conflicting gain-of-function results could be due to the use of small numbers of animals, nonphysiological peptide concentrations resulting in nonspecific behavioral effects, or the use of in vitro synthesized peptides that may lack posttranslational modifications required for function. Other indirect evidence supporting a role for EGFR signaling in mammalian sleep was provided by the recent finding that inhibition of ERK results in reduced sleep in rodents (39). However, ERK participates in many signal transduction pathways, so it is unclear whether this effect is specific to EGFR signaling. Thus, before our study, it remained unclear whether EGFR signaling plays a role in vertebrate sleep.

Consistent with a conserved role for EGFR in sleep is its expression in the hindbrain and hypothalamus in both rodents and zebrafish. These brain structures are anatomically and molecularly conserved in mammals and zebrafish (10) and play prominent roles in sleep regulation (50). In rodents, *Egfr* is expressed in juxtaventricular radial-glia-like cells named tanycytes (51), which are responsive to photoperiod and regulate energy balance (52), and *Tgfa* is expressed in nearby ependymal cells (27, 53). Similarly, in zebrafish, *egfra* is expressed in juxtaventricular cells that express the glial cell markers *sox2* and glial fibrillary acidic protein, and *tgfa* is expressed in an adjacent *sox2*-expressing glial cell population. These observations suggest a conserved neuroanatomical basis for *Tgfa/Egfr*-regulated sleep. This possibility is supported by the observations that TGF $\alpha$  overexpression in zebrafish results in *c-fos* expression in *egfra*-expressing juxtaventricular cells, while intracerebroventricular injection of either TGF $\alpha$  or EGF in hamsters similarly results in ERK activation in *Egfr*-expressing juxtaventricular tanycytes (15, 51). Studies in mammals have shown that these tanycytes can sense and transport molecules from the cerebrospinal fluid (CSF) to brain regions implicated in regulating sleep (54) and that CSF contains sleep-inducing factors (55). Together, these observations suggest that *Tgfa/Egfr* signaling may regulate sleep through interactions between CSF, specialized glial cells, and neurons. This possibility contrasts with most previously described neuropeptidergic and neuromodulatory mechanisms that regulate sleep, which act via communication between neurons, but is consistent with the model that tanycytes act as gatekeepers that monitor homeostatic signals in CSF (52), including those that regulate sleep (55).

Before this study, the molecular and cellular mechanisms through which EGFR regulates invertebrate sleep were unclear. It has been shown that EGF-induced sleep in *C. elegans* is suppressed by mutation of the RFamide neuropeptides *flp-13* and *flp-24* (21, 22). Likewise, in *Drosophila*, cellular stress-induced sleep, which is thought to be mediated by EGFR signaling, is suppressed by mutation of the RFamide neuropeptide *fmrfamide* (40). However, these studies only demonstrated genetic interactions, and the mechanism linking the EGFR and RFamide systems was unknown. We previously showed that NPVF, a member of the vertebrate RFamide family of neuropeptides, as well

as *npvf*-expressing neurons, are necessary and sufficient for normal sleep amount in zebrafish (35), similar to our findings for EGFR signaling in this study. Our demonstration that EGFR signaling promotes sleep in part via the NPVF system in zebrafish suggests that this aspect of EGFR-regulated sleep is conserved, although it is unclear whether the vertebrate and invertebrate RFamide proteins are true orthologs owing to low sequence conservation (35). In addition to showing a genetic interaction, we found that EGFR signaling regulates both NPVF expression and the activity of *npvf*-expressing neurons, providing a molecular and cellular link between the two pathways. However, we did not observe *egfra* expression in *npvf*-expressing neurons, so it remains unclear how EGFR signaling affects NPVF expression and the activity of *npvf*-expressing neurons. One possibility is that *egfra* is expressed in *npvf*-expressing neurons at levels too low to detect using ISH. Alternatively, presently unknown genes and/or neurons may act to link the EGFR and NPVF systems. Further studies are also needed to determine whether EGFR signaling regulates the expression of invertebrate RFamide neuropeptides and the activity of neurons that express them, and whether mechanisms that act upstream of EGFR signaling to promote invertebrate sleep, such as cellular stress, are conserved in vertebrates.

Are these experimental findings in model systems relevant to humans? Our analysis of several self-reported sleep traits ( $N = 450,000$ ) in humans of European ancestry from the U.K. Biobank (4–8, 24) identified genome-wide significant associations of several sleep traits with common variants in the *KSR2* and *ERBB4* genomic loci. *KSR2* is a scaffolding protein that enables interaction of the EGFR signaling components RAF, MEK, and ERK, and thus is crucial for EGFR signaling (46). *ERBB4* is a member of the *ERBB1/EGFR* receptor tyrosine kinase subfamily that binds several EGFR ligands and forms heterodimers with other *ERBB* family members, including EGFR (45). Our results suggest that common variants in EGFR signaling pathway genes underlie some of the variation observed in human sleep, including traits such as sleep duration, daytime napping, and EDS, suggesting gain-of-function effects leading to an increased sleep propensity. We found that genetic loss of *egfra*, *egf*, or *tgfa*, as well as pharmacological inhibition of EGFR, *KSR2*, or *ERBB4*, results in reduced sleep in zebrafish, consistent with the human data. Our results are consistent with a report that elevated TGF $\alpha$  serum levels are associated with fatigue in humans (56). Furthermore, our results suggest that the disrupted sleep commonly observed in RASopathies (57), a group of human developmental syndromes caused by mutations in genes involved in Ras-MAPK signaling, may be directly caused by abnormal EGFR signaling, rather than occurring as a secondary consequence of abnormal development. Last, we note that except for narcolepsy, which is due to loss of *Hcrt* neurons (1), and circadian disorders that affect sleep timing (1), the basis for human sleep disorders is largely unknown, and our study provides one of the first examples of a genetic mechanism that likely affects human sleep. Further studies using animal models, such as zebrafish, will provide a deeper mechanistic understanding of the role of EGFR signaling in sleep and how abnormal EGFR signaling results in sleep disorders.

## MATERIALS AND METHODS

### Experimental model

Zebrafish experiments and husbandry followed standard protocols (58) in accordance with Caltech Institutional Animal Care and Use

Committee guidelines. Larval zebrafish were studied before the onset of sexual differentiation, and all behavioral experiments were performed using siblings with the same genetic background, differing only in the presence of a transgene, mutation of a specific gene, or treatment with drugs and appropriate vehicle control. The age of animals used in each experiment is described in the manuscript, in each figure legend, or in this section. All experiments used animals from a TLAB strain. Unless otherwise noted, larvae were raised and behaviorally tested in a 14-hour light/10-hour dark cycle, with white lights on from 9 a.m. to 11 p.m.

### Transgenic and mutant zebrafish

#### **Tg(*hs:tgfa*)**

Zebrafish *tgfa* complementary DNA (cDNA) was isolated using 5' and 3' RACE (FirstChoice RLM-RACE, AM1700, Thermo Fisher Scientific), and the open reading frame (ORF) was cloned downstream of the zebrafish *hsp70c* promoter (59) in a vector containing flanking I-SceI endonuclease recognition sites. We generated stable transgenic animals by injecting plasmids with I-SceI (New England Biolabs Inc.) into the cell of embryos at the one-cell stage (60). To identify transgenic founders, we outcrossed potential founders, heat-shocked their progeny at 5 dpf, fixed animals 30 min later, and performed *tgfa* ISH. At least two lines that produced strong and ubiquitous overexpression and lacked pre-HS expression were selected and tested for behavioral phenotypes. The two independent *Tg(hs:tgfa)* transgenic lines produced similar behavioral phenotypes, but all data in the paper are from the line that produced stronger phenotypes. Transgenic animals were identified by PCR using the primers 5'-CGGGCCACCATGATGTAT-3' and 5'-GGTTTGTC-CAAACATCAATGT-3' with a product size of 647 base pairs (bp).

#### **Tg(*npvf:GCaMP6s-2A-tdTomato*)**

Using Gibson assembly, the 3.9-kb *npvf* promoter (35) was cloned upstream of GCaMP6s (61), a T2A sequence (62), and tdTomato. Transgenic lines were generated using the I-SceI method (60).

#### ***tgfa* mutant**

Using TALENs (30), we isolated a zebrafish *tgfa* mutant that contains a 7-bp deletion (ORF nucleotides 25 to 31: 5'-ACAATAT-3') that shifts the translational reading frame after the eighth amino acid and introduces a premature stop codon that should truncate the protein from 189 to 61 amino acids (fig. S2A). The predicted mutant protein lacks the signal peptide, protease cleavage sites, and epidermal growth factor-like domain and should thus be nonfunctional. *Tgfa*<sup>-/-</sup> zebrafish are viable, and homozygous mutant males and females are fertile, with no gross morphological abnormalities. Mutant animals were genotyped using the primers 5'-GACGACGCT-GAGAATCTTTCATC-3' and 5'-CAATCGTTTTGGTCTTTGCAT-3', which produce a 188- or 181-bp band for the WT or mutant allele, respectively.

#### ***egf* mutant**

Using CRISPR-Cas9 (31), we isolated a zebrafish *egf* mutant that contains a 19-bp insertion and a 2-bp deletion after ORF nucleotide 426 that shifts the translational reading frame and introduces a premature stop codon that should truncate the protein from 1177 to 457 amino acids (fig. S2B). The predicted mutant protein lacks five EGF-like domains and the transmembrane domain and should thus be nonfunctional. *Egf*<sup>-/-</sup> zebrafish are viable, and homozygous mutant males and females are fertile, with no gross morphological abnormalities. Mutant animals were genotyped using the primers

5'-GCTGTGGATTGTGATCATGC-3' and 5'-TGCAGAAAA-CAGCTGAAA-3', which produce a 204- or 221-bp band for the WT or mutant allele, respectively.

#### ***egfra* mutant**

Using CRISPR-Cas9 (31), we generated a zebrafish *egfra* mutant that contains an 11-bp deletion and a 27-bp insertion after nucleotide position 792 that shifts the translational reading frame and introduces a premature stop codon upstream of the transmembrane and intracellular domains, and should thus produce a non-functional protein (fig. S2C). *egfra*<sup>-/-</sup> animals were viable and lacked obvious morphological abnormalities. *egfra*<sup>-/-</sup> males were fertile, but *egfra*<sup>-/-</sup> females were seemingly infertile. Mutant animals were genotyped using the primers 5'-ACTTCCAGGATGAAGGGACA-3' and 5'-CTTACGTGGGCATGTCTTGA-3', which produced a 137- or 153-bp band for the WT or mutant allele, respectively.

*Tg(hs:hcr1) zf12Tg* (28), *Tg(hs:npv) ct853Tg* (43), and *npvf* ct845 mutant (35) zebrafish have been previously described.

### Zebrafish behavioral assays

#### **Locomotor activity assay**

Individual larvae were placed into each well of a 96-well plate (7701-1651, Whatman) containing 650  $\mu$ l of E3 embryo medium [5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl<sub>2</sub>, and 0.33 mM MgSO<sub>4</sub> (pH 7.4)] at 4 dpf and acclimated overnight in the video tracker before behavioral analysis beginning at 5 dpf. Plates were sealed with an optical adhesive film (4311971, Applied Biosystems) to prevent evaporation, except for drug experiments. Animals in unsealed plates can appear to have higher amounts of sleep compared to those in sealed plates because as water evaporates from each well, it forms a concave meniscus at the top of the well, which creates a slight visual distortion that causes the tracking system to miss some animal movements. To mitigate this issue, each well was filled to the top with water each morning during unsealed plate experiments. The sealing process introduces air bubbles in some wells, which are discarded from analysis. Animals were blindly assigned a position in the plate and were genotyped after the behavioral experiment was complete. Locomotor activity was monitored using a video tracking system (Viewpoint Life Sciences) with a Dinion one-third inch monochrome camera (Dragonfly 2, Point Grey) fitted with a fixed-angle megapixel lens (M5018-MP, Computar) and infrared filter. For HS-induced overexpression experiments, larvae were heat-shocked at 37°C for 1 hour starting at 3 p.m., 4 p.m., or 9:45 p.m. at 5 dpf. The movement of each larva was captured at 15 Hz and recorded using the quantization mode with 1-min time bins. The 96-well plate and camera were housed inside a custom-modified ZebraBox (Viewpoint Life Sciences) that was continuously illuminated with infrared LEDs and illuminated with white LEDs from 9 a.m. to 11 p.m., except as noted in experiments where animals were entrained in a 14-hour light/10-hour dark cycle and then assayed in constant light or constant dark. The 96-well plate was housed in a chamber filled with recirculating water to maintain a constant temperature of 28.5°C. The parameters used for detection were as follows: detection threshold, 15; burst, 29; freeze, 3, which were determined empirically. Data were processed using custom PERL and MATLAB (MathWorks Inc.) scripts, and statistical tests were performed using Prism (GraphPad) for ANOVA analysis and Excel (Microsoft) for Student's *t* test.

A movement was defined as a pixel displacement between adjacent video frames preceded and followed by a period of inactivity of at least 67 ms (the limit of temporal resolution). Any 1-min period with

no movement was defined as 1 min of sleep based on arousal threshold changes (28). A sleep bout was defined as a continuous string of sleep minutes. Sleep latency was defined as the length of time from lights on or off to the start of the first sleep bout. Average activity was defined as the average amount of locomotor activity in seconds per hour, including sleep bouts. Average waking activity was defined as the average amount of activity in seconds per hour, excluding sleep bouts.

### Arousal threshold assay

We used an Arduino-based automated driver to control two solenoids (28P-I-12, Guardian Electric) that delivered a tap over an intensity range of 4095 settings (from 0.01 to 40.95) to a 96-well plate in the video tracker (35). We used taps ranging from a power setting of 1 to 36.31. Taps of 14 different intensities were applied in a random order with an intertrial interval of 1 min during the day or night for *Tg(hs:tgfa)* or gefitinib and erlotinib experiments, respectively. Previous studies showed that a 15-s intertrial interval is sufficient to prevent behavioral habituation (63). The background probability of movement was calculated by identifying for each genotype the fraction of animals that moved 5 s before all stimuli delivered during an experiment (14 different tap powers  $\times$  30 trials at each intensity = 420 data points per animal). This value was subtracted from the average response fraction value for each tap event. The response of larvae to the stimuli was monitored using the video tracking software and was analyzed using MATLAB and Excel. Statistical analysis was performed using the Variable Slope log(dose) response curve fitting module of Prism.

### SD assay

At 4 dpf, up to 84 animals were placed in a 96-well plate that had been cut into two separable halves that can be rejoined to fit inside the video tracker. Animals were monitored for 48 hours (5 to 6 dpf) to measure baseline sleep. Experiments in which significant differences were observed between the two halves of the plate during the 24 hours before perturbation were excluded from analyses. Immediately before the start of the third night of analysis (7 dpf), a half-plate from each of two video trackers was transferred and fastened to the center of a 245-mm<sup>2</sup> square dish (431301, Corning) containing 20 glass beads, 6 plastic marbles, and 8 fixed pegs that prevent the marbles and beads from only circling the perimeter of the plate during nutation. For night-perturbed animals, the plates were wrapped in aluminum foil to shield the animals from light, fastened to a nutator (B3D1020, Benchmark Mini BioMixer), and nutated in a dark 28°C incubator for 6 hours. For day-perturbed experiments, animals of the same developmental stage were treated the same, except the plates were exposed to light during perturbation during the day at 7 dpf. In addition to mechanical stimulation, we simultaneously subjected the animals to auditory stimuli by playing songs from a percussive recording in random order (*Death Magic*, Health) at maximum volume using a portable Bluetooth speaker (DKnight MagicBox II) that was housed in the incubator. For each SD half-plate, the corresponding half remained inside the video tracker to serve as a nonperturbed control. At the end of the 6-hour perturbation, the SD half-plates were rejoined with their nonperturbed control halves in their respective video tracker. Night-perturbed samples were kept in the dark when transferred from the SD assay into video trackers, and both night- and day-perturbed samples were maintained in dark for the rest of the experiment to avoid masking effects of light on any sleep recovery response.

### Behavioral data analysis

`sort_fish_sttime_192.pl` is a Perl script (28) that converts data acquired by the Viewpoint system to a format that is useful for MATLAB analysis. `perl_batch_192well.m` is a MATLAB script that allows the `sort_fish_sttime_192.pl` script to run using MATLAB. `TapAnalysis.m` is a MATLAB script that analyzes tapping assay data and generates a table that lists the number of animals that moved following each tap. `VT_analysis.m` is a MATLAB script [modified from (28)] that performs sleep/wake analyses. These scripts are posted online (35), and detailed instructions on their use will be provided upon request.

### Histology

#### In situ hybridization

Samples were fixed in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) for 16 hours at room temperature. ISH was performed using digoxigenin (DIG)-labeled antisense riboprobes (DIG RNA Labeling Kit, Roche) (35). Double FISH was performed using DIG- and fluorescein-labeled riboprobes and the TSA Plus Fluorescein and Cyanine 3 Systems kit (PerkinElmer). Probes specific for *npvf* (35) and *c-fos* (35) have been described. The *tgfa* probe was transcribed using a PCR product amplified from zebrafish cDNA using the primers 5'-CGCGTGCCTTCATCTTTATT-3' (forward) and 5'-TCCCCTGCCCATATTGAAC-3' (reverse) and then serially amplified with the same forward primer and a T7 sequence (lowercase) added to the primer 5'-ggatcctaatacactcattataggTCCCCTGCCCATATTGAAC-3' (reverse). The *egfra* probe was transcribed using a PCR product amplified from zebrafish cDNA using the primers 5'-TCTGCTTGACCTCCACAGTG-3' (forward) and 5'-CAAGGGCAAAGTACCTGATCG-3' (reverse) and then serially amplified with the same forward primer and a T7 sequence (lowercase) added to a reverse primer located within that amplicon: 5'-ggatcctaatacactcattataggTGTAAACAAGCTGAAGTAACGTGC-3'. The *sox2* probe was transcribed using a PCR product amplified from zebrafish cDNA using the primers 5'-GGTCAAC-CAGAGGATGGACA-3' (forward) and 5'-AACTGCGTACAGGCCACATT-3' (reverse) and then serially amplified with the inner primer 5'-GGCCGACAGCTACGCGCATATGAA-3' (forward) and a T7 sequence (lowercase) added to a primer located within that amplicon: 5'-taatacactcattataggGCTGGAGAGTGCCTCTGTTC-3'.

#### Immunohistochemistry

Samples were fixed in 2% PFA in PBS overnight at 4°C and then washed with 0.25% Triton X-100/PBS (PBTx). Brains were dissected and blocked in 2% goat serum/2% DMSO/PBTx overnight at 4°C. Antibody incubations were performed in blocking solution overnight at 4°C using rabbit anti-NPVF (1:200; ab122738, Abcam) or rabbit anti-Hcrt (1:500; AB3704, MilliporeSigma) primary antibodies and Alexa Fluor 568 secondary antibody (1:500; Life Technologies). Samples were mounted in Vectashield (H-1000, Vector Labs) and imaged using a Zeiss LSM 780 confocal microscope.

### Pharmacology

A dose-response experiment was first performed for each drug to determine its optimal concentration. Each drug was dissolved in DMSO and then added to E3 medium. The following drug concentrations were tested: DMSO vehicle control, 0.1, 0.5, 1, 5, 10, 50, and 100  $\mu$ M. Gefitinib (13166, Cayman Chemical) and erlotinib (10483, Cayman Chemical) were not soluble at 50 or 100  $\mu$ M, and treatment

of zebrafish with 10  $\mu\text{M}$  of either drug produced developmental defects in a small number of animals. Treatment with 5  $\mu\text{M}$  gefitinib or erlotinib did not cause developmental defects but did result in increased locomotor activity and decreased sleep compared to DMSO vehicle-treated controls. Optimal concentrations for other drugs were determined in a similar fashion and were used in experiments at the following concentrations: 15  $\mu\text{M}$  U0126 (1144, Tocris), 3  $\mu\text{M}$  SL327 (1969, Tocris), 10  $\mu\text{M}$  RF9 (3672, Tocris), 0.2  $\mu\text{M}$  spironolactone (2968, Tocris), and 5  $\mu\text{M}$  APS-2-79 (S8355, Selleck Chemicals). Drug solutions were freshly prepared before each experiment. Controls were exposed to DMSO at concentrations identical to experimental conditions, at no more than 0.1% DMSO. Drugs and DMSO vehicle were loaded into alternating rows of a 96-well plate or loaded on opposite sides of a 96-well plate. WT larvae from the same clutch were then added to the plate. For RF9 experiments and gefitinib constant dark and constant light experiments, drug and DMSO vehicle were loaded into separate sides of a 96-well plate with an empty row in between. WT larvae from the same clutch were added to the plate, and plates were sealed with an optical adhesive film (4311971, Applied Biosystems) to prevent evaporation.

### Quantitative polymerase chain reaction

Zebrafish were raised on a 14-hour light/10-hour dark cycle at 28.5°C with lights on at 9 a.m. and lights off at 11 p.m. At 5 dpf, total RNA was collected using TRIzol reagent (15596-026, Life Technologies) from 20 pooled larvae per biological replicate every 6 hours for 36 hours, with three independent biological replicates at each time point. cDNA was synthesized from 5  $\mu\text{g}$  of total RNA using SuperScript III Reverse Transcriptase (18080-051, Invitrogen), and qPCR was carried out using SYBR Green Master Mix (4364346, Life Technologies) using an ABI PRISM 7900HT (Life Technologies) instrument using the primers 5'-GTGTGTGGTGGCAGTGTC-3' and 5'-CCAA-CAGGAGAGGGTGTGAC-3' for *tgfa* and 5'-ATCCAGAC-CCCAATACAAC-3' and 5'-GGGAGACTCTGCTCCTTCT-3' for *per1b*. *ef1a* was used as a reference gene using the primers 5'-CAGCTGATCGTTGGAGTCAA-3' and 5'-TGTATGCGCT-GACTTCCTTG-3'. Relative expression levels were determined using the  $\Delta\Delta C_t$  method (64), normalized to the highest cycle threshold ( $C_t$ ) value for each gene.

### Imaging

Dissected brains were coverslip-mounted in Vectashield (H-1000, Vector Labs) or 80% glycerol in PBS and imaged using a compound microscope [AxioImager with EC Plan-Neofluar 10 $\times$ /0.30-numerical aperture (NA) air objective or Plan-Apochromat 20 $\times$ /0.8-NA air objective, Zeiss] for chromogenic ISH samples or using a confocal microscope (LSM 780 with Plan-Apochromat 10 $\times$ /0.45-NA air objective, LD LCI Plan-Apochromat 25 $\times$ /0.8-NA Imm Corr objective, or LD C-Apochromat 40 $\times$ /1.1-NA water objective; Zeiss) for FISH samples. Fluorescein and cyanine 3 were imaged in separate channels with 488- and 561-nm lasers, respectively. Confocal images were displayed as an optical section of approximately 1-airy unit thickness.

For *c-fos* and *npvf* mRNA quantification, chromogenic ISH was performed using *c-fos* and *npvf* riboprobes on 5-dpf dissected brains, with experimental and control samples developed in the same tube blinded to genotype (Fig. 4, H to L, and fig. S7, A and B) or with experimental and control samples developed in different tubes in parallel and treated identically (fig. S7, C and D). Samples were imaged using a Zeiss compound microscope and then geno-

typed by PCR. Regions of interest (ROIs) were drawn over hypothalamic or hindbrain regions containing ISH signal, along with a nonlabeled area of the brain that was selected for background correction. The ISH signal was quantified using ImageJ (65), with background subtraction correction. Average pixel intensity within the ROI was quantified blind to genotype or treatment.

For NPVF (Fig. 5, A and B, and fig. S7, E to H) and Hcrt (fig. S7, I and J) protein quantification, experimental and control brains were immunostained in the same tube, imaged using a Zeiss confocal microscope, and then genotyped by PCR. For NPVF protein quantification on drug-treated animals (Fig. 5, C and D), experimental and control brains were immunostained in different tubes in parallel and treated identically. An ROI containing all NPVF soma was drawn in each brain hemisphere to calculate mean fluorescence intensity, and an unlabeled area in the image was selected for background correction using ImageJ (65). ROIs were kept the same size and shape but moved to contain as many NPVF soma as possible across images. Average pixel fluorescence intensity was quantified blind to genotype or treatment.

For GCaMP6s experiments, 5-dpf *Tg(npvf:GCaMP6s-2A-tdTomato)* animals were mounted in low-melting agarose and then imaged once per minute for 5 min. The average pixel intensity of GCaMP6s and tdTomato fluorescence in each neuron was averaged for the five time points. tdTomato fluorescence was used to normalize GCaMP6s fluorescence to control for movement artifacts or changes in transgene expression. For *Tg(hs:tgfa)* experiments, these animals and their WT siblings were either not heat-shocked or heat-shocked for 30 min at 37°C and then embedded in agarose 30 min later. Animals were imaged 2 hours after heat shock. For gefitinib experiments, animals were head-fixed in agarose and imaged, then treated with either DMSO or 10  $\mu\text{M}$  gefitinib in E3 medium, and then imaged again 30 min later. A rolling ball algorithm (ImageJ) was used for background correction. ROIs were drawn over each cell in the red channel, and the average pixel intensity of the green channel (GCaMP6s) was divided by the average pixel intensity of the red channel (tdTomato). NPVF soma with saturated fluorescence were excluded from quantification. Images were quantified blind to treatment or condition.

### Human study population and phenotypes

This study was conducted on participants of European ancestry from the U.K. Biobank study of more than 500,000 people aged 40 to 69 living in the United Kingdom (24). We identified two genome-wide significant signals that localize within genes broadly implicated in EGFR/MAPK signaling using summary statistics of GWAS available at the Sleep Disorder Knowledge Portal: <http://sleepdisordergenetics.org/> (4–8). Self-reported health questionnaire, anthropometric assessments, and demographic information were collected at baseline. To assess chronotype, subjects were asked, “Do you consider yourself to be...” with response options “Definitely a ‘morning’ person,” “More a ‘morning’ than ‘evening’ person,” “More an ‘evening’ than a ‘morning’ person,” “Definitely an ‘evening’ person,” “Do not know,” and “Prefer not to answer,” which were coded as 2, 1, –1, –2, 0, and missing, respectively. To assess difficulty waking up, participants were asked, “On an average day, how easy do you find getting up in the morning?” with responses “Not at all easy,” “Not very easy,” “Fairly easy,” “Very easy,” “Do not know,” and “Prefer not to say.” To assess self-reported sleep duration, participants were asked, “About how many hours sleep do you get in every 24 hours? (please include naps),” with responses in hour increments. To assess frequent insomnia

symptoms, subjects were asked, “Do you have trouble falling asleep at night or do you wake up in the middle of the night?” with responses “never/rarely,” “sometimes,” “usually,” and “prefer not to answer.” Subjects who responded “Prefer not to answer” ( $n = 637$ ) were set to missing. Insomnia symptoms were dichotomized into controls (“never/rarely”) and cases with frequent insomnia symptoms (“usually”), with those reporting “sometimes” excluded. To assess daytime napping, subjects were asked, “Do you have a nap during the day?” with response options “Never/rarely,” “Sometimes,” “Usually,” and “Prefer not to answer.” Self-reported EDS was ascertained using the question “How likely are you to dose off or fall asleep during the daytime when you don’t mean to? (e.g., when working, reading, or driving)” with response options “Never/rarely,” “Sometimes,” “Often,” “All of the time,” “Do not know,” and “Prefer not to answer.” Subjects reporting “Do not know” and “Prefer not to answer” were set to missing. Other responses were coded continuously as 1 to 4 corresponding to the severity of EDS. Subjective overall health was reported in response to the question “In general, how would you rate your overall health?” Those who self-reported “Excellent” or “Good” overall health were classified as “healthy,” and those who self-reported “Fair” or “Poor” overall health were classified as “unhealthy.”

### Genetic association analysis

Genome-wide genotyping was performed by the U.K. Biobank, and genotyping, quality control, and imputation procedures have been previously described in detail (66). We further clustered subjects into four ancestry clusters using *K*-means clustering on the principal components, identifying 453,964 subjects of European ancestry. Genetic association analysis was performed, as previously described (7), in related subjects of European ancestry using BOLT (67) linear mixed models and an additive genetic model adjusted for age, sex, 10 principal components, genotyping array, and genetic correlation matrix with a maximum per-SNP missingness of 10% and per-sample missingness of 40%. We tested for an association between SNPs in the genomic intervals encompassing *ERBB4* and *KSR2*. Genomic intervals were defined using recombination hotspots upstream and downstream of the gene of interest and correspond to the National Center for Biotechnology Information Build 37/hg19 positions 211.94 to 213.49 Mb for *ERBB4* and 117.81 to 118.47 Mb for *KSR2*. Linear/logistic regression was performed adjusting for age, sex, 10 principal components of ancestry, and genotyping array. Additional adjustment for body mass index was performed in the regression model for difficulty waking up and EDS. Association testing for daytime napping excluded those with self-reported poor overall health (“unhealthy”;  $n = 60,712$ ). Reported SNPs were significant at the stringent genome-wide significant level used in GWAS ( $P < 5 \times 10^{-8}$ ). Locus zoom plots were generated using LocusZoom using the HG19 1000 genomes European genetic panel for genetic linkage data (68).

### *ERBB4* and *KSR2* SNP eQTL analysis in Genome-Tissue Expression

The Genome-Tissue Expression (GTEx) Project (69) data portal (accessed 1 July 2019) was used to investigate genetic effects of the genome-wide significant association signals at *ERBB4* and *KSR2* on gene expression across 48 tissues (data source: GTEx data version V7;  $n = 620$  individuals). GTEx version V7 eQTL analysis (described in the data portal) used gene expression measured using RNA se-

quencing (Illumina TruSeq) with inclusion of samples with an RNA integrity number (as measured by Agilent Bioanalyzer) of 6.0 or higher, and quality control measures included genes  $>0.1$  transcripts per kilobase million in at least 20% of samples and  $\geq 6$  reads in at least 20% of samples. Expression values were quantile-normalized between samples, and for each gene, expression values were normalized across samples using an inverse normal transform. SNPs were genotyped in GTEx using whole-genome sequencing from blood samples. SNPs with a minor allele frequency  $\geq 1\%$  were used. Association analyses were linear regression analyses that were adjusted for sex, genotyping platform, and the top three genotyping-based principal components to adjust for potential population stratification by race/ethnicity and sample size-dependent complex nongenetic PEER (probabilistic estimation of expression residuals) factors (70). The normalized effect size of the eQTLs was defined as the slope of the linear regression based on the human genome reference GRCh37/hg19. The normalized effect sizes were computed in a normalized space where magnitude had no direct biological interpretation. A false discovery rate  $< 0.05$  was used to correct for multiple hypothesis testing.

### Statistical analysis

Behavioral data were processed using MATLAB (MathWorks), graphs were generated using Excel (Microsoft), and statistical analyses were performed using Prism 6 (GraphPad). The number of humans, animals, and statistical tests used are stated in each figure or figure legend. Line graphs in Figs. 1 to 4 and figs. S1 and S3 to S8 were generated from raw data and averaged over 1-hour bins in 10-min intervals to show underlying behavioral trends. Bar graphs show mean  $\pm$  SEM. Box plots extend from the 25th to the 75th percentile, with the median marked by a horizontal line in the box, and whiskers extend from the 10th to the 90th percentile of data points. In all zebrafish statistical tests, the significance threshold was set to  $P < 0.05$ , and  $P$  values were adjusted for multiple comparisons where appropriate. Parametric analyses were applied because the data followed an approximately normal distribution in nearly all cases. Unpaired two-tailed Student’s *t* test was performed using Excel. One-way and two-way ANOVA and post hoc tests to correct for multiple comparisons were performed using Prism. The Holm-Sidak post hoc test was used to correct for multiple comparisons and to allow pairwise comparison of means for all samples. Asterisks in figures denote statistics for pairwise or multiple comparisons as indicated. The GCaMP6s/tdTomato fluorescence data and sleep rebound data were not normally distributed, so a Wilcoxon rank-sum non-parametric test was used to test for significant differences.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/11/eaax4249/DC1>

Fig. S1. Effects of gain and loss of EGFR signaling on sleep architecture.

Fig. S2. Amino acid alignment of human and zebrafish TGF $\alpha$ , EGF, and EGFR.

Fig. S3. Gefitinib does not enhance *egfra*<sup>-/-</sup> phenotype and effects of EGFR inhibitors on sleep architecture.

Fig. S4. EGFR signaling is not required for behavioral circadian rhythms.

Fig. S5. Validation of an SD assay, and EGFR signaling is required for normal homeostatic regulation of sleep.

Fig. S6. Inhibition of MAPK/ERK signaling suppresses TGF $\alpha$  overexpression-induced sleep.

Fig. S7. EGFR signaling regulates *npvf* expression, and TGF $\alpha$  overexpression-induced sleep is suppressed in *npvf* mutant animals.

Fig. S8. Association of *ERBB4* sleepiness allele with increased *ERBB4* expression in humans and pharmacological inhibition of *KSR2* or *ERBB4* decrease sleep in zebrafish.

Table S1. Variants at *ERBB4* and *KSR2* associate with self-reported measures of sleep quality and quantity in U.K. Biobank subjects.

Table S2. Descriptive characteristics of U.K. Biobank subjects of European ancestry used for sleep trait analysis.

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- H. R. Colten, Institute of Medicine, Board on Health Sciences Policy, Committee on Sleep Medicine and Research, in *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*, B. M. Altevogt, H. R. Colten, Eds. (National Academies Press, 2006).
- S. E. Jones, J. Tyrrell, A. R. Wood, R. N. Beaumont, K. S. Ruth, M. A. Tuke, H. Yaghoorakar, Y. Hu, M. Teder-Laving, C. Hayward, T. Roenneberg, J. F. Wilson, F. del Greco, A. A. Hicks, C. Shin, C. H. Yun, S. K. Lee, A. Metspalu, E. M. Byrne, P. R. Gehrman, H. Tiemeier, K. V. Allebrandt, R. M. Freathy, A. Murray, D. A. Hinds, T. M. Frayling, M. N. Weedon, Genome-wide association analyses in 128,266 individuals identifies new morningness and sleep duration loci. *PLoS Genet.* **12**, e1006125 (2016).
- Y. Hu, A. Shmygelska, D. Tran, N. E. Eriksson, J. Y. Tung, D. A. Hinds, GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat. Commun.* **7**, 10448 (2016).
- 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. Spiegelhalter, 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).
- H. S. Dashti, S. E. Jones, A. R. Wood, J. M. Lane, V. T. van Hees, H. Wang, J. A. Rhodes, Y. Song, K. Patel, S. G. Anderson, R. N. Beaumont, D. A. Bechtold, J. Bowden, B. E. Cade, M. Garaulet, S. D. Kyle, M. A. Little, A. S. Loudon, A. I. Luik, F. A. J. L. Scheer, K. Spiegelhalter, J. Tyrrell, D. J. Gottlieb, H. Tiemeier, D. W. Ray, S. M. Purcell, T. M. Frayling, S. Redline, D. A. Lawlor, M. K. Rutter, M. N. Weedon, R. Saxena, Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat. Commun.* **10**, 1100 (2019).
- 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. Kocovska, J. Tyrrell, R. N. Beaumont, M. Hillsdon, K. S. Ruth, M. A. Tuke, H. Yaghoorakar, 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).
- J. M. Lane, S. E. Jones, H. S. Dashti, A. R. Wood, K. G. Aragom, V. T. van Hees, L. B. Strand, B. S. Winsvold, H. Wang, J. Bowden, Y. Song, K. Patel, S. G. Anderson, R. N. Beaumont, D. A. Bechtold, B. E. Cade, M. Haas, S. Kathiresan, M. A. Little, A. I. Luik, A. S. Loudon, S. Purcell, R. C. Richmond, F. A. J. L. Scheer, B. Schormair, J. Tyrrell, J. W. Winkelmann, J. Winkelmann; HUNT All In Sleep, K. Hveem, C. Zhao, J. B. Nielsen, C. J. Willer, S. Redline, K. Spiegelhalter, S. D. Kyle, D. W. Ray, J. A. Zwart, B. Brumpton, T. M. Frayling, D. A. Lawlor, M. K. Rutter, M. N. Weedon, R. Saxena, Biological and clinical insights from genetics of insomnia symptoms. *Nat. Genet.* **51**, 387–393 (2019).
- 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. Yaghoorakar, 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).
- W. J. Joiner, Unraveling the evolutionary determinants of sleep. *Curr. Biol.* **26**, R1073–R1087 (2016).
- G. Oikonomou, D. A. Prober, Attacking sleep from a new angle: Contributions from zebrafish. *Curr. Opin. Neurobiol.* **44**, 80–88 (2017).
- H. Bringmann, Sleep-active neurons: Conserved motors of sleep. *Genetics* **208**, 1279–1289 (2018).
- C. Van Buskirk, P. W. Sternberg, Epidermal growth factor signaling induces behavioral quiescence in *Caenorhabditis elegans*. *Nat. Neurosci.* **10**, 1300–1307 (2007).
- 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).
- A. Kramer, F.-C. Yang, P. Snodgrass, X. Li, T. E. Scammell, F. C. Davis, C. J. Weitz, Regulation of daily locomotor activity and sleep by hypothalamic EGF receptor signaling. *Science* **294**, 2511–2515 (2001).
- J. Gilbert, F. C. Davis, Behavioral effects of systemic transforming growth factor- $\alpha$  in Syrian hamsters. *Behav. Brain Res.* **198**, 440–448 (2009).
- J. M. Donlea, N. Raman, P. J. Shaw, Use-dependent plasticity in clock neurons regulates sleep need in *Drosophila*. *Science* **324**, 105–108 (2009).
- S. T. Harbison, L. J. McCoy, T. F. Mackay, Genome-wide association study of sleep in *Drosophila melanogaster*. *BMC Genomics* **14**, 281 (2013).
- S. T. Harbison, Y. L. Serrano Negron, N. F. Hansen, A. S. Lobell, Selection for long and short sleep duration in *Drosophila melanogaster* reveals the complex genetic network underlying natural variation in sleep. *PLoS Genet.* **13**, e1007098 (2017).
- A. J. Hill, R. Mansfield, J. M. Lopez, D. M. Raizen, C. Van Buskirk, Cellular stress induces a protective sleep-like state in *C. elegans*. *Curr. Biol.* **24**, 2399–2405 (2014).
- T. Kushikata, J. Fang, Z. Chen, Y. Wang, J. M. Krueger, Epidermal growth factor enhances spontaneous sleep in rabbits. *Am. J. Physiol.* **275**, R509–R514 (1998).
- R. D. Nath, E. S. Chow, H. Wang, E. M. Schwarz, P. W. Sternberg, *C. elegans* stress-induced sleep emerges from the collective action of multiple neuropeptides. *Curr. Biol.* **26**, 2446–2455 (2016).
- M. D. Nelson, K. H. Lee, M. A. Churgin, A. J. Hill, C. van Buskirk, C. Fang-Yen, D. M. Raizen, FMRamide-like FLP-13 neuropeptides promote quiescence following heat stress in *Caenorhabditis elegans*. *Curr. Biol.* **24**, 2406–2410 (2014).
- I. V. Zhdanova, S. Y. Wang, O. U. Leclair, N. P. Danilova, Melatonin promotes sleep-like state in zebrafish. *Brain Res.* **903**, 263–268 (2001).
- C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T. Peakman, R. Collins, UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
- J. Salvatierra, D. A. Lee, C. Zibetti, M. Duran-Moreno, S. Yoo, E. A. Newman, H. Wang, J. L. Bedont, J. de Melo, A. L. Miranda-Angulo, S. Gil-Perotin, J. M. Garcia-Verdugo, S. Blackshaw, The LIM homeodomain factor Lhx2 is required for hypothalamic tanyocyte specification and differentiation. *J. Neurosci.* **34**, 16809–16820 (2014).
- P. Snodgrass-Belt, J. L. Gilbert, F. C. Davis, Central administration of transforming growth factor- $\alpha$  and neuregulin-1 suppress active behaviors and cause weight loss in hamsters. *Brain Res.* **1038**, 171–182 (2005).
- Y. J. Ma, M. P. Junier, M. E. Costa, S. R. Ojeda, Transforming growth factor- $\alpha$  gene expression in the hypothalamus is developmentally regulated and linked to sexual maturation. *Neuron* **9**, 657–670 (1992).
- 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).
- G. B. Mann, K. J. Fowler, A. P. Gabriel, E. C. Nice, R. L. Williams, A. R. Dunn, Mice with a null mutation of the TGF  $\alpha$  gene have abnormal skin architecture, wavy hair, and curly whiskers and often develop corneal inflammation. *Cell* **73**, 249–261 (1993).
- S. Chen, G. Oikonomou, C. N. Chiu, B. J. Niles, J. Liu, D. A. Lee, I. Antoshechkin, D. A. Prober, A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. *Nucleic Acids Res.* **41**, 2769–2778 (2013).
- W. Y. Hwang, Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J. R. Yeh, J. K. Joung, Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
- M. Sibilia, E. F. Wagner, Strain-dependent epithelial defects in mice lacking the EGF receptor. *Science* **269**, 234–238 (1995).
- J. A. Laisney, I. Braasch, R. B. Walter, S. Meierjohann, M. Schartl, Lineage-specific co-evolution of the Egf receptor/ligand signaling system. *BMC Evol. Biol.* **10**, 27 (2010).
- S. R. Hubbard, W. T. Miller, Receptor tyrosine kinases: Mechanisms of activation and signaling. *Curr. Opin. Cell Biol.* **19**, 117–123 (2007).
- D. A. Lee, A. Andreev, T. V. Truong, A. Chen, A. J. Hill, G. Oikonomou, U. Pham, Y. K. Hong, S. Tran, L. Glass, V. Sapin, J. Engle, S. E. Fraser, D. A. Prober, Genetic and neuronal regulation of sleep by neuropeptide VF. *eLife* **6**, e25727 (2017).
- S. Chen, C. N. Chiu, K. L. McArthur, J. R. Fetcho, D. A. Prober, TRP channel mediated neuronal activation and ablation in freely behaving zebrafish. *Nat. Methods* **13**, 147–150 (2016).
- W. M. Vanderheyden, J. R. Gerstner, A. Tanenhaus, J. C. Yin, P. J. Shaw, ERK phosphorylation regulates sleep and plasticity in *Drosophila*. *PLoS ONE* **8**, e81554 (2013).
- C. C. Hong, Q. P. Peterson, J. Y. Hong, R. T. Peterson, Artery/vein specification is governed by opposing phosphatidylinositol-3 kinase and MAP kinase/ERK signaling. *Curr. Biol.* **16**, 1366–1372 (2006).
- C. Mikhail, A. Vaucher, S. Jimenez, M. Tafti, ERK signaling pathway regulates sleep duration through activity-induced gene expression during wakefulness. *Sci. Signal.* **10**, eaa19219 (2017).
- O. Lenz, J. Xiong, M. D. Nelson, D. M. Raizen, J. A. Williams, FMRamide signaling promotes stress-induced sleep in *Drosophila*. *Brain Behav. Immun.* **47**, 141–148 (2015).
- F. Simonin, M. Schmitt, J. P. Laulin, E. Laboureyras, J. H. Jhamandas, D. MacTavish, A. Matifas, C. Mollereau, P. Laurent, M. Parmentier, B. L. Kieffer, J. J. Bourguignon,

- G. Simonnet, RF9, a potent and selective neuropeptide FF receptor antagonist, prevents opioid-induced tolerance associated with hyperalgesia. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 466–471 (2006).
42. R. Madelaine, M. Lovett-Barron, C. Halluin, A. S. Andalman, J. Liang, G. M. Skariah, L. C. Leung, V. M. Burns, P. Mourrain, The hypothalamic NPVF circuit modulates ventral raphe activity during nociception. *Sci. Rep.* **7**, 41528 (2017).
  43. C. Singh, J. Rihel, D. A. Prober, Neuropeptide Y regulates sleep by modulating noradrenergic signaling. *Curr. Biol.* **27**, 3796–3811.e5 (2017).
  44. R. W. Daniels, A. J. Rossano, G. T. Macleod, B. Ganetzky, Expression of multiple transgenes from a single construct using viral 2A peptides in *Drosophila*. *PLoS ONE* **9**, e100637 (2014).
  45. B. D. Cohen, J. M. Green, L. Foy, H. P. Fell, HER4-mediated biological and biochemical properties in NIH 3T3 cells. Evidence for HER1-HER4 heterodimers. *J. Biol. Chem.* **271**, 4813–4818 (1996).
  46. N. S. Dhawan, A. P. Scopton, A. C. Dar, Small molecule stabilization of the KSR inactive state antagonizes oncogenic Ras signalling. *Nature* **537**, 112–116 (2016).
  47. D. L. Costanzo-Garvey, P. T. Pfluger, M. K. Dougherty, J. L. Stock, M. Boehm, O. Chaika, M. R. Fernandez, K. Fisher, R. L. Kortum, E. G. Hong, J. Y. Jun, H. J. Ko, A. Schreiner, D. J. Volle, T. Treece, A. L. Swift, M. Winer, D. Chen, M. Wu, L. R. Leon, A. S. Shaw, J. McNeish, J. K. Kim, D. K. Morrison, M. H. Tschöp, R. E. Lewis, KSR2 is an essential regulator of AMP kinase, energy expenditure, and insulin sensitivity, and insulin sensitivity. *Cell Metab.* **10**, 366–378 (2009).
  48. M. C. Wehr, W. Hinrichs, M. M. Brzózka, T. Unterbarnscheidt, A. Herholt, J. P. Wintgens, S. Papiol, M. C. Soto-Bernardini, M. Kravchenko, M. Zhang, K. A. Nave, S. P. Wichert, P. Falkai, W. Zhang, M. H. Schwab, M. J. Rossner, Spironolactone is an antagonist of NRG1-ERBB4 signaling and schizophrenia-relevant endophenotypes in mice. *EMBO Mol. Med.* **9**, 1448–1462 (2017).
  49. S. M. Lee, S. Tole, E. Grove, A. P. McMahon, A local Wnt-3a signal is required for development of the mammalian hippocampus. *Development* **127**, 457–467 (2000).
  50. F. Weber, Y. Dan, Circuit-based interrogation of sleep control. *Nature* **538**, 51–59 (2016).
  51. Y. J. Ma, D. F. Hill, M. P. Junier, M. E. Costa, S. E. Felder, S. R. Ojeda, Expression of epidermal growth factor receptor changes in the hypothalamus during the onset of female puberty. *Mol. Cell. Neurosci.* **5**, 246–262 (1994).
  52. D. A. Lee, S. Blackshaw, Feed your head: Neurodevelopmental control of feeding and metabolism. *Annu. Rev. Physiol.* **76**, 197–223 (2014).
  53. X. Li, N. Sankrithi, F. C. Davis, Transforming growth factor- $\alpha$  is expressed in astrocytes of the suprachiasmatic nucleus in hamster: Role of glial cells in circadian clocks. *Neuroreport* **13**, 2143–2147 (2002).
  54. C. Y. Feng, L. M. Wiggins, C. S. von Bartheld, The locus ceruleus responds to signaling molecules obtained from the CSF by transfer through tanycytes. *J. Neurosci.* **31**, 9147–9158 (2011).
  55. J. M. Krueger, D. M. Rector, L. Churchill, Sleep and cytokines. *Sleep Med. Clin.* **2**, 161–169 (2007).
  56. T. Rich, P. F. Innominato, J. Boerner, M. C. Mormont, S. Iacobelli, B. Baron, C. Jasmin, F. Lévi, Elevated serum cytokines correlated with altered behavior, serum cortisol rhythm, and dampened 24-hour rest-activity patterns in patients with metastatic colorectal cancer. *Clin. Cancer Res.* **11**, 1757–1764 (2005).
  57. K. A. Rauen, S. M. Huson, E. Burkitt-Wright, D. G. Evans, S. Farschtschi, R. E. Ferner, D. H. Gutmann, C. O. Hanemann, B. Kerr, E. Legius, L. F. Parada, M. Patton, J. Peltonen, N. Ratner, V. M. Riccardi, T. van der Vaart, M. Vikkula, D. H. Viskochil, M. Zenker, M. Upadhyaya, Recent developments in neurofibromatosis and RASopathies: Management, diagnosis and current and future therapeutic avenues. *Am. J. Med. Genet. A* **167A**, 1–10 (2015).
  58. M. Westerfield, *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Brachydanio rerio)*, M. Westerfield, O. R. Eugene, Eds. (University of Oregon Press, 1993).
  59. M. C. Halloran, M. Sato-Maeda, J. T. Warren, F. Su, Z. Lele, P. H. Krone, J. Y. Kuwada, W. Shoji, Laser-induced gene expression in specific cells of transgenic zebrafish. *Development* **127**, 1953–1960 (2000).
  60. V. Thernes, C. Grabher, F. Ristoratore, F. Bourrat, A. Choulika, J. Wittbrodt, J. S. Joly, I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **118**, 91–98 (2002).
  61. T. W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, D. S. Kim, Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
  62. E. Provost, J. Rhee, S. D. Leach, Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos. *Genesis* **45**, 625–629 (2007).
  63. I. G. Woods, D. Schoppik, V. J. Shi, S. Zimmerman, H. A. Coleman, J. Greenwood, E. R. Soucy, A. F. Schier, Neuropeptidergic signaling partitions arousal behaviors in zebrafish. *J. Neurosci.* **34**, 3142–3160 (2014).
  64. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup> $\Delta\Delta$ CT method. *Methods* **25**, 402–408 (2001).
  65. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
  66. C. Bycroft, C. Freeman, D. Petkova, G. Band, L. T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. O'Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. McVean, S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
  67. P. R. Loh, G. Tucker, B. K. Bulik-Sullivan, B. J. Vilhjálmsón, H. K. Finucane, R. M. Salem, D. I. Chasman, P. M. Ridker, B. M. Neale, B. Berger, N. Patterson, A. L. Price, Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
  68. R. J. Pruim, R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Glied, M. Boehnke, G. R. Abecasis, C. J. Willer, LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
  69. GTEx Consortium, The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
  70. O. Stegle, L. Parts, R. Durbin, J. Winn, A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS Comput. Biol.* **6**, e1000770 (2010).
- Acknowledgments:** We thank members of the Prober Lab, P. Sternberg, R. Nath, H. Wang, C. Oikonomou, and J. Bedont for feedback. We also thank H. S. Dashti, C. Singh, D. Chilin, V. Sapin, T. Cammidge, H. Hurley, and C. Cook for technical assistance. **Funding:** This work was supported by grants from the National Institutes of Health (NIH) (D.A.L.: K99NS097683 and F32NS084769; G.O.: F32NS082010; D.A.P.: R01NS070911 and R01NS101158; J.M.L.: T32HL007567; R.S.: R01DK107859, R01HL113338, R01DK102696, and R01DK105072) and the Mallinckrodt (to D.A.P.), Rita Allen (to D.A.P.), and Brain and Behavior Research (to D.A.P. and D.A.L.) foundations. The GTEx Project was supported by the Common Fund of the Office of the Director of the NIH and by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. **Author contributions:** D.A.L., J.L., and D.A.P. conceived the project and designed the experiments. D.A.L. performed most of the experiments and analyzed the data, with assistance from J.L., Y.H., A.J.H., S.L.H., G.O., U.P., J.E. and D.A.P. J.M.L., H.W., and R.S. performed the sleep genetic association analysis. D.A.L. and D.A.P. wrote the manuscript with input from J.M.L. and R.S. **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. Additional data related to this paper may be requested from the authors. Custom PERL and MATLAB code used for zebrafish behavioral analysis is available upon request and posted online (35). Human sleep genetic association analysis was conducted using the U.K. Biobank Resource under application number 6818.
- Submitted 21 March 2019  
Accepted 17 September 2019  
Published 13 November 2019  
10.1126/sciadv.aax4249
- Citation:** 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).

## Evolutionarily conserved regulation of sleep by epidermal growth factor receptor signaling

Daniel A. Lee, Justin Liu, Young Hong, Jacqueline M. Lane, Andrew J. Hill, Sarah L. Hou, Heming Wang, Grigorios Oikonomou, Uyen Pham, Jae Engle, Richa Saxena and David A. Prober

*Sci Adv* 5 (11), eaax4249.  
DOI: 10.1126/sciadv.aax4249

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/5/11/eaax4249>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2019/11/08/5.11.eaax4249.DC1>

### REFERENCES

This article cites 68 articles, 15 of which you can access for free  
<http://advances.sciencemag.org/content/5/11/eaax4249#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2019 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).