

# Molecular codes and in vitro generation of hypocretin and melanin concentrating hormone neurons

Ali Seifinejad<sup>a,b,1</sup>, Sha Li<sup>a</sup>, Cyril Mikhail<sup>b</sup>, Anne Vassalli<sup>a</sup>, Sylvain Pradervand<sup>c</sup>, Yoan Arribat<sup>a</sup>, Hassan Pezeshgi Modarres<sup>d</sup>, Bridget Allen<sup>e</sup>, Rosalind M. John<sup>e</sup>, Francesca Amati<sup>a</sup>, and Mehdi Tafti<sup>a,1</sup>

<sup>a</sup>Department of Physiology, Faculty of Biology and Medicine, University of Lausanne, 1005 Lausanne, Switzerland; <sup>b</sup>Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland; <sup>c</sup>Genomic Technologies Facility, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland; <sup>d</sup>Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; and <sup>e</sup>Biomedicine Division, Cardiff School of Biosciences, Cardiff University, CF10 3AX Cardiff, United Kingdom

Edited by Joseph S. Takahashi, University of Texas Southwestern Medical Center, Dallas, TX, and approved July 10, 2019 (received for review February 8, 2019)

**Hypocretin/orexin (HCRT) and melanin concentrating hormone (MCH) neuropeptides are exclusively produced by the lateral hypothalamus and play important roles in sleep, metabolism, reward, and motivation. Loss of HCRT (ligands or receptors) causes the sleep disorder narcolepsy with cataplexy in humans and in animal models. How these neuropeptides are produced and involved in diverse functions remain unknown. Here, we developed methods to sort and purify HCRT and MCH neurons from the mouse late embryonic hypothalamus. RNA sequencing revealed key factors of fate determination for HCRT (*Peg3*, *Ahr1*, *Six6*, *Nr2f2*, and *Prrx1*) and MCH (*Lmx1*, *Gbx2*, and *Peg3*) neurons. Loss of *Peg3* in mice significantly reduces HCRT and MCH cell numbers, while knock-down of a *Peg3* ortholog in zebrafish completely abolishes their expression, resulting in a 2-fold increase in sleep amount. We also found that loss of HCRT neurons in *Hart-ataxin-3* mice results in a specific 50% decrease in another orexigenic neuropeptide, QRFP, that might explain the metabolic syndrome in narcolepsy. The transcriptome results were used to develop protocols for the production of HCRT and MCH neurons from induced pluripotent stem cells and ascorbic acid was found necessary for HCRT and BMP7 for MCH cell differentiation. Our results provide a platform to understand the development and expression of HCRT and MCH and their multiple functions in health and disease.**

HCRT/OREXIN | MCH | transcription factor | iPSC | *Peg3*

The lateral and posterior hypothalamus play a central role in the regulation of food intake (1), vigilance states (2, 3), reward, and motivation (4). Among major cell types of this region are hypocretin/orexin- (HCRT) and melanin concentrating hormone- (MCH) producing cells. These neuropeptides are not colocalized but their cells are spatially intermingled (5).

HCRT and MCH neurons send projections throughout the central nervous system (CNS) (6, 7) and orchestrate their effects via their respective receptors; HCRT receptor 1 and 2 and MCH receptor 1 and 2 (8, 9). HCRT neurons project to all wake-promoting nuclei of the brain located in the brainstem, basal forebrain, hypothalamus, and thalamus (7). As physiological integrators they act as intermediates between different regulatory pathways, such as the autonomic, endocrine, feeding, reward, and stress systems (10, 11). MCH neurons are also involved in multiple functions including sleep, especially rapid-eye movement sleep (12), regulating energy balance (13), beat frequency of ependymal cilia, and ventricular volume (14), and are associated in controlling stress, and learning and memory (15, 16).

There is considerable evidence that sleep disorders are lined to alteration in HCRT signaling. Narcolepsy in dogs is linked to recessive loss-of-function mutations in HCRT receptor 2 (17). Patients with narcolepsy with cataplexy (also called narcolepsy type1 or NT1) have reduced or undetectable cerebrospinal fluid levels of HCRT-1 peptide and immunostaining of postmortem hypothalamus shows no or dramatically reduced HCRT cells, while MCH cells are intact (18). Hypocretin deficiency is accompanied by sudden loss of muscle tone (cataplexy) and daytime sleepiness, the 2 main features of narcolepsy with cataplexy. In addition, *Hcrt* gene deletion in mice results in a phenotype similar to human

narcolepsy (19). How narcolepsy patients lose the expression of the *Hcrt* gene or HCRT cells remains a mystery, but one hypothesis is an autoimmune-driven destruction of HCRT cells (20). The specific peptide expressed by HCRT cells and targeted by the immune attack is unknown, although autoreactive T cells to HCRT were recently reported (21) in narcolepsy with and without cataplexy. The important role of MCH in the pathophysiology of obesity, especially diet-induced obesity, is proven (9, 13, 22). How these multiple functions are accomplished by a small number of hypothalamic neuronal cell types needs careful investigation of their intrinsic properties.

Transcriptional maps for MCH cells were lacking until very recently (23) and attempts to obtain profiles for HCRT neurons relied on total RNA or mRNA of HCRT cells through transcript precipitation and comparing brains with and without HCRT (24–27), except in the recent single-cell transcriptome study (23). For the mRNAs obtained from HCRT neurons, some portion of transcripts is lost due to technical problems and because transcripts are not precipitated from purified cells. In the comparative study between HCRT-deficient and normal brain tissues, the differentially expressed genes (DEGs) can be masked by transcripts of other cell types. Finally, microarray technology was used to

## Significance

**Hypocretin (HCRT) and melanin concentrating hormone are brain neuropeptides involved in multiple functions, including sleep and metabolism. Loss of HCRT causes the sleep disorder narcolepsy. To understand how these neuropeptides are produced and contribute to diverse functions in health and disease, we purified their cells from mouse embryonic brains and established their molecular machinery. We discovered that partial removal of PEG3 (a transcription factor) in mice significantly reduces the number of HCRT and melanin concentrating hormone neurons, and its down-regulation in zebrafish completely abolishes their expression. We used our molecular data to produce these neurons in vitro from mouse fibroblasts, a technique that can be applied to cells from narcolepsy patients to generate an in vitro cell-based model.**

Author contributions: A.S., R.M.J., F.A., and M.T. designed research; A.S., S.L., C.M., Y.A., and B.A. performed research; A.S., S.L., A.V., S.P., Y.A., H.P.M., and M.T. analyzed data; and A.S., A.V., R.M.J., and M.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, <https://www.ncbi.nlm.nih.gov/geo> (accession nos. GSE126330 and GSE126345).

<sup>1</sup>To whom correspondence may be addressed. Email: Ali.Seifinejad@unil.ch or Mehdi.Tafti@unil.ch.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1902148116/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1902148116/-DCSupplemental).

compare the samples, which is less sensitive than RNA sequencing (RNA-seq), a technique that allows the detection of differentially spliced transcripts and transcripts with low abundance. In zebrafish, RNA-seq was used on purified HCRT cells and led to the identification of the *Kcnh4a* gene, which has a sleep modulatory role (28). Complementary to these approaches, emerging technologies, like single-cell sequencing of specific (MCH and HCRT) (29) or whole LH cell types (23), have substantially improved our understanding of molecular properties of these cell types in the adult mouse brain.

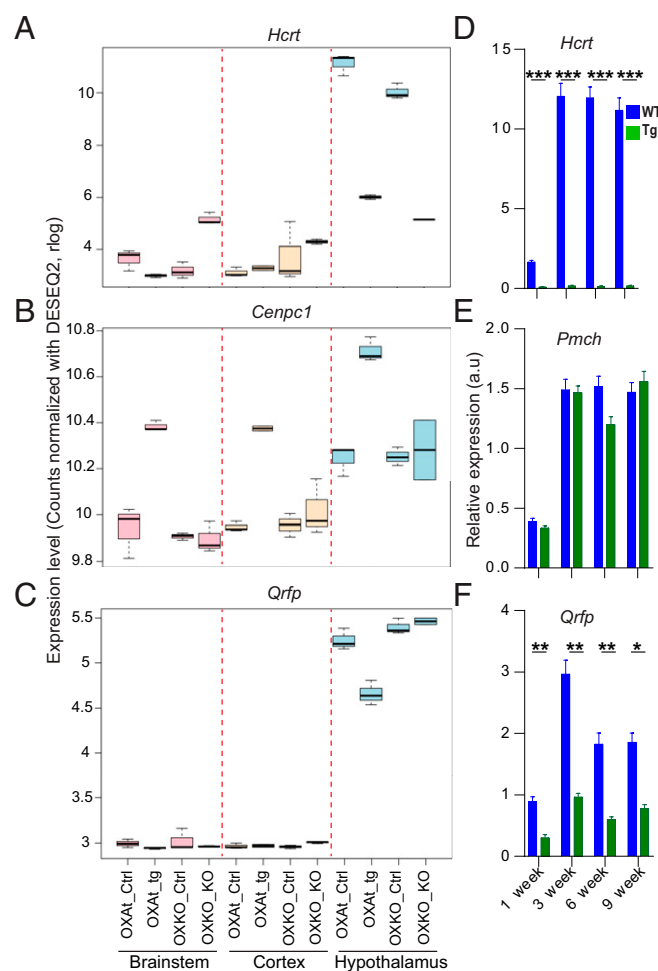
We hypothesized that a specific transcriptional machinery is responsible for the unique features of HCRT and MCH cells. To identify molecules that orchestrate the early specification of these cell types and to characterize molecules unique to them, we purified these cells from the mouse brain. We identified a set of transcription factors (TF) with important roles in the fate determination of both HCRT and MCH cells. One of these, PEG3, was identified with a substantial role in regulating vigilance states. The result of this high-throughput study and the discovery of these specific TFs led us to develop a protocol to derive these neuronal populations from mouse fibroblasts, that could potentially be applied to human cells to establish an in vitro model or for cell-replacement therapy in narcolepsy with cataplexy. Through this process, we have also identified molecules with potential importance in understanding the pathophysiology of narcolepsy.

## Result

**Transcriptome Analysis of Different Brain Areas in *Hcrt* Gene-, or *Hcrt* Cell-, Ablated Mice Reveals *Qrfp* as a Putative Marker of Narcolepsy with Cataplexy.** To identify cell type-specific regulators of HCRT cells, we used 2 mouse models of narcolepsy with cataplexy: *Hcrt* gene knockout (*Hcrt*<sup>ko/ko</sup>) mice (19) and *Hcrt-ataxin-3* mice in which HCRT cells are ablated by the expression of the *ataxin-3* gene (30). We hypothesized that comparing the brain transcriptome of these mice and their controls might identify transcriptional changes due both to a lack of the *Hcrt* gene and HCRT neurons altogether (as in narcolepsy). We dissected the hypothalamus, brainstem, and neocortex of adult mice belonging to 4 genotype groups: *Hcrt*<sup>ko/ko</sup> and *Hcrt*<sup>+/+</sup> littermates, and *Hcrt-Ataxin-3* hemizygous transgenic (*Tg*<sup>+</sup>) and WT littermates. RNA was purified and submitted to high-throughput RNA-seq.

As expected, the *Hcrt* transcript was found to be severely depleted in the hypothalamus of *Hcrt*<sup>ko/ko</sup> and *Hcrt-ataxin-3* transgenic mice (Fig. 1A). Nevertheless, very few DEGs except *Hcrt* were identified (adjusted  $P < 0.05$ ) (SI Appendix, Table S1). Analysis showed that in the cortex of *Hcrt*<sup>ko/ko</sup> mice *Kcnh4* and *Krt12* gene transcripts were down-regulated by more than 2-fold. *Kcnh4* encodes a voltage-gated potassium channel proposed to play a role in sleep regulation in zebrafish (28). In *Hcrt-Ataxin-3* mice, comparing all 3 brain areas with respective controls identified the *Cenpc1* gene with significant up-regulation in transgenic animals (Fig. 1B). *Cenpc1* encodes a protein that modulates DNA methylation and the histone code at centromeric regions (31).

One of the very few genes that were affected by HCRT cell loss in *Hcrt-Ataxin-3* but not *Hcrt* gene deletion was *Qrfp*, which attracted our attention due to its neuropeptidergic identity and anorexic function. *Hcrt-Ataxin-3 Tg*<sup>+</sup> mice exhibited a 50% decrease in *Qrfp* transcript levels in their hypothalamus relative to controls, while levels in *Hcrt*<sup>ko/ko</sup> mice did not differ from WT littermates (Fig. 1C). To confirm *Hcrt-ataxin-3 Tg*<sup>+</sup> mice hypothalamic *Qrfp* reduction, the hypothalamus was dissected at the age of 1, 3, 6, and 9 wk, and analyzed by qRT-PCR. A massive reduction of *Hcrt* gene expression during all ages was observed, while there were no changes in *Pmch* gene expression (Fig. 1D and E). More than 50% reduction was observed in *Qrfp* gene expression at all ages (1 wk: WT =  $0.9 \pm 0.07$  vs. tg =  $0.31 \pm 0.05$ ,  $P < 0.01$ ; 3 wk: WT =  $2.9 \pm 0.2$  vs. tg =  $0.9 \pm 0.06$ ,  $P < 0.01$ ; 6 wk: WT =  $1.8 \pm 0.1$  vs. tg =  $0.6 \pm 0.04$ ,  $P < 0.01$ ; 9 wk: WT =  $1.8 \pm 0.1$  vs. tg =  $0.7 \pm 0.06$ ,  $P < 0.05$ ;  $t$  test) (Fig. 1F). It was shown in zebrafish and mouse that QRFP and HCRT most probably comprise different subgroups of hypothalamic cells with no



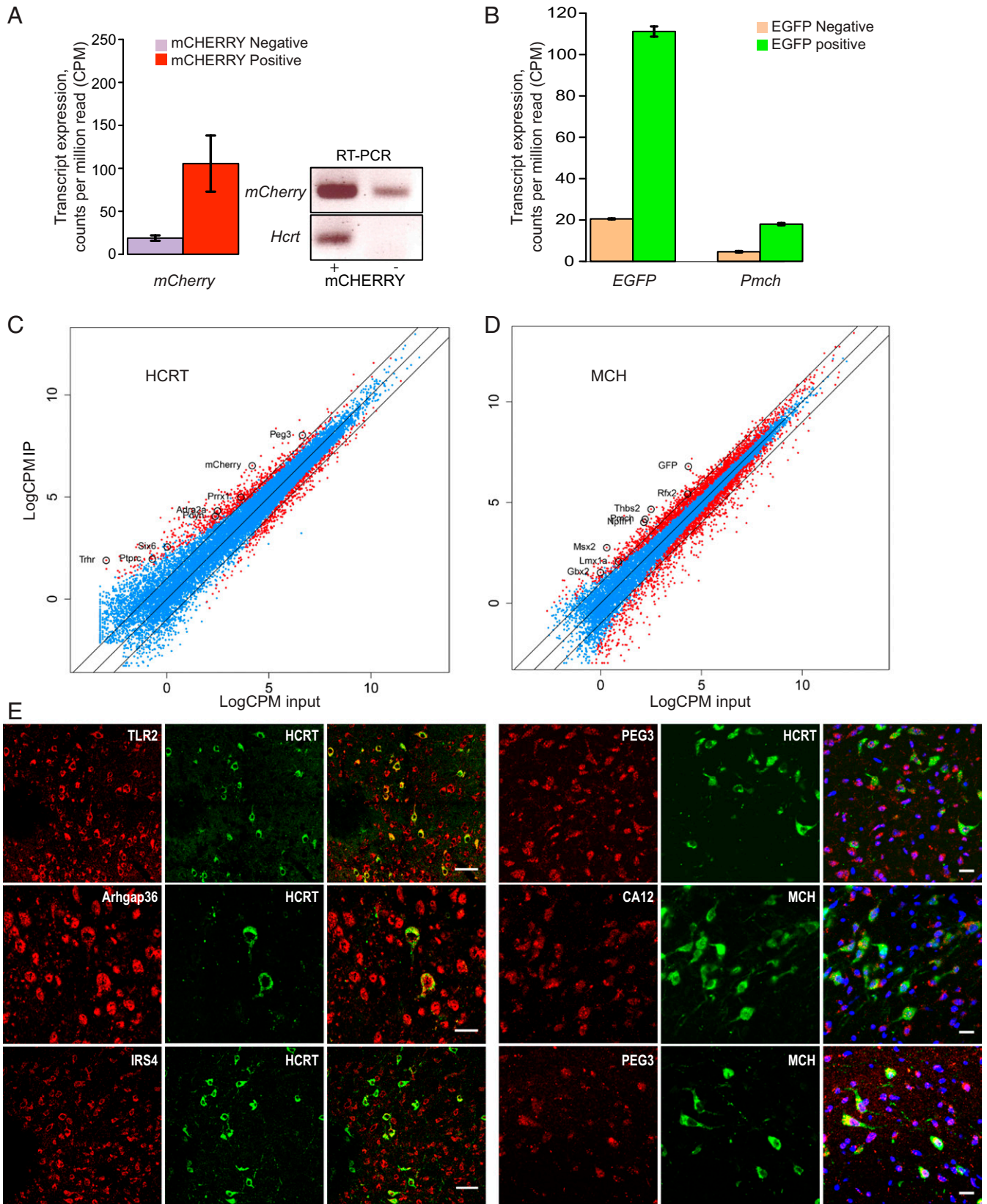
**Fig. 1.** Identification of *Cenpc1* and *Qrfp* as putative narcolepsy gene markers. (A–C) Differential RNA sequencing analysis between *Hcrt*<sup>ko/ko</sup> (OXKO), or *Hcrt-Ataxin-3 Tg* (OXAt) mice, and their respective control littermates (Ctrl), was performed in 3 brain regions (brainstem, cortex, and hypothalamus). Depicted are the differential expressions of *Hcrt*, *Cenpc1*, and *Qrfp* in each region. Note that *Qrfp* mRNA levels are about 50% reduced in the hypothalamus of *Hcrt-ataxin-3* transgenic mice relative to WT littermates. Read counts has been normalized with DESeq2 package. (D–F) Expression of *Hcrt*, *Pmch*, and *Qrfp* in the hypothalamus of *Hcrt-Ataxin-3* transgenic and WT littermates at 4 developmental time points (1, 3, 6, and 9 wk after birth). Shown are values (mean  $\pm$  SEM, arbitrary units) derived from qRT-PCR analysis ( $t$  test,  $n = 3$  mice; \* $P = 0.023$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

colocalization (25, 32). Nevertheless, single-cell RNA-seq analysis of the hypothalamus indicated colocalization of *Qrfp* and *Hcrt* and hierarchical clustering of hypothalamic neuronal subtypes defined by unique molecular fingerprints showed *Hcrt* and *Qrfp* among the closest neuronal subtypes, with the difference that QRFP neurons also highly express brain-derived neurotrophic factor (33). Our finding that *Qrfp* is lost together with HCRT neurons in *Hcrt*-ablated mice suggests that *Qrfp* might play a role not only in the control of sleep and metabolism but potentially in narcolepsy.

**Sorting of Intact Lateral Hypothalamic Cells.** Despite the discovery of *Qrfp*, the above experiment suffers several shortcomings as described in our introduction. To define the specific transcriptional repertoires of HCRT and MCH neurons, we set out to purify each cell type, and compare their transcriptomes to the rest of hypothalamus and to each other. Cell survival rate was too low during and after sorting of adult brains, while it was feasible for late embryonic stages (embryonic day [E] 18 to E19).

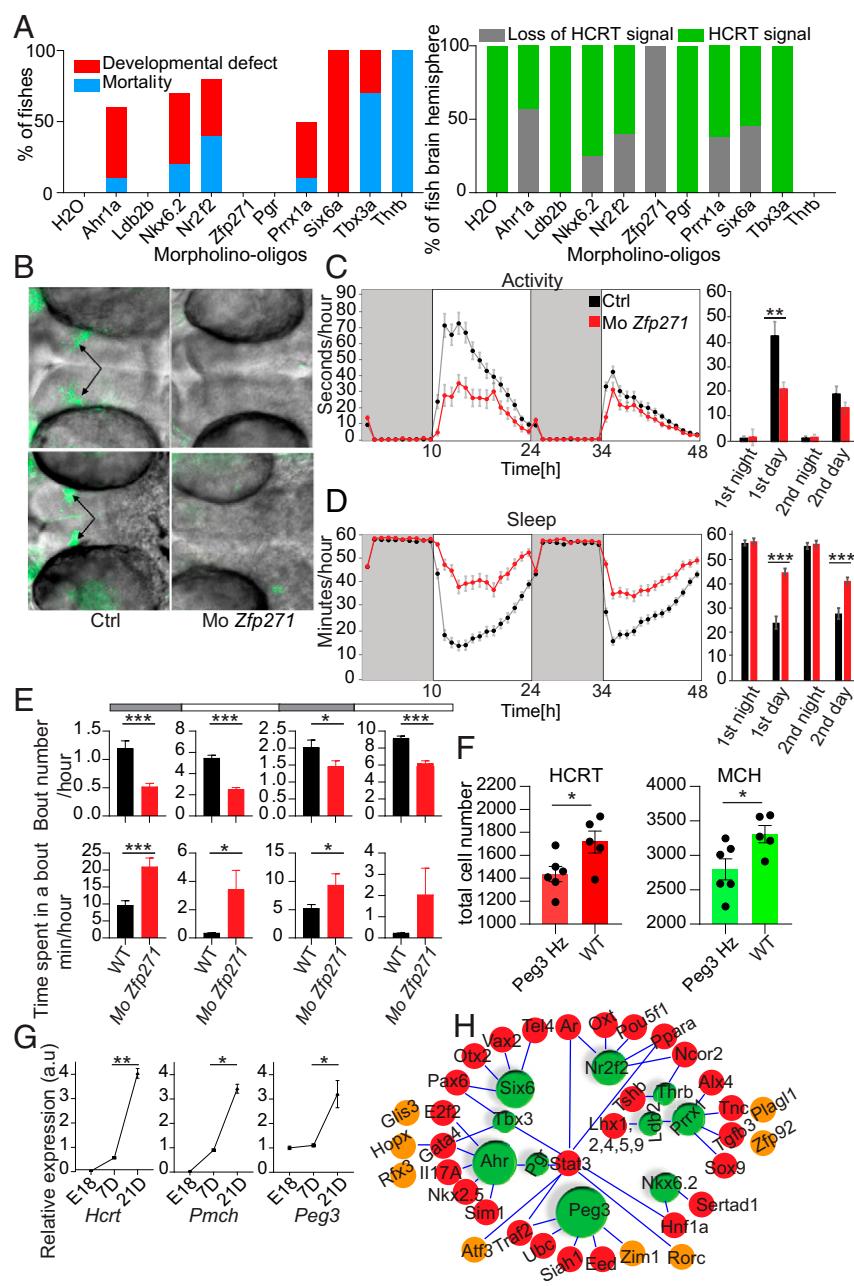






**Fig. 3.** Verification of the methods and enrichment for the genes expressed in HCRT and MCH cells. (A) HCRT cell sorting: cells sorted for mCherry fluorescence displayed enrichment for *Hcrt* transcript (Left). Due to very low expression of the *Hcrt* gene, we performed PCR to confirm the enrichment of HCRT cells and mCherry<sup>+</sup> cells were also positive for *Hcrt* expression (Right). The results are RNA-seq results and the transcript expression level is displayed as cpm reads. (B) MCH cell sorting: EGFP enriched cells showed enrichment for *EGFP* and *Pmch* transcripts. (C and D) Plotting of log counts per million reads (LogCPM) of immunoprecipitation against LogCPM input. Significant DEG genes are highlighted in red and nonsignificant ones in blue ( $P > 0.05$ ). Outer lines are 2-fold increase/decrease in expression. Some representative genes are marked. (E) Colocalization of candidate genes (red) with HCRT and MCH neuropeptides (green). (Scale bars: Left, 50  $\mu$ m; Right, 20  $\mu$ m.)





**Fig. 5.** Knockdown of critical TFs identified for *Hcrt* and *Pmch* gene expression and sleep behavior in fish. (A, Left) Mortality and developmental defect rate of 48-h postfertilization zebrafish larvae injected with 2.5 ng of candidate Mo (SI Appendix, Table S7). Data are presented as percentage of fish that exhibit the phenotype ( $10 < n < 20$ ). (Right) Visualization of HCRT signal in brain hemispheres of zebrafish embryos by immunostaining and confocal microscopy. The percentage of brain hemispheres positive or negative for HCRT signal for each candidate Mo is presented. (B) Representative image of loss of HCRT signal in Mo-Zfp271-injected zebrafish brains compared with controls. (10× confocal magnification.) (C) Hourly distribution of locomotor activity of Zfp271 morphants and their respective controls during 48 h (night [10 h] and day [14 h]). Dark periods are indicated in gray and light periods in white. As shown in the bar graph (Right) Mo-Zfp271-injected zebrafish exhibit about 50% reduced locomotor activity during the light period. \*\*P < 0.01; \*\*\*P < 0.001. (D) Hourly distribution of sleep during 48 h. Mo-Zfp271-injected zebrafish exhibit profound increase in sleep time compared with controls. (E) Sleep bout numbers and bout durations in Mo-Zfp271-injected fish compared with controls. During both light and dark periods, Zfp271 morphants exhibit reduced bout number. \*P < 0.05; \*\*\*P < 0.001; n = 48, mean ± SEM, t test. (F) HCRT and MCH neuron counts in Peg3 hemizygous KO mice (Hz; Peg3<sup>+/-</sup>) and their WT littermates. The entire hypothalamus was covered from anterior to posterior regions and all sections were counted. \*P = 0.04 for HCRT and \*P = 0.02 for MCH; total cell number ± SEM, Mann-Whitney U test; n = 5 to 6. (G) Quantitative PCR of brain *Hcrt*, *Pmch*, and *Peg3* expression during different time points. There is a parallel increase in the expression of *Hcrt* and *Pmch* along with *Peg3* up to 3 wk after birth. \*P = 0.019 for *Peg3* and P = 0.03 for *Pmch*; \*\*P < 0.01; n = 3, mean ± SEM; t test. (H) Potential gene network for identified TFs within HCRT cells. Their interactions are extracted from databases. The more effective TFs in zebrafish study (green) have bigger sizes. Identified TFs but not tested are in brown and potential targets in red.

duration both during the light (day 1: WT =  $0.3 \pm 0.02$  vs. Mo-Zfp271 =  $3.4 \pm 1.3$ ;  $P = 0.025$ , t test) and the dark period (night 1: WT =  $9.5 \pm 1.3$  vs. Mo-Zfp271 =  $20.9 \pm 2.5$ ,  $P < 0.0001$ ; night 2: WT =  $5.2 \pm 0.6$  vs. Mo-Zfp271 =  $9.3 \pm 2$ ,  $P = 0.05$ , t test) (Fig. 5E).

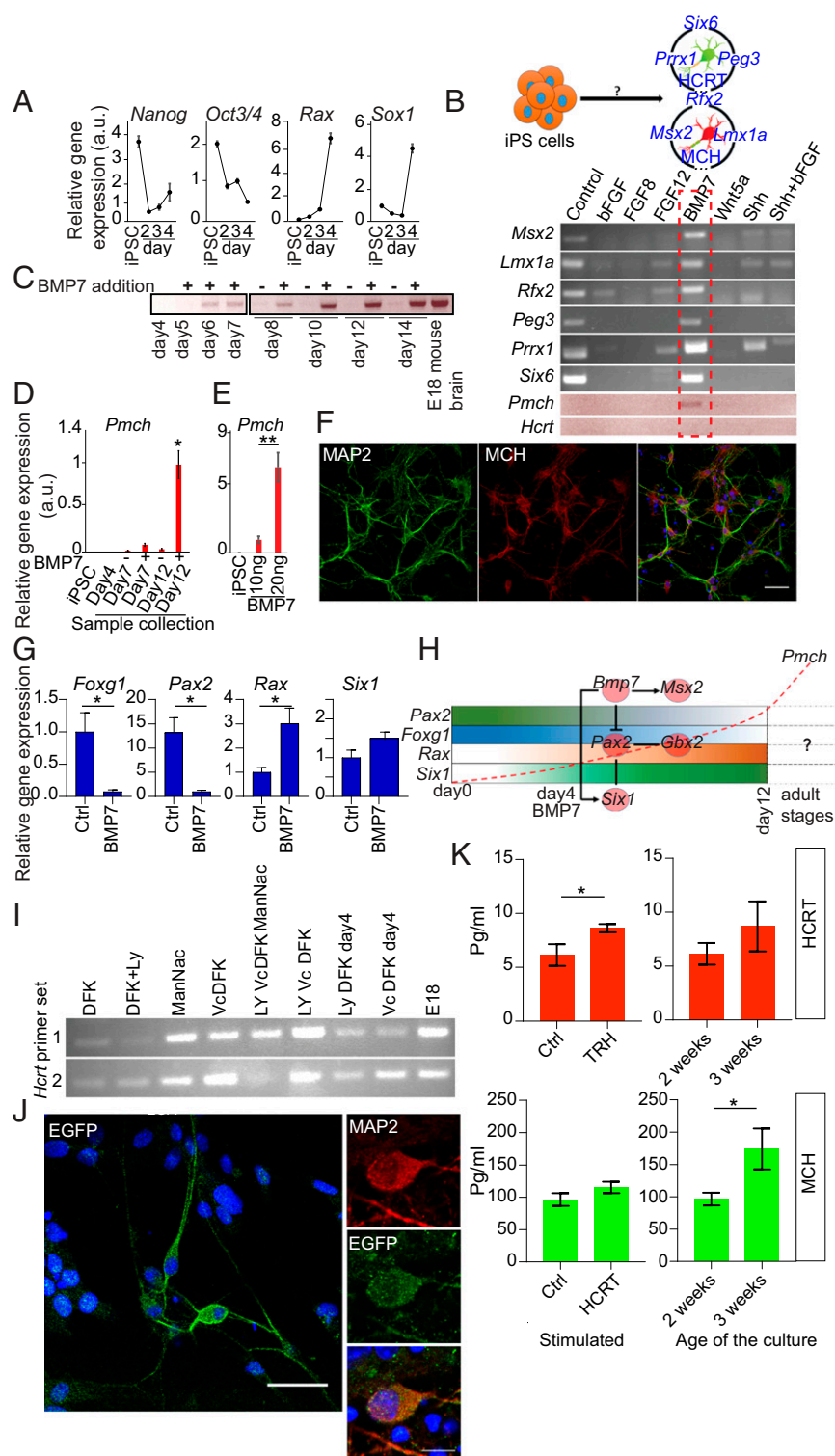
**Peg3 Loss-of-Function in Mice Reduces the Number of HCRT and MCH Cells.** *Peg3* is expressed from the paternal allele and female mice inheriting a targeted deletion of this allele display striking impairment of maternal behavior and decreased number of oxytocin neurons (43). In the same model system, we found a 16.31% decrease in the number of HCRT cells (WT [n = 5]:  $1716.4 \pm 95.7$  vs. KO [n = 6]:  $1436.5 \pm 66.39$ , mean ± SEM,  $P = 0.04$ , Mann-Whitney U test) and a 15.44% reduction in MCH cell number (WT [n = 5]:  $3307.4 \pm 123.6$  vs. KO [n = 6]:  $2796.5 \pm 154.2$ , mean ± SEM,  $P = 0.02$ , Mann-Whitney U test) (Fig. 5F). These significant reductions confirmed a role for *Peg3* in regulating these critically important cell types.

We also investigated *Peg3* gene-expression levels in mice at different postnatal time points. The results showed that at postnatal day 21, when *Hcrt* and *Pmch* reach the peak of expression, *Peg3* also reaches its maximum level, indicating a strong temporal expression relationship between the 3 genes (Fig. 5G).

**Transcriptome Data as a Roadmap to Generate MCH and HCRT Cells from Mouse Induced Pluripotent Stem Cells.** To further understand the development of HCRT and MCH neurons, we aimed at establishing these neuronal cell types from pluripotent stem cells. We first focused on the generation of HCRT cells to monitor them easily in culture; we established induced pluripotent stem cells (iPSCs) from skin fibroblast of *Hcrt-eGFP* mice. This was performed by introducing 4 TFs (OCT3/4, SOX2, c-MYC, and KLF4) in fibroblasts and after 2 wk embryonic stem cell-like colonies were picked up and expanded. The colonies expressed known pluripotency markers (SI Appendix, Fig. S5 A and C) and silenced their transgenes, which normally should







**Fig. 6.** Generation of HCRT and MCH neurons from iPSCs. (A) qRT-PCR for pluripotency (*Oct3/4*, *Nanog*), hypothalamic (*Rax*), and early neuroepithelial (*Sox1*) markers. A large decrease is observed in the expression level of pluripotent markers within 4 d while hypothalamic genes exhibit marked increase. (B, Upper) Schematic figure, which indicates that induction of candidate genes enriched in HCRT and MCH cells, with the help of a specific molecule or condition(?), might induce the expression of *Hcrt* and *Pmch* genes. (Lower) RT-PCR for TFs enriched in HCRT and MCH cells in cultures exposed to different signaling molecules and growth factors. iPSCs using protocol in *SI Appendix*, Fig. S5F were exposed to different molecules with potential role in the hypothalamus development. BMP7 showed strong potential to induce all indicated TFs and also *Pmch*. (C) The expression pattern of *Pmch* in the samples with and without BMP7 after different days in culture. Differentiating neurospheres were treated with and without 10 ng/mL BMP7 from day 4. Only in BMP7-treated samples a large expression of *Pmch* was observed and not in controls. (D) qRT-PCR for *Pmch* expression. Samples were treated as in C. (E) The effect of different concentrations of BMP7 on expression pattern of *Pmch*. Results show that the effect of BMP7 on the expression of *Pmch* is dose-dependent. (F) Neurospheres were kept in differentiation medium for 12 to 14 d before being dissociated and plated on PDL-coated dishes to produce mature neurons. Two weeks after plating the cells were stained for MCH and MAP2 expression. (G) qRT-PCR for representative genes among 30 tested genes downstream of *Bmp7*. (H) Schematic model of how TFs interact with each other to induce *Pmch* expression. For details see Discussion. (I) qRT-PCR for *Hcrt* expression under different differentiation conditions. The same method as for MCH differentiation was used to drive *Hcrt* expression except that ascorbic acid was used instead of BMP7 from day 3. Ascorbic acid induced *Hcrt* expression alone and in combination with other tested molecules (*SI Appendix*, Table S5). Due to short RNA sequence of the *Hcrt* gene, 2 different primer sets were used. (J) Hcrt-eGFP iPSC-derived, ascorbic acid-treated hypothalamic cells were stained for EGFP and MAP2 expression. (K) Two weeks after plating the neurospheres, differentiated HCRT and MCH cells were stimulated with TRH for HCRT and HCRT-1 for MCH cells for 3 h and the medium were collected to measure secreted molecules in the medium. Bar graphs show ELISA results for secretion analysis of HCRT and MCH neuropeptides. \**P* = 0.04 for HCRT and *P* = 0.039 for MCH; \*\**P* < 0.01; *n* = 3, mean ± SEM, *t* test. For D, E, G, and K: \**P* < 0.05; \*\**P* < 0.01. Scale bars: F and J, Left, 50 μm; J, Right, 10 μm.)

their specific markers the amplification of the reporter gene is required for which our developed method is of high efficiency.

We showed that the development of an intact hypocretin system requires PEG3, AHR, NR2F2, SIX6, PRRX1, and NKX6-2 as TFs, and that of the MCH system requires MSX2, GBX2, LMX1A, RFX2, SIX1, RAX, as well as PEG3. The respective role of these transcriptional factors in the development of HCRT and MCH needs further investigations by loss-of-function experiments in mice. Recently, in a comprehensive study by using single-cell

sequencing (23), the molecular signatures of HCRT and MCH neurons in the adult mouse brain were identified, which can be complementary to our embryonic findings in constructing an overall transcriptional blueprint for these cell types. Although due to the developmental age differences we do not expect identical findings between these 2 studies, we did find important genes and TFs shared between the 2 for HCRT cells (*Slc17a6*, *Pdyn*, *Scg2*, *Rasgrp1*, and *Plagl1*). Interestingly, *PLAGL1*, which was recently reported to have a role in *Hcrt* transcription (51), is enriched in



both datasets, suggesting that this TF is substantially expressed from embryonic stage until adulthood and might be involved in controlling the *Hcrt* expression. We established the transcriptional machinery for these 2 cell types and illustrated the interaction of TFs with each other. The identified networks shed light on how these cell types might use their machinery during normal development and adult functioning. These networks might also explain how these cells integrate diverse inputs and function properly on demand. We subgrouped TFs for development, sleep, and feeding behavior and showed how they might interact with each other. The important role of LHX9 and EBF2 both in development and sleep are documented (24, 50), and we add here PEG3, which displayed a critical role in development and sleep. The precise fine-tuning of these networks during day and night guarantees the proper functioning of these cells that otherwise may lead to sleep disruption, obesity, and metabolic disorders.

Loss-of-function studies confirm the importance of our findings at the developmental level, indicating that the TFs found here are required for proper development of these cell types. Removal of the *Peg3* ortholog (*Zfp271*) totally abolished HCRT and MCH expression in zebrafish and deletion of the paternal allele of *Peg3* significantly decreased the number of HCRT and MCH neurons in mice. Note that *Zfp271* is an ortholog of *Peg3* and shares similarities to several other zinc finger proteins. Therefore, *Zfp271* in zebrafish may not be functionally identical to *Peg3* in mice. Additionally, the *Zfp271*-related sleep phenotype seen in zebrafish might be different in mice. We showed that knockdown of *Zfp271* affects locomotor activity in zebrafish; however, in mice it might affect other aspects of sleep, such as sleep fragmentation and the sleep EEG and not its total amount, as shown in *Hcrt*<sup>KO/KO</sup> mice. Such species differences need further investigation. *Peg3* is an imprinted gene with close connectivity to *Zim1* (52). It was established that the *Magel2* imprinted gene, which is similar to *Peg3*, a paternally expressed gene, also regulates *Hcrt* expression (53). *Peg3* has sex- and tissue-specific promoters (54) and is believed, like many other imprinted genes, to have a monoallelic pattern of expression, although it was shown recently that in some regions of the brain, specifically in the hypothalamus, there is a biallelic pattern of expression (55). Our finding that the deletion of the paternal allele significantly decreases but does not abolish the expression of HCRT and MCH suggests that some compensation mechanisms, including a role for the maternal allele, are at play. Furthermore, *Peg3* might play a role in the early development of HCRT and MCH neurons and not at the adult stage. How HCRT and MCH systems depend on imprinted genes and whether *Peg3* also depends on the core clock machinery remain to be discovered. The involvement of *Peg3* in the modulation of sleep in zebrafish seems dependent on its effect on *Hcrt* and *Mch* expression, but whether HCRT and MCH are the only affected cell types and whether both sexes are affected at the same level remains to be investigated, as it was shown that hypothalamic oxytocin neurons and maternal behavior are affected in *Peg3* female KO mice (43). *Peg3* in association with *Traf2* is involved in the TNF- $\kappa$ B signal transduction pathway, which leads to cell survival or death (56). Therefore, it can be speculated that in some pathological conditions, such as in narcolepsy where HCRT neurons seem to be selectively destroyed, *Peg3* can participate in HCRT cell death.

The difficulties in accessing intact neurons from the human brain prompted us to establish protocols to produce these cells from iPSCs. We established iPSCs from mouse fibroblasts and differentiated them toward lateral hypothalamic cell types. We identified BMP7 as the master signaling molecule that regulates the main TFs of both cell types and the expression of *Pmch*. RNA-seq of MCH cells indicated that GBX2 is a TF up-regulated in these cells. *Gbx2* expression highly influences the expression of *Pax2* (57, 58), which itself is suppressed by BMP7 (Fig. 6G). *Six1*, which is also up-regulated by BMP7 treatment (Fig. 6G), has strong genetic interaction with *Pax2* (59). On the other hand, *MSX2*, which is the TF with the highest (5.5-fold change) overexpression in MCH cells, is highly expressed under

BMP7 exposure (Fig. 6B). Based on these findings we propose a model for the generation of MCH cells from stem cells (Fig. 6H). This model suggests BMP7 as the major player in fate determination of MCH cells.

To induce the expression of the HCRT gene we applied vitamin C, of which its role in CNS development is well-documented. It was shown in the fetal rat brain that vitamin C increases from 374  $\mu$ g/g on the 15th day of gestation to 710  $\mu$ g/g by the 20th day and remains at this level until birth (60). Our findings are complementary and add induction markers to the other studies that successfully generated hypothalamic cells from human and mouse stem and iPSCs (48, 61–63). Since we derived HCRT neurons from *Hcrt-EGFP* mice expressing eGFP under the control of a 3.4-kb human *Hcrt* promoter (64), the generated cells may not be identical to the authentic mouse HCRT neurons. However, we believe that the authentic promoter is also activated as we reported the expression of *Hcrt* at mRNA and protein levels.

We identified genes and molecules which might have a role in the pathophysiology of narcolepsy with cataplexy. QRFP is the latest hypothalamic neuropeptide identified and implicated in feeding, locomotor activity, arousal, and metabolism (65–67). More recently QRFP and its receptor GPR103 were found to regulate sleep in zebrafish (42). Additionally, the QRFP receptor GPR103 heterodimerizes with HCRT receptors and is involved in protection against Alzheimer's disease (68). QRFP seems not to be colocalized with HCRT, suggesting that its reduced expression in *Hcrt-Ataxin-3* mice is secondary to the lack of HCRT cells. Since the expression of *Qrfp* is normal in *Hcrt*<sup>KO/KO</sup> mice, we speculate that its expression is under the control of other factors produced by HCRT neurons. A secondary reduction of QRFP in narcolepsy, if confirmed, may explain auxiliary symptoms, such as food intake behavior and obesity (30, 69). The fact that the identified TFs here are also highly enriched in the immune system suggests that more investigation on their role in the pathophysiology of narcolepsy with cataplexy is warranted.

## Methods

**Animals.** Two mouse models of narcolepsy [*Hcrt*<sup>KO/KO</sup>: B6.129S6-*Hcrt*<sup>tm1Ywaj</sup> and *Hcrt-ataxin-3*: B6;D2-Tg(HCRT-Ataxin-3)] were used to perform RNA sequencing of the different brain regions. *Hcrt-Cre* [B6;D2-Tg(HCRT-Cre-IRES-eGFP)] mice were used to purify HCRT cells and WT C57BL/6J mice were used for purification of MCH cells. *Hcrt-EGFP* (B6;D2-Tg(HCRT-eGFP)) mice were used to establish iPSCs. *Peg3* KO (129Sv.MF1-*Peg3*<sup>tm1</sup>) mice were used for HCRT and MCH staining and counting. All animals were back-crossed to C57BL/6J for at least 10 generations and provided by international collaborators, except for C57BL/6J mice, which were purchased from Charles Rivers Laboratory. Animals were maintained under standard animal housing conditions with free access to food and water. All experimental protocols were approved by the Veterinary Ethical commission of the state of Vaud, Switzerland.

**RNA Sequencing.** Transcriptome analysis from mouse models of narcolepsy were performed as described in Nikolaeva et al. (70) and aligned against *Mus musculus* GRCh38.75. The R package DESeq2 was used for statistical tests. For sorted hypothalamic cells, double-stranded cDNA for RNA-seq library preparation was generated using SMART-Seq v4 Ultra Low Input RNA reagents (Catalog no. 634888, Clontech) according to the protocol provided with the reagents beginning with 2 ng (MCH cells) or 0.8 ng (for HCRT cells) of total RNA and using 11 cycles of PCR amplification. Next, 150 pg of the resulting cDNA were used for library preparation with the Illumina Nextera XT DNA Library reagents (Catalog no. 15032354, Illumina) using the single-cell RNA-seq library preparation protocol developed for the Fluidigm C1 (Fluidigm). Data analysis is summarized in *SI Appendix, Supplementary Information Text*.

**Behavioral Assays.** Sleep and activity of zebrafish were assessed based on refs. 41 and 42. Briefly, the behavioral tests were performed on zebrafish injected with 2.5 ng of *Zfp271*-Mo, or with H<sub>2</sub>O as control. From 110 to 158 h post-fertilization, larval zebrafish were placed into a 96-well plate, 1 per well, in 300  $\mu$ L of E3 embryo medium. Locomotor activity was analyzed by the video tracking module of the classic version of Zebrafish (Viewpoint Lifescience). The system applied a continuous infrared detection under a 14/10-h cycle of

light/dark. The quantification mode of Zebrabox was set up following the parameters described in Chen et al. (42): integration time bins, 1 min; detection threshold, 15; burst, 29; freeze, 3. Any 1-min epoch with less than 0.1 s of movement was considered as 1 min of sleep (42). Finally, a sleep bout was defined as a continuous string of sleep minutes. Data were processed using MATLAB.

1. M. W. Schwartz, S. C. Woods, D. Porte, Jr, R. J. Seeley, D. G. Baskin, Central nervous system control of food intake. *Nature* **404**, 661–671 (2000).
2. O. K. Hassani, P. Henny, M. G. Lee, B. E. Jones, GABAergic neurons intermingled with orexin and MCH neurons in the lateral hypothalamus discharge maximally during sleep. *Eur. J. Neurosci.* **32**, 448–457 (2010).
3. R. R. Konadhode et al., Optogenetic stimulation of MCH neurons increases sleep. *J. Neurosci.* **33**, 10257–10263 (2013).
4. G. C. Harris, M. Wimmer, G. Aston-Jones, A role for lateral hypothalamic orexin neurons in reward seeking. *Nature* **437**, 556–559 (2005).
5. C. Broberger, L. De Lecea, J. G. Sutcliffe, T. Hökfelt, Hypocretin/orexin- and melanin-concentrating hormone-expressing cells form distinct populations in the rodent lateral hypothalamus: Relationship to the neuropeptide Y and agouti gene-related protein systems. *J. Comp. Neurol.* **402**, 460–474 (1998).
6. C. F. Elias et al., Chemically defined projections linking the mediobasal hypothalamus and the lateral hypothalamic area. *J. Comp. Neurol.* **402**, 442–459 (1998).
7. C. Peyron et al., Neurons containing hypocretin (orexin) project to multiple neuronal systems. *J. Neurosci.* **18**, 9996–10015 (1998).
8. T. Sakurai et al., Orexins and orexin receptors: A family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585 (1998).
9. Y. Chen et al., Targeted disruption of the melanin-concentrating hormone receptor-1 results in hyperphagia and resistance to diet-induced obesity. *Endocrinology* **143**, 2469–2477 (2002).
10. J. Li, Z. Hu, L. de Lecea, The hypocretins/orexins: Integrators of multiple physiological functions. *Br. J. Pharmacol.* **171**, 332–350 (2014).
11. W. J. Giardino, L. de Lecea, Hypocretin (orexin) neuromodulation of stress and reward pathways. *Curr. Opin. Neurobiol.* **29**, 103–108 (2014).
12. S. Jégo et al., Optogenetic identification of a rapid eye movement sleep modulatory circuit in the hypothalamus. *Nat. Neurosci.* **16**, 1637–1643 (2013).
13. M. Shimada, N. A. Tritos, B. B. Lowell, J. S. Flier, E. Maratos-Flier, Mice lacking melanin-concentrating hormone are hypophagic and lean. *Nature* **396**, 670–674 (1998).
14. G. Conduetier et al., Melanin-concentrating hormone regulates beat frequency of ependymal cilia and ventricular volume. *Nat. Neurosci.* **16**, 845–847 (2013).
15. B. I. Baker, D. J. Bird, J. C. Buckingham, Salmonid melanin-concentrating hormone inhibits corticotrophin release. *J. Endocrinol.* **106**, R5–R8 (1985).
16. A. Adamantidis, L. de Lecea, A role for Melanin-Concentrating Hormone in learning and memory. *Peptides* **30**, 2066–2070 (2009).
17. L. Lin et al., The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**, 365–376 (1999).
18. C. Peyron et al., A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nat. Med.* **6**, 991–997 (2000).
19. R. M. Chemelli et al., Narcolepsy in orexin knockout mice: Molecular genetics of sleep regulation. *Cell* **98**, 437–451 (1999).
20. R. S. Liblau, A. Vassalli, A. Seifinejad, M. Tafti, Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy. *Lancet Neurol.* **14**, 318–328 (2015).
21. D. Latorre et al., T cells in patients with narcolepsy target self-antigens of hypocretin neurons. *Nature* **562**, 63–68 (2018).
22. D. Qu et al., A role for melanin-concentrating hormone in the central regulation of feeding behaviour. *Nature* **380**, 243–247 (1996).
23. L. E. Mickelsen et al., Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nat. Neurosci.* **22**, 642–656 (2019).
24. J. Dalal et al., Translational profiling of hypocretin neurons identifies candidate molecules for sleep regulation. *Genes Dev.* **27**, 565–578 (2013).
25. J. Liu et al., Evolutionarily conserved regulation of hypocretin neuron specification by *Lhx9*. *Development* **142**, 1113–1124 (2015).
26. M. Honda et al., IGFBP3 colocalizes with and regulates hypocretin (orexin). *PLoS One* **4**, e4254 (2009).
27. V. Cvetkovic-Lopes et al., Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. *J. Clin. Invest.* **120**, 713–719 (2010).
28. L. Yelin-Bekerman et al., Hypocretin neuron-specific transcriptome profiling identifies the sleep modulator *Kcnh4a*. *eLife* **4**, e08638 (2015).
29. L. E. Mickelsen et al., Neurochemical heterogeneity among lateral hypothalamic hypocretin/orexin and melanin-concentrating hormone neurons identified through single-cell gene expression analysis. *eNeuro* **4**, ENEURO.0013-17.2017 (2017).
30. J. Hara et al., Genetic ablation of orexin neurons in mice results in narcolepsy, hyperphagia, and obesity. *Neuron* **30**, 345–354 (2001).
31. S. Gopalakrishnan, B. A. Sullivan, S. Trazzi, G. Della Valle, K. D. Robertson, DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. *Hum. Mol. Genet.* **18**, 3178–3193 (2009).
32. K. Okamoto et al., QRFP-deficient mice are hypophagic, lean, hypoactive and exhibit increased anxiety-like behavior. *PLoS One* **11**, e0164716 (2016).
33. R. A. Romanov et al., Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
34. T. Matsuki et al., Selective loss of GABA(B) receptors in orexin-producing neurons results in disrupted sleep/wakefulness architecture. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4459–4464 (2009).
35. Y. Li, X. B. Gao, T. Sakurai, A. N. van den Pol, Hypocretin/Orexin excites hypocretin neurons via a local glutamate neuron-A potential mechanism for orchestrating the hypothalamic arousal system. *Neuron* **36**, 1169–1181 (2002).
36. T. Shimogori et al., A genomic atlas of mouse hypothalamic development. *Nat. Neurosci.* **13**, 767–775 (2010).
37. A. Weyer, K. Schilling, Developmental and cell type-specific expression of the neuronal marker NeuN in the murine cerebellum. *J. Neurosci. Res.* **73**, 400–409 (2003).
38. K. K. Kim, R. S. Adelstein, S. Kawamoto, Identification of neuronal nuclei (NeuN) as Fox-3, a new member of the Fox-1 gene family of splicing factors. *J. Biol. Chem.* **284**, 31052–31061 (2009).
39. J. A. González, E. Horjales-Araujo, L. Fugger, C. Broberger, D. Burdakov, Stimulation of orexin/hypocretin neurones by thyrotropin-releasing hormone. *J. Physiol.* **587**, 1179–1186 (2009).
40. F. Lu et al., Rax is a selector gene for mediobasal hypothalamic cell types. *J. Neurosci.* **33**, 259–272 (2013).
41. 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).
42. A. Chen et al., QRFP and its receptors regulate locomotor activity and sleep in zebrafish. *J. Neurosci.* **36**, 1823–1840 (2016).
43. L. Li et al., Regulation of maternal behavior and offspring growth by paternally expressed *Peg3*. *Science* **284**, 330–333 (1999).
44. T. Wataya et al., Minimization of exogenous signals in ES cell culture induces rostral hypothalamic differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11796–11801 (2008).
45. G. Alvarez-Bolado, F. A. Paul, S. Blaess, Sonic hedgehog lineage in the mouse hypothalamus: From progenitor domains to hypothalamic regions. *Neural Dev.* **7**, 4 (2012).
46. H. Wichterle, I. Lieberam, J. A. Porter, T. M. Jessell, Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**, 385–397 (2002).
47. N. Maeda et al., Aortic wall damage in mice unable to synthesize ascorbic acid. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 841–846 (2000).
48. K. Hayakawa et al., Epigenetic switching by the metabolism-sensing factors in the generation of orexin neurons from mouse embryonic stem cells. *J. Biol. Chem.* **288**, 17099–17110 (2013).
49. J. P. Silva et al., Regulation of adaptive behaviour during fasting by hypothalamic *Foxa2*. *Nature* **462**, 646–650 (2009).
50. A. K. De La Herrán-Arita et al., Aspects of the narcolepsy-cataplexy syndrome in *O/E3*-null mutant mice. *Neuroscience* **183**, 134–143 (2011).
51. S. Tanaka et al., Involvement of *PLAGL1/ZAC1* in hypocretin/orexin transcription. *Int. J. Mol. Med.* **43**, 2164–2176 (2019).
52. A. Ye, H. He, J. Kim, Paternally expressed *Peg3* controls maternally expressed *Zim1* as a trans factor. *PLoS One* **9**, e108596 (2014).
53. S. V. Kozlov et al., The imprinted gene *Mage12* regulates normal circadian output. *Nat. Genet.* **39**, 1266–1272 (2007).
54. B. P. Perera, J. Kim, Sex and tissue specificity of *Peg3* promoters. *PLoS One* **11**, e0164158 (2016).
55. B. P. Perera, R. Teruyama, J. Kim, *Yy1* gene dosage effect and bi-allelic expression of *Peg3*. *PLoS One* **10**, e0119493 (2015).
56. F. Relaix, X. J. Wei, X. Wu, D. A. Sassoon, *Peg3/Pw1* is an imprinted gene involved in the TNF-NFκB signal transduction pathway. *Nat. Genet.* **18**, 287–291 (1998).
57. M. Hidalgo-Sánchez, R. Alvarado-Mallart, I. S. Alvarez, *Pax2*, *Otx2*, *Gbx2* and *Fgf8* expression in early otic vesicle development. *Mech. Dev.* **95**, 225–229 (2000).
58. T. Katahira et al., Interaction between *Otx2* and *Gbx2* defines the organizing center for the optic tectum. *Mech. Dev.* **91**, 43–52 (2000).
59. P. X. Xu et al., *Six1* is required for the early organogenesis of mammalian kidney. *Development* **130**, 3085–3094 (2003).
60. C. C. Kratzing, J. D. Kelly, J. E. Kratzing, Ascorbic acid in fetal rat brain. *J. Neurochem.* **44**, 1623–1624 (1985).
61. K. Ogawa et al., Vasopressin-secreting neurons derived from human embryonic stem cells through specific induction of dorsal hypothalamic progenitors. *Sci. Rep.* **8**, 3615 (2018).
62. F. T. Merkle et al., Generation of neuropeptidergic hypothalamic neurons from human pluripotent stem cells. *Development* **142**, 633–643 (2015).
63. L. Wang et al., Differentiation of hypothalamic-like neurons from human pluripotent stem cells. *J. Clin. Invest.* **125**, 796–808 (2015).
64. A. Yamanaka et al., Hypothalamic orexin neurons regulate arousal according to energy balance in mice. *Neuron* **38**, 701–713 (2003).
65. N. Chartrel et al., The RFamide neuropeptide 26Rfa and its role in the control of neuroendocrine functions. *Front. Neuroendocrinol.* **32**, 387–397 (2011).
66. S. Takayasu et al., A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7438–7443 (2006).
67. K. Ukena, T. Osugi, J. Leprince, H. Vaudry, K. Tsutsui, Molecular evolution of GPCRs: 26Rfa/GPR103. *J. Mol. Endocrinol.* **52**, T119–T131 (2014).
68. J. Davies et al., Orexin receptors exert a neuroprotective effect in Alzheimer's disease (AD) via heterodimerization with GPR103. *Sci. Rep.* **5**, 12584 (2015).
69. A. Schulz, J. Hebebrand, F. Geller, T. Pollmächer, Increased body-mass index in patients with narcolepsy. *Lancet* **355**, 1274–1275 (2000).
70. S. Nikolaeva et al., Nephron-specific deletion of circadian clock gene *Bmal1* alters the plasma and renal metabolome and impairs drug disposition. *J. Am. Soc. Nephrol.* **27**, 2997–3004 (2016).