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

Ali Seifinejad, Sha Li, Cyril Mikhail, Anne Vassalli, Sylvain Pradervand, Yoan Arribat, Hassan Pesehgi Modarres, Bridget Allen, Rosalind M. John, Francesca Amati, and Mehdi Tafiti

Abstract

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.

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.


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).

To whom correspondence may be addressed. Email: Ali.Seifinejad@unil.ch or Mehdi.Tafiti@unil.ch.

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

www.pnas.org/cgi/doi/10.1073/pnas.1902148116

PNAS Latest Articles | 1 of 10
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 to both 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). An overabundance of genes 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 anorexigenic function. Hcrt-Ataxin-3 Tg<sup>+</sup> mice exhibited a 50% reduction of Qrfp mRNA levels are about 50% reduced in the hypothalamus of Hcrt<sup>ko/ko</sup> mice (19) and 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 have 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 ± SEM, arbitrary units) derived from qRT-PCR analysis (t test, n = 3 mice; *P = 0.023; **P < 0.01; ***P < 0.001).

**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 have 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 ± SEM, arbitrary units) derived from qRT-PCR analysis (t test, n = 3 mice; *P = 0.023; **P < 0.01; ***P < 0.001).

**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).
We established a protocol in which the hypothalamus of E18 to E19 Hcrt-Cre transgenic (34) or C57BL/6J mice embryos were dissected and dissociated single hypothalamic cells were prepared and transduced with viral vectors with reporter proteins (Fig. 2A and SI Appendix, Supplemental Information Text). The challenging issue was the purification of HCRT cells due to the low number of cells and the low expression level of the Hcrt gene at this developmental stage. Therefore, regular reporter markers, like eGFP in Orexin-eGFP mice (35), are not suitable. To improve the sorting, we set up a system enabling the amplification of the reporter gene. We took advantage of an adenovirus-associated virus (AAV) vector where an mCherry gene is inserted downstream of Eef1a1 promoter in the opposite orientation between 2 loxP sites (Fig. 2B). The neurobasal medium containing cells and viral vector were dispensed in low adherent 96-well plates. This method forces cells to quickly aggregate, which improves tremendously the survival rate (Fig. 2A). Aggregates were kept in neurosphere medium for 7 d to allow mCherry accumulation and cell sorting (Fig. 2C). We were able to sort 234.87 ± 126.29 mCherry+ cells from each E18 to E19 brain and collected about more than 5,000 cells for each biological replicate (3 replicates of mCherry+ and 3 mCherry−) (SI Appendix, Fig. S1A). A total amount of high-quality RNA, ranging from 1 to 2.5 μg, was extracted from samples (SI Appendix, Fig. S1C) and RNA-seq was performed to identify DEGs.

To purify the MCH cells, which show enough E18 to E19 expression level, we isolated them based on the level of EGFP expression. To this end, we inserted the eGFP sequence between 2 long-terminal repeats of the pRRLSIN lentiviral backbone under the mouse Pnch gene promoter and the mCherry sequence under the human PGK promoter (Fig. 2D). To verify if the lentivector acts specifically, we transduced the hypothalamus, cortex, and cortex+hippocampus cell cultures. One week later the cultures were stained for the expression of EGFP and MCH. Only in hypothalamic cultures was the expression and colocalization of both peptides detected, indicating that our construct is specific (Fig. 2C). MCH cells were sorted based on the expression of this lentivirus system and approximately the same number of cells was sorted and high-quality RNA was extracted for RNA-seq.

Identification of Enriched Genes by FACS Sorting of HCRT and MCH Cells. As expected, RNA-seq results showed a large enrichment of mCherry RNA in HCRT cells (mCherry+ = 67.1 ± 21.9 vs. mCherry− = 11.2 ± 1.9, counts per million reads [cpm], mean ± SEM) and for EGFP in MCH cells (EGFP+ = 82.5 ± 2.4 vs. EGFP− = 15.6 ± 0.25, cpm, mean ± SEM) (Fig. 3A and B). This analysis revealed 652 DEGs, among them 350 genes up-regulated in HCRT cells by more than 2-fold. In addition to Hcrt, Pdyn (3-fold increase), Dlk1 (4.5-fold increase), and Adra2a (3-fold increase) are known markers of HCRT cells (24). We also identified 2,665 DEGs in MCH cells compared with other hypothalamic cell types; 468 of them were up-regulated more than 2-fold (Fig. 3C and D and SI Appendix, Table S2). A higher number of DEGs for MCH neurons might be due to the fact that these neurons appear to be more mature at E18 to E19 than HCRT neurons. In silico verification (Allen atlas) of the expression of some of these DEGs confirmed hypothalamic and lateral hypothalamic expression patterns in the adult mouse brain (SI Appendix, Fig. S2). We also confirmed the colocalization of some of these genes with HCRT and MCH by immunohistochemistry (Fig. 3E).

Unique Gene Sets Distinguish HCRT and MCH Cell Types. We identified 2 gene sets which might underlie the unique characteristics of these 2 cell types: (i) a set of genes with TF and DNA binding activity, and (ii) specific genes that are part of some important pathways and are differentially regulated between the 2 cell types.

Eighteen genes with DNA binding and TF activity were up-regulated more than 2-fold in HCRT cells and 18 in MCH cells. Only Ldb2 and Ahr were shared between the 2 cell types, with the remaining TFs being cell-type specific (Fig. 4A and SI Appendix, Table S3). Lhx9, proposed as a regulator of HCRT cell development (24, 25, 36), was not significantly changed between cell types. Of the other Lhx gene family members, only Lhx3 and Lhx6 were less expressed in MCH cells (SI Appendix, Table S2).

Pathway analysis for differentially regulated genes highlighted signaling, disease, and metabolic pathways. In this analysis, we included all significantly differentially regulated genes and identified shared but differential pathways between the 2 cell types. The gene list with opposite direction of expression is shown in Fig. 4B. These genes have exosomal, hormonal, and signaling but mainly unknown functions and have lateral hypothalamic expression. Pathway analysis indicated that some of these genes are involved in infectious diseases, autoimmune thyroid disease, mRNA surveillance, amyotrophic lateral sclerosis, and the PI3K-Akt signaling pathway (SI Appendix, Table S4). Opposite direction of expression of these genes in the 2 cell types might also explain their opposite contribution to behaviors, such as sleep (HCRT promoting wakefulness and MCH sleep).

By analyzing the expression pattern of HCRT and MCH neurons with immunohistochemistry and with the mature neuronal marker NeuN (anti-NeuN antibody against purified nuclear proteins from mouse brain), we made a serendipitous discovery. HCRT neurons do not express the NeuN marker, while MCH neurons and other hypothalamic neurons do (Fig. 4C). To test the specificity of this finding we also tested another NeuN antibody (against the first 100 amino acids of the human NeuN) and found exactly the same lack of staining of HCRT neurons. This is surprising since NeuN is expressed by all CNS mature neurons except for Purkinje and Golgi cells of the cerebellum (37). NeuN is encoded...
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+ 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 μm; Right, 20 μm.)
by Rbox3 (38), a splicing regulator, suggesting that adult HCRT neurons might have a set of differentially spliced mRNAs.

Several ion channels and receptors were also identified for each cell type (SI Appendix, Fig. S3). Finally, we detected a strong up-regulation of the Trhr gene in HCRT cells and it is known that thyrotropin-releasing hormone (TRH) stimulates these neurons (39) (SI Appendix, Table S2).

Morpholino Knockdown of Candidate TFs. To identify a possible role of the candidate TFs to modulate Hcrt and Pnch gene expression, we knocked down their expression by morpholino oligonucleotides (Mo) (SI Appendix, Supplementary Information Text and Table S7) in zebrafish embryos at the 1-cell stage. Among candidate TFs, PEG3 is not found in the zebrafish genome. We therefore used its ortholog in Danio rerio (SI Appendix, Table S2). Among 10 top TFs enriched in HCRT cells, Mo-Zfp271 at 2.5 ng strikingly abolished Hcrt gene expression (Fig. S A and B) but did not result in any developmental defect (Fig. S A and SI Appendix, Fig. S4A). Mo-Si66 (33%) and Mo-Prx1 (10%) at 1.25 ng also reduced Hcrt expression. At 2.5 ng, Mo-Si66 (45%) and Mo-Prx1 (38%) had the most pronounced effect after Mo-Zfp271 (100%), Mo-Ahr (57%), and Mo-Nr2f2 (40%), indicating that HCRT cells might arise from the Six family-expressing domain of the embryonic hypothalamus (40). The phenotype obtained from knockdown of several TFs enriched in MCH cells was partial loss of MCH signal for Mo-Gbx2, Mo-Lmx1a, and Mo-Msx2a (50%, 60%, and 22% respectively) (SI Appendix, Fig. S4B). Peg3 was not enriched in MCH neurons but to test if it has any effect on Pmch expression, we also injected Mo-Zfp271 and surprisingly observed that it also totally abolished MCH expression (SI Appendix, Fig. S4B).

Altered Locomotor Activity and Sleep Architecture in Zfp271 (Peg3 Ortholog) Knocked-Down Zebrafish. Enrichment of Peg3 in HCRT cells and its effect on the expression of HCRT and MCH neuropeptides suggested a potential involvement in sleep regulation. The effect of Zfp271 knockdown on zebrafish behavior was assessed using a locomotor activity assay, as previously described (41, 42). Mo-Zfp271 larvae exhibited significantly less locomotor activity mainly during daytime (day 1: WT = 42.6 ± 5.8 vs. Mo-Zfp271 = 21.1 ± 2.7; P < 0.01, t test) (Fig. S 5C). This reduction in activity is accompanied with 2-fold increase in sleep time during the day (day 1: WT = 24.2 ± 2.6 vs. Mo-Zfp271 = 45.2 ± 1.5; day 2: WT = 28 ± 2.2 vs. Mo-Zfp271 = 41.6 ± 1.4; P < 0.0001, t test) (Fig. S 5D). Analysis of sleep bout number and duration showed that the increase in sleep amount is mainly due to increased bout
duration both during the light (day 1: WT = 0.3 ± 0.02 vs. Mo-Zfp271 = 3.4 ± 1.3; P = 0.025, t test) and the dark period (night 1: WT = 9.5 ± 1.3 vs. Mo-Zfp271 = 20.9 ± 2.5, P < 0.0001; night 2: WT = 5.2 ± 0.6 vs. Mo-Zfp271 = 9.3 ± 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 ± 95.7 vs. KO [n = 6]: 1436.5 ± 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 ± 123.6 vs. KO [n = 6]: 2796.5 ± 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
occur after activation of internal stemness genes (SI Appendix, Fig. S5B). The established iPSCs also had strong potential to differentiate toward neural lineage (SI Appendix, Fig. S5D).

To differentiate our iPSCs toward hypothalamic neuronal progenitors, we took advantage of a protocol established by Wataya et al. (44). This method applies a strictly chemically defined medium with minimum level of inductive signals called “growth factor-free chemically defined medium” together with quick aggregation where the seeded cells are quickly precipitated in wells of low-adherent 96-well plates. Preliminary screening using this protocol with different conditions (more than 100 conditions) failed to induce Hct and Pmch gene expression (SI Appendix, Fig. S5E). Next, we tested if any specific signaling molecule can induce the expression of some TFs enriched in HCRT (Peg3, Pax2, and S6k6) and MCH (Mx2, Lmx1a, and Rfx2) cells from our RNA-seq experiment. To this end and based on the days that hypothalamic (Rax gene) and neuroepithelial (Sox1 gene) markers start to be up-regulated during neuronal differentiation (Fig. 6A), we changed the protocol and from day 4 applied different signaling molecules and growth factors individually (SI Appendix, Fig. S5F). Screening of several molecules identified BMP7 as a highly potent molecule to induce the expression of TFs of interest (Fig. 6B). Almost all 6 TFs that were tested were strongly expressed in BMP7-treated cells (Fig. 6B). Intriguingly, BMP7 induced the expression of Pmch but not Pmch-like gene (Pmcll). We next performed a detailed analysis of the effect of BMP7 on Pmch expression. Exposure of the neurospheres from day 4 showed that BMP7 exerts its effect 2 d after addition and Pmch expression is increased during several days (Fig. 6C). The effect of BMP7 was more pronounced only from day 4 (SI Appendix, Fig. S6A) and was dose-dependent (Fig. 6E and SI Appendix, Fig. S6D). Sonic hedgehog, which is one of the major regulators of hypothalamic patterning (45), was not able to induce Pmch expression, either in different media or at different concentrations (SI Appendix, Fig. S6D). The effect of the other members of BMP family, BMP4, was not as strong as BMP7 (SI Appendix, Fig. S6B). Addition of the retinoic acid, which is the important player for caudal CNS patterning (46), suppressed Pmch expression (SI Appendix, Fig. S6C). After 2 wk, the neurospheres were dissociated and cultured on coated dishes for 2 extra weeks to allow mature neuronal growth. Staining for MAP2 and MAP2 showed that MCH-expressing cells also express MAP2, indicating that these cells are mature neurons (Fig. 6F). We tested if any specific combination of TFs other than MCH cell-specific TFs, are affected by BMP7, which results in the induction of Pmch gene expression. qPCR expression of about 30 candidate genes showed that, while genes from the Pax family are strongly down-regulated under BMP7 condition, Rax, an important hypothalamic marker, is highly expressed. Six1 also showed increased expression. Foxg1, a major marker of telencephalon development, was strongly suppressed as expected (Fig. 6G). Based on our data and the literature, we propose a model to explain how different combinations of TFs might interact with each other and how their up- or down-regulation induces Pmch expression (Fig. 6H).

All above-mentioned approaches failed to induce any expression of HCRT, even though the treatment with BMP7 induced HCRT-enriched TFs (SI6k6, PEG3, and PRRX1). We performed another large screen of molecules (SI Appendix, Fig. S6E and Table S5) and included vitamin C as one of the treatments. The reason for the inclusion of vitamin C was the enrichment of the Gulo (L-gulonolactone oxidase) gene in our HCRT cell RNA-seq experiment as well as our RNA-seq of Hct-ataxin-3 mice (SI Appendix, Fig. S6F). L-Gulonolactone oxidase is the enzyme producing the precursor to ascorbic acid (47). Although a few of the screened molecules induced HCRT expression, only vitamin C treatment could be reproduced (SI Appendix, Fig. S6G). Inclusion of vitamin C in combination with different molecules showed its potential to induce the expression of the Hct gene (Fig. 6I). Vitamin C displays almost the same effect as ManNac, which has previously been shown to also induce the expression of Hct gene (48).

After plating the cells in poly-d-lysine (PDL)-coated dishes, EGFP-expressing neurons derived from Hct-EGFP iPSCs were detected in our cultures (Fig. 6J). To test if our in vitro MCH and HCRT cells are functional, we measured the secretion of their respective neuropeptide in the cell culture medium. MCH and HCRT cells plated on PDL dishes for 2 wk secreted about 96.61 ± 9.87 and 6.14 ± 1 pg/mL of their neuropeptides, respectively (Fig. 6K). After stimulation with an inducer molecule (TRH [125 nM] for HCRT cells and HCRT-1 [1 μM] for MCH cells), the levels increased to 115.35 ± 8.84 and 8.63 ± 0.38 pg/mL. Maturation stage also affects their functionality and more mature neurons showed more secretion of MCH (3-wk culture: 174.03 ± 31.67 vs. 2-wk culture: 96.61 ± 9.87 pg/mL) and HCRT (3-wk culture: 8.68 ± 2.3 vs. 2-wk culture: 6.14 ± 1 pg/mL, mean ± SEM, n = 3) neuropeptides (Fig. 6K).

Network of TFs in HCRT and MCH Cells. Partial or complete effect of the removal of a TF depends on the gene network. Based on the literature and different databases (BioGRID, TRRUST, and STRING), we predicted a potential network for the 18 identified TFs and their targets (physical and genetic interaction) in HCRT cells, which might explain the results of our knockdown experiment in zebrafish. PEG3 has the largest weight on the network since its removal totally abolishes HCRT expression (Figs. 5B, E, and F), which was the most effective among all of our RNAi screening data (iv) TFs regulating genes up-regulated more than 2-fold, and (iii) TFs with known functions in a specific behavior, such as FOXA2 (49) in feeding and LHX9 (24) and EBF2 (50) in sleep. For all TFs we also extracted if they have a salient role in a specific phenotype and if they are part of a specific pathway. Some TFs participated in known functions, such as development, sleep, and feeding behavior and obesity in both HCRT and MCH cells (SI Appendix, Fig. S7). Not only are there shared TFs between different functional groups, but they also interact with each other. Normally the output of this gene network is the expression or inhibition of Hct, Pmch (Dynorphin), and glutaminase for HCRT cells and the expression or inhibition of Pmch and Gaba for MCH cells, and proper functioning of these cell types (SI Appendix, Fig. S7). If these outputs are altered under different physiological and pathological conditions, it requires alteration in the gene network, as well. No TF can be excluded from any function but the figure shows how the network can change the gene output depending on the addition or removal of TFs. As such, the cell types might be maintained or altered to change their behavior and integration with other cell types.

Discussion Using the end-stage markers of lateral hypothalamic neuropeptides, which are neuropeptides HCRT and MCH, we established their transcriptional machinery. To our knowledge, this study is unique in purifying intact HCRT and MCH cells from late embryonic brains of mice. We identified molecules that might be the driving machinery of these neurons and used these markers as a guideline to develop protocols to establish these cells from mouse iPSCs. Our aim was to establish a method to dissociate and culture live neurons and to have a specific marker to isolate specific cell types. With reporter proteins under specific promoters and quick aggregation method, it was possible to specifically purify them. Nevertheless, in the case of cell types with weak expression of...
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 (Fig. 5). Thus, the findings between these 2 studies, we did find important genes and TFs shared between the 2 for HCRT cells (Fig. 5).

**Discussion**

Results show that the effect of BMP7 on the expression of Pmch and Hcrt is dose-dependent. The effect of different concentrations of BMP7 in the sample 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. The 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. (I) Schematic model of how TFs interact with each other to induce Pmch expression. For details see Discussion.

**(B) 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.**

**(C) The expression pattern of Pmch in the samples with and without BMP7 after different days in culture.**

**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 pluripotency 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.
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. This indicates that Hcrt expression is influenced functionally by 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 programming of these cells that otherwise might be more imprinted genes to have a monolaelic 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 at the adult stage. How HCRT and MCH systems depend on either genetic or whether Peg3 also depends on the collective 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 Trn2β is involved in the TNF–NF–eB signal transduction pathway, which leads to cell survival or death (50). 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 μg/g on the 15th day of gestation to 710 μg/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>koxk</sup> (B6;129S6-Hcrt<sup>-tm1Yvy</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 (129SvMF1-Peg3<sup>−/−</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 River Laboratory. Animals were maintained under standard animal housing conditions and had 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* GRCm38.75. 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 postfertilization, larval zebrafish were injected into a Stock embryo with 300 µL 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 quantitative fraction of zebrafish 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.

64. A. Yamanaka et al., Hypothalamic orexin neurons regulate arousal according to energy balance in mice. Neuron 38, 701–713 (2003).