3-15-2015

Evolutionarily Conserved Regulation of Hypocretin Neuron Specification by Lhx9

Justin Liu
Florian T. Merkle
Avni V. Gandhi
James A. Gagnon
Ian Woods
Ithaca College, iwoods@ithaca.edu

See next page for additional authors

Follow this and additional works at: http://digitalcommons.ithaca.edu/biology_faculty_pubs

Part of the Animal Sciences Commons, Biology Commons, and the Genetics and Genomics Commons

Recommended Citation
http://digitalcommons.ithaca.edu/biology_faculty_pubs/20

This Article is brought to you for free and open access by the Biology Department at Digital Commons @ IC. It has been accepted for inclusion in Biology Faculty Publications and Presentations by an authorized administrator of Digital Commons @ IC.
Evolutionarily conserved regulation of hypocretin neuron specification by Lhx9

Justin Liu1, Florian T. Merkle2,3, Avni V. Gandhi1, James A. Gagnon2, Ian G. Woods2,*, Cindy N. Chiu1, Tomomi Shimogori4, Alexander F. Schier2,3,5 and David A. Prober1,†

ABSTRACT

Loss of neurons that express the neuropeptide hypocretin (Hcrt) has been implicated in narcolepsy, a debilitating disorder characterized by excessive daytime sleepiness and cataplexy. Cell replacement therapy, using Hcrt-expressing neurons generated in vitro, is a potentially useful therapeutic approach, but factors sufficient to specify Hcrt neurons are unknown. Using zebrafish as a high-throughput system to screen for factors that can specify Hcrt neurons in vivo, we identified the LIM homeobox transcription factor Lhx9 as necessary and sufficient to specify Hcrt neurons. We found that Lhx9 can directly induce hcart expression and we identified two potential Lhx9 binding sites in the zebrafish hcart promoter. Akin to its function in zebrafish, we found that Lhx9 is sufficient to specify Hcrt-expressing neurons in the developing mouse hypothalamus. Our results elucidate an evolutionarily conserved role for Lhx9 in Hcrt neuron specification that improves our understanding of Hcrt neuron development.

KEY WORDS: Hypocretin, Lhx9, Zebrafish

INTRODUCTION

The neuropeptide hypocretin (Hcrt) is conserved among vertebrates and plays key roles in regulating sleep, metabolism, feeding, anxiety, reward and addiction (Bonnavi and de Lecea, 2010; Tsujino and Sakurai, 2009). Hcrt is particularly important in promoting arousal, as loss of Hcrt neurons is thought to cause narcolepsy (Peyron et al., 2000; Thannickal et al., 2000), a disorder characterized by daytime sleepiness, fragmented sleep-wake states and cataplexy. Narcolepsy affects approximately 1 in 2000 individuals, but treatments are limited to symptom management (Dauvilliers et al., 2007). Despite the importance of the Hcrt system, little is known about the developmental processes that give rise to Hcrt neurons. A recent study found that mice lacking the LIM domain homeobox transcription factor Lhx9 had fewer Hcrt neurons (Dalal et al., 2013), suggesting that Lhx9 is required to specify a subset of Hcrt neurons. However, overexpression of Lhx9 in adult mice or in a mouse neoblastoma cell line had no effect on Hcrt cell number or expression. Therefore, the role of Lhx9 in Hcrt neuron specification remains unclear and the set of factors sufficient to specify Hcrt neurons remains unknown. Identifying these factors would help elucidate how a key neural circuit that governs sleep is established, and could lead to novel therapies for narcolepsy.

The zebrafish Danio rerio is a powerful genetic model of vertebrate development that provides several advantages for studying Hcrt neuron specification. First, the hypothalamus is remarkably conserved (Blackshaw et al., 2010; Machluf et al., 2011; Tessmar-Raible et al., 2007), suggesting that developmental mechanisms identified in zebrafish are likely to be relevant to mammals. Several studies have shown that the mammalian Hcrt system is functionally and anatomically conserved in zebrafish (Chiu and Prober, 2013; Elbaz et al., 2013). Whereas the rodent hypothalamus contains thousands of Hcrt neurons, larval and adult zebrafish contain only approximately 10 and 40 Hcrt neurons, respectively (Faraco et al., 2006; Kaslin et al., 2004), making zebrafish a more tractable system to study Hcrt neuron development. Second, the external development and transparency of zebrafish embryos facilitate the observation of developing Hcrt neurons. Third, high-throughput genetic gain- and loss-of-function assays facilitate efficient screens to identify developmental regulators. We exploited these features of zebrafish to identify genes that regulate Hcrt neuron specification.

RESULTS

Microarray analysis identifies transcripts enriched in Hcrt neurons

Previous studies showed that the number of Hcrt neurons in zebrafish and rodents increases as animals develop and mature to adulthood (Fanco et al., 2006; Kaslin et al., 2004; Sawai et al., 2010). We reasoned that cell-autonomous factors required to specify Hcrt neurons might still be expressed in Hcrt neurons shortly after they are specified. To identify these factors, we generated transgenic zebrafish that express monomeric red fluorescent protein (mRFP) in Hcrt neurons and enhanced green fluorescent protein (EGFP) in neurons that express the hypothalamic neuropeptide QRFP (supplementary material Fig. S1). QRFP has been implicated in regulating locomotor activity (Takayasu et al., 2006), feeding (Chartrel et al., 2003; Takayasu et al., 2006) and nociception (Yamamoto et al., 2009) in rodents, and sleep/wake behaviors in zebrafish (C.N.C., A. Chen and D.A.P., unpublished).

Expression of hcart and qrfp (si:ch211-185o22.2, incorrectly annotated as lineRNA) is first detected in zebrafish embryos at 24 h post-fertilization (hpf) in bilateral hypothalamic nuclei of 4-6 cells (Fig. 1A,B), which expand to 10-15 cells by 120 hpf(Fig. 1C,D). Hcrt and QRFP are expressed in neighboring neurons throughout development, but are never co-expressed within the same cells (Fig. 1B,D). To identify genes with enriched expression in Hcrt neurons, we dissociated pools of 100-300 Tg(hcrt:mRFP, qrfp: EGFP) embryos at 26 hpf into single cells and isolated EGFP- and mRFP-expressing neurons by fluorescence-activated cell sorting (FACS) (Fig. 1E; supplementary material Fig. S2). FACS gates for
EGFP and mRFP populations were set using wild-type embryos (0/10,000 EGFP⁺ or mRFP⁺ events). In a representative experiment, we obtained 250 EGFP⁺ cells and 528 mRFP⁺ cells from 150 Tg(hcrt: mRFP, qrfp:EGFP) double-heterozygous embryos. To verify the fidelity of FACS, we visually screened for fluorescence in sorted cells (Fig. 1F). In the sorted qrfp:EGFP population, we observed EGFP in 99/117 cells (85%) but no mRFP (0/117). In the sorted hcrt:mRFP population, we observed mRFP in 110/146 cells (75%) but no EGFP (0/146). These values are likely to underestimate the purity of the sorted cells because FACS is more sensitive than visual inspection.

We extracted total mRNA from each cellular fraction and used cDNA microarrays to compare gene expression in Hcrt and QRFP neurons. We also compared gene expression in Hcrt neurons with expression in neurons labeled by a pan-neuronal marker, Tg(elavl3:EGFP), and with expression in subtypes of sensory neurons: Tg(trpa1b:EGFP), Tg isl1:Gα4VP16, 14xUAS:EGFP) and Tg(p2rx3b:EGFP). elavl3 encodes a RNA-binding protein that is expressed in most postmitotic neurons (Park et al., 2000); trpa1b encodes a transient receptor potential (TRP) channel that is activated by chemical irritants (Prober et al., 2008); p2rx3b encodes an ATP-gated ion channel in non-peptidergic nociceptors (Kucenas et al., 2006); and isl1 (isl1) encodes a LIM homeobox transcription factor that is expressed in sensory neurons and motoneurons (Higashijima et al., 2000). We used an isl1 enhancer that drives expression in a subset of sensory neurons (Sagasti et al., 2005). The Tg(trpa1b:EGFP) and Tg isl1:Gα4VP16, 14xUAS:EGFP) lines express EGFP in largely non-overlapping subsets of trigeminal and Rohon-Beard sensory neurons (Pan et al., 2012). p2rx3b is expressed in all cells labeled in Tg(trpa1b:EGFP) embryos and in a quarter of cells labeled in Tg isl1:Gα4VP16, 14xUAS:EGFP) embryos. These samples allowed five pairwise comparisons of Hcrt neurons with different purified neuron populations (supplementary material Fig. S2 and Table S1). This analysis provided greater statistical power than previous studies that compared Hcrt neurons with a single outgroup (Cvetkovic-Lopes et al., 2010; Dalal et al., 2013). An additional study examined changes in gene expression across multiple brain regions after the onset of narcolepsy (Honda et al., 2009); additional study examined changes in gene expression across multiple brain regions after the onset of narcolepsy (Honda et al., 2009); however, this approach is unlikely to identify developmentally relevant transcripts. We focused on 19 highly ranked genes that encode transcription factors or secreted proteins (Tables 1 and 2), as both classes of proteins have well-established roles in neural development (Blackshaw et al., 2010; Wilson and Houart, 2004); as expected, hcrt was the most highly enriched gene in Hcrt neurons (Table 2). The complete microarray dataset is available through ArrayExpress, accession number E-MTAB-3317.

Expression patterns of candidate genes validate the microarray results

High quality in situ hybridization (ISH) images were available in the ZFIN database for seven candidate genes (supplementary material Fig. S3). We determined the expression patterns of the remaining 12 genes using ISH on 24 hpf embryos (Fig. 2; supplementary material Fig. S3 and Movie 1). We found that 11 of 14 genes encoding transcription factors and five of five genes encoding secreted proteins are expressed in a similar or overlapping domain to Hcrt neurons. Some genes, such as the transcription factor lhx9, are expressed in all Hcrt cells throughout early development (supplementary material Fig. S4). The microarray therefore accurately reported the expression of most candidate genes.

Lhx9 is sufficient to specify Hcrt neurons

To determine whether any candidate gene is sufficient to induce specification of Hcrt neurons, we cloned each gene downstream of a heat shock-inducible promoter in a vector containing Tol2 transposase sites, and injected the plasmid together with tol2 transposase mRNA into zebrafish embryos at the 1-cell stage. As a result, the plasmid inserts into the genome of a random subset of cells. We transiently overexpressed each gene by performing a heat shock at 24 hpf (Fig. 3A). This approach provides an efficient, inducible method for overexpressing different candidate genes in 10-20% of cells (Fig. 3C).

Quantification of hypothalamic Hcrt neurons using ISH at 120 hpf showed no significant differences between larvae overexpressing candidate genes and controls (supplementary material Fig. S5). However, 14% of larvae overexpressing lhx9 contained additional, ectopic Hcrt expression in the medial hindbrain (7/50 larvae; Fig. 3E,F). These larvae had 2.7 ectopic
Hcrt cells on average, a 17% increase in the total number of Hcrt neurons. No other candidate gene was sufficient to specify ectopic Hcrt neurons. Like endogenous hypothalamic Hcrt neurons, the ectopic Hcrt neurons expressed vesicular glutamate transporter 1 (vglut1; slc17a7a) and prodynorphin (pdyn) (Fig. 4). vglut1 is widely expressed in the hypothalamus and hindbrain, but its expression is particularly strong in Hcrt neurons (Fig. 4A,B; supplementary material Movie 2). Hcrt neurons also exhibit intense, punctate labeling (Fig. 4D,E; supplementary material Movie 3) that is likely to indicate sites of transcription (Hanisch et al., 2012; Kosman et al., 2004). We did not detect significant expression of vglut2a (slc17a6b) or vglut2b (slc17a6a) in Hcrt neurons (supplementary material Figs S8 and S9).

lhx9 overexpression generated more cells in the medial hindbrain with strong vglut1 expression or punctate pdyn expression than ectopic Hcrt cells (Fig. 4C,F), suggesting that Lhx9 can specify multiple cell types. Indeed, most lhx9-overexpressing neurons also expressed qrgf at 1 h post-heat shock (data not shown), and we observed an average of three ectopic QRFP neurons in the medial hindbrain 96 h after Lhx9 overexpression (8/40 larvae, supplementary material Fig. S6; note that endouc is widely expressed in the hypothalamus and hindbrain, but its expression is particularly strong in Hcrt neurons (Fig. 4A,B; supplementary material Movie 2). Hcrt neurons also exhibit intense, punctate labeling (Fig. 4D,E; supplementary material Movie 3) that is likely to indicate sites of transcription (Hanisch et al., 2012; Kosman et al., 2004). We did not detect significant expression of vglut2a (slc17a6b) or vglut2b (slc17a6a) in Hcrt neurons (supplementary material Figs S8 and S9).

lhx9 overexpression generated more cells in the medial hindbrain with strong vglut1 expression or punctate pdyn expression than ectopic Hcrt cells (Fig. 4C,F), suggesting that Lhx9 can specify multiple cell types. Indeed, most lhx9-overexpressing neurons also expressed qrgf at 1 h post-heat shock (data not shown), and we observed an average of three ectopic QRFP neurons in the medial hindbrain 96 h after Lhx9 overexpression (8/40 larvae, supplementary material Fig. S6; note that lhx9 was enriched in hcrtr neurons compared with elav3, isl1, trpa1b and p2rx3b neurons, but not compared with qrgf neurons, in the microarray analysis). Similar to Hcrt neurons, both hypothalamic and ectopic QRFP neurons express Lhx9 (supplementary material Figs S6 and S7). Ectopic QRFP neurons did not co-express hcrtr (0/24 QRFP neurons), suggesting that they remain developmentally distinct populations.

Both mammalian and zebrafish Hcrt neurons project to several brain regions, including the noradrenergic locus coeruleus (LC) (Horvath et al., 1999; Prober et al., 2006). To determine whether ectopic Hcrt neurons project to this endogenous Hcrt neuron target, we overexpressed lhx9 in transgenic zebrafish that expressed the photoconvertible fluorescent protein Kaede in Hcrt neurons and EGFP in dopamine beta-hydroxylase (dbh)-expressing LC neurons (supplementary material Fig. S1). We used a 405 nm laser to photoconvert Hcrt neurons in the hindbrain, but not in the hypothalamus, from green to red fluorescence. We observed that all red fluorescent ectopic Hcrt neurons project to the LC (15/15 neurons) (Fig. 5A,B). By contrast, stochastic labeling of neurons in the medial hindbrain with an elav3:Kaede transgene indicates that only ~20% of cells in this region target the LC (9/50 neurons) (Fig. 5B,C). These experiments indicate that Lhx9 is sufficient to

<table>
<thead>
<tr>
<th>Table 1. Transcription factors enriched in Hcrt neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>hmx2</td>
</tr>
<tr>
<td>hmx3</td>
</tr>
<tr>
<td>ngas4</td>
</tr>
<tr>
<td>lhx9</td>
</tr>
<tr>
<td>sox1a</td>
</tr>
<tr>
<td>nr4a2</td>
</tr>
<tr>
<td>mybpb1a</td>
</tr>
<tr>
<td>tshz2</td>
</tr>
<tr>
<td>bhle40</td>
</tr>
<tr>
<td>tead1</td>
</tr>
<tr>
<td>cited2</td>
</tr>
<tr>
<td>ets2</td>
</tr>
<tr>
<td>cbx4</td>
</tr>
</tbody>
</table>

*Fold increase indicates the expression level of a gene in Hcrt neurons relative to its expression in the other neuronal subtypes analyzed by microarray.

<table>
<thead>
<tr>
<th>Table 2. Secreted proteins enriched in Hcrt neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>hct</td>
</tr>
<tr>
<td>cart4</td>
</tr>
<tr>
<td>igsf21</td>
</tr>
<tr>
<td>trh</td>
</tr>
<tr>
<td>endouc</td>
</tr>
<tr>
<td>glpr1b</td>
</tr>
</tbody>
</table>

*Fold increase indicates the expression level of a gene in Hcrt neurons relative to its expression in the other neuronal subtypes analyzed by microarray.
specify Hcrt neurons in vivo that are genetically and anatomically similar to endogenous Hcrt neurons.

Lhx9 is necessary for Hcrt neuron specification

To test whether Lhx9 is necessary for Hcrt neuron specification, we used a splice-blocking morpholino to knock down lhx9 expression. RT-PCR revealed robust inhibition of lhx9 mRNA splicing, and thus of functional Lhx9 protein production, up to 72 hpf (supplementary material Fig. S10). Morphants had ∼40% fewer Hcrt neurons when assessed by ISH (Fig. 6C), and the remaining Hcrt neurons had reduced hcart expression compared with embryos injected with a 5 bp mismatch control morpholino (Fig. 6A,B). qRT-PCR revealed that lhx9 morphants express 62% less hcart transcript than control morphants (s.e.m.=5%, n=3 replicates), confirming that Hcrt neurons that persist in lhx9 morphants contain less hcart transcript than controls.

To determine whether the missing Hcrt neurons in lhx9 morphants lack hcart expression or were absent, we examined the effect of lhx9 knockdown on vglut1 and pdyn (supplementary material Fig. S11). lhx9 morphants had no gross defects in vglut1 expression compared with controls. However, the number of cells with intense punctate pdyn expression, which includes all Hcrt cells, was reduced by ∼40%. The loss of Hcrt cells was not caused by nonspecific morpholino-induced apoptosis, as embryos co-injected with an apoptosis-suppressing p53 (tp53) morpholino (Robu et al., 2007) and the lhx9 morpholino showed the same phenotype (Fig. 6C). Furthermore, lhx9 morpholino-injected embryos stained with Acridine Orange, which labels apoptotic cells, showed no increase in apoptosis (supplementary material Fig. S12A-E). The morpholino phenotype was partially rescued by co-injecting the lhx9 morpholino with the hs:lhx9 plasmid and performing a heat shock at 24 hpf (supplementary material Fig. S12F,G), indicating that the reduction in Hcrt cells was a specific effect of lhx9 knockdown. hcart expression was weaker in rescued cells than in endogenous Hcrt neurons, presumably because rescued cells only received a pulse of lhx9...
whereas endogenous Hcrt neurons continuously express \textit{lhx9} (supplementary material Fig. S4). We also tested morpholinos against \textit{hmx2} or \textit{hmx3}, which were more highly enriched in Hcrt neurons than \textit{lhx9} in our microarray analysis. Although RT-PCR confirmed that these morpholinos were effective, they had no effect on Hcrt neuron specification (data not shown).

Some Hcrt neurons persisted in \textit{lhx9} morphants, possibly owing to incomplete \textit{lhx9} knockdown (supplementary material Fig. S10) or because Lhx9 is only necessary to specify a subset of Hcrt neurons, as in mice (Dalal et al., 2013). To distinguish between these possibilities, we used the CRISPR/Cas9 system (Jao et al., 2013; Hwang et al., 2013) to introduce mutations into \textit{lhx9}. We co-injected Cas9 protein (Gagnon et al., 2014) with a set of short guide RNAs (sgRNAs) that target \textit{lhx9} into embryos at the 1-cell stage to generate biallelic mutations and a loss-of-function phenotype in the injected animals (Jao et al., 2013). Embryos injected with Cas9+\textit{lhx9} sgRNAs had 90% fewer Hcrt cells than embryos injected with Cas9 alone (Fig. 6D-F). Furthermore, over half of brain hemispheres of embryos injected with Cas9+\textit{lhx9} sgRNAs completely lacked Hcrt cells, whereas embryos injected with Cas9 alone had at least three Hcrt cells in each brain hemisphere (Fig. 6G). This phenotype is unlikely to be due to off-target effects of particular sgRNAs, as we observed a similar, albeit weaker, phenotype in embryos injected with Cas9 and independent subsets of \textit{lhx9} sgRNAs (supplementary material Fig. S13). We conclude that \textit{lhx9} is necessary to specify all Hcrt neurons in zebrafish embryos.

Because \textit{lhx9} overexpression was sufficient to specify QRFP neurons, we asked whether \textit{lhx9} is also necessary for QRFP neuron specification. We observed 63% fewer QRFP neurons in embryos injected with Cas9+\textit{lhx9} sgRNAs compared with Cas9 alone, and over 20% of brain hemispheres lacked QRFP neurons (supplementary material Fig. S13). Injection of embryos with Cas9 and subsets of \textit{lhx9} sgRNAs produced a similar, albeit weaker, phenotype. We conclude that \textit{lhx9} is necessary to specify all QRFP neurons.

\textbf{Lhx9 directly promotes hcret expression}

As Lhx9 is a transcription factor, we hypothesized that it might promote \textit{hcrt} expression directly. A previous study tested this...
A previous study identified an Lhx9 binding site upstream of the hcxrt promoter-luciferase reporter in a neuroblastoma cell line, and in vivo by lentiviral transduction of Lhx9 into the hypothalamus of adult mice, but observed no effect on hcxrt expression (Dalal et al., 2013). However, when we performed double-fluorescent ISH against lsx9 and hcxrt on zebrafish embryos fixed 1 h after heat shock, we observed hcxrt expression in almost all lsx9-overexpressing cells (Fig. 7A; supplementary material Fig. S14 and Movie 4). The number of ectopic lsx9- and hcxrt-expressing cells is reduced at 8 h after heat shock (supplementary material Fig. S14D), and few ectopic cells are observed 24 h after heat shock (supplementary material Fig. S14F), similar to the number observed at 120 hpf (supplementary material Fig. S14H). Importantly, ectopic hcxrt-expressing neurons also express ectopic lsx9 at all time points. These findings support the hypothesis that Lhx9 can directly regulate hcxrt expression.

A previous study identified an Lhx9 binding site upstream of the Wilms tumor 1 gene (Wilhelm and Englert, 2002), and we observed two similar sites in the zebrafish hcxrt promoter (sites A and B) (Fig. 7B,C). Notably, site A corresponds to a region previously shown to be important for hcxrt expression in zebrafish (Faraco et al., 2013). We tested whether these sites are necessary for endogenous (i.e. hypothalamic) and Lhx9-induced ectopic hcxrt expression by injecting wild-type embryos with plasmids containing the hcxrt promoter, in which one or both putative Lhx9 binding sites were mutated, placed upstream of EGFP. Each plasmid also contained a heat shock-inducible lhx9 transgene downstream of the EGFP reporter. Thus, any cell that contains the plasmid will have both the hcxrt:EGFP reporter and the lhx9:lsx9 transgene. Injected embryos were heat-shocked at 24 hpf and analyzed at 120 hpf for hypothalamic and ectopic hcxrt:EGFP-expressing neurons (Fig. 7D, HS samples). Some injected embryos were not heat-shocked (Fig. 7D, no HS samples) to determine whether the putative Lhx9 binding sites are required for EGFP expression in the endogenous Hcxrt domain alone. Mutating site A, either by deletion or by scrambling every third nucleotide, reduced the number and intensity of endogenous and ectopic cells labeled with EGFP as compared with the wild-type hcxrt promoter (Fig. 7C,D,F-H). Scrambling the sequences of both sites A and B virtually abolished EGFP expression (Fig. 7D,I). Notably, ectopic hcxrt cells were never observed for the double-mutant reporter. These experiments indicate that sites A and B are crucial for hcxrt expression in vivo.

To test whether Lhx9 can interact with these sites, we performed an EMSA using the zebrafish Lhx9 homeodomain (Lhx9 HD) and radiolabeled oligonucleotides that include the wild-type or scrambled sequences for sites A and B. We found that Lhx9 HD binds to the wild-type site A probe, but not to the scrambled site A probe (Fig. 7J). We failed to observe an interaction between Lhx9 HD and the site B probe (data not shown), possibly owing to non-optimal in vitro binding or electrophoresis conditions. This result indicates that Lhx9 can bind to site A in vitro and suggests that Lhx9 directly regulates hcxrt expression in vivo via binding to this site.

**Lhx9 overexpression in mouse embryos induces Hcxrt neuron specification**

A previous study that overexpressed Lhx9 in the hypothalamus of adult mice observed no effect on Hcxrt neuron specification (Dalal et al., 2013). To determine whether Lhx9 can promote Hcxrt neuron specification earlier in mammalian development, we used micro in utero electroporation (Matsumoto et al., 2011) to focally overexpress EYFP and the murine Lhx9 ortholog, or EYFP alone, in the developing murine diencephalon at embryonic day (E) 10.5 and assayed Hcxrt expression at postnatal day (P) 6. Embryos overexpressing Lhx9 had significantly more Hcxrt-expressing neurons in the lateral hypothalamus than embryos overexpressing EYFP alone (Fig. 8; supplementary material Figs S15 and S16). This effect appears to be specific, since Lhx9 overexpression had no effect on the expression of other hypothalamic markers, including Carpt, Foxp2 and Gal (data not shown). Lhx9 overexpression in the subthalamic zona incerta (supplementary material Fig. S17) or cerebral cortex (data not shown) did not induce ectopic Hcxrt expression. We conclude that Lhx9 is capable of promoting Hcxrt neuron specification exclusively in its endogenous domain during mouse embryogenesis.

**DISCUSSION**

Using microarray gene expression analysis and high-throughput gene overexpression assays in zebrafish, we found that the LIM homeobox transcription factor Lhx9 is both necessary and sufficient to specify Hcxrt neurons in zebrafish and is sufficient to specify Hcxrt neurons in mouse embryos. We found that Lhx9 is also necessary and sufficient to specify QRFP neurons in zebrafish, which are located adjacent to Hcxrt neurons in the hypothalamus. To our knowledge, this is the first study to identify a factor that is capable of inducing the specification of these neurons, or of any terminal neural subtype in the lateral hypothalamus.
lhx9 was identified as enriched in Hcrt neurons by both our analysis of zebrafish embryos and by a previous study that used adult mice (Dalal et al., 2013). We analyzed zebrafish neurons just after the initiation of hclrt expression, enabling us to screen for transcripts that are likely to play a role in the specification of Hcrt neurons. We isolated purified cell populations by FACS, analyzed their gene expression patterns by microarray, and performed multiple pairwise comparisons of purified Hcrt neurons with a closely related cell type (QRFP-expressing neuron) and to more distantly related neurons (pan-neuronal or sensory neurons). By contrast, Dalal and colleagues used a translational profiling approach in which a tagged ribosomal subunit is expressed in Hcrt neurons. Biochemical purification of this subunit from a whole brain homogenate isolates transcripts that are actively translated in Hcrt neurons. Although this approach allowed profiling of the rare Hcrt cell population, its statistical power was diminished by contamination with nonspecific transcripts, as indicated by the presence of glial transcripts. In addition to lhx9, the 112 genes most enriched in Hcrt neurons in our study and the 188 most enriched genes identified by Dalal and colleagues include the definitive Hcrt neuron markers hclrt and pdyn, the transcription factors hmox2 and rfx4, and scg2, agrp, glipr1 and fam46a. The absence of more overlapping genes can be attributed in part to ambiguity in the gene assignment of microarray probes and the imperfect annotation of the zebrafish genome. Furthermore, the stringent criteria for significance used by both studies are likely to underestimate the true complement of Hcrt enriched genes shared between zebrafish embryos and adult mice.

Lhx9 belongs to the LIM homeobox family of transcription factors, which is conserved from invertebrates to mammals. These proteins have essential roles in tissue patterning and differentiation, particularly in the brain (Hobert and Westphal, 2000). In mice, several LIM homeobox proteins are expressed dynamically to demarcate regions of the developing hypothalamus (Shimogori et al., 2010). However, the developmental roles of specific LIM homeobox genes have been difficult to distinguish; loss-of-function phenotypes are subtle and similar LIM homeobox family members often exhibit redundancy. For instance, double knockdown of lhx9 and lhx2 dramatically altered thalamus and forebrain patterning in zebrafish, but knockdown of either gene alone had no gross effects (Peukert et al., 2011). Similarly, Lhx9 knockout mice survive to adulthood without gross brain defects (Birk et al., 2000).

Although we tested morpholinos against several candidate genes, only the lhx9 morpholino decreased the number of Hcrt neurons, with an average decrease of 40%. The remaining Hcrt neurons, which expressed hclrt at reduced levels, are likely to result from incomplete lhx9 knockdown. Indeed, co-injection of Cas9 protein with sgRNAs targeting lhx9 completely abolished hclrt expression in over half of brain hemispheres analyzed, indicating that lhx9 is required for the specification of all Hcrt neurons. This result contrasts with Lhx9 knockout mice, in which Hcrt neurons are only reduced by 39% (Dalal et al., 2013). This discrepancy is likely to be due to the expression of lhx9 in all Hcrt neurons in zebrafish but in only a subset in mice (Shimogori et al., 2010).

In zebrafish, lhx9 overexpression at 24 hpf was sufficient to produce ectopic Hcrt neurons in the medial hindbrain, but the number of Hcrt neurons in the hypothalamus remained unchanged. We characterized these ectopic Hcrt neurons at 120 hpf, when the Hcrt neuronal circuit is functional (Prober et al., 2006; Elbaz et al., 2012), and confirmed that all zebrafish Hcrt neurons in the hypothalamus and hindbrain express pdyn and vglut1, two markers of mammalian Hcrt neurons (Chou et al., 2001; Rosin et al., 2003). Unlike previous studies (Appelbaum et al., 2009; Rosin et al., 2003), we did not observe expression of vglut2a or vglut2b in hypothalamic or hindbrain Hcrt neurons, indicating that Hcrt neurons in larval zebrafish express different VGLUT family genes than adult zebrafish or rats. We also observed that all ectopic Hcrt neurons project to the LC, a target of Hcrt neurons in zebrafish and mammals (Horvath et al., 1999; Prober et al., 2006). Thus,
Despite their location in the hindbrain, Lhx9-induced ectopic Hcrt neurons express the same genetic markers and project to the same target as endogenous Hcrt neurons.

Although it might be surprising that we did not detect an effect of lhx9 overexpression on the number of hypothalamic Hcrt neurons, there are several possible explanations for this result. First, because lhx9 induces few ectopic Hcrt neurons in our assay and the number of endogenous Hcrt neurons is variable, it is possible that lhx9 induces additional Hcrt neurons in the hypothalamus but the difference was not large enough to detect. Second, endogenous lhx9 is likely to act in concert with other factors to specify Hcrt neurons because only a subset of endogenous lhx9-expressing cells expresses hcrt. If these co-factors are only present in hypothalamic neurons that express hcrt, lhx9 overexpression will have no effect on the number of Hcrt neurons in the hypothalamus. Other genes identified by our microarray analysis might encode these co-factors. However, overexpressing two or more candidate genes using our assay was not feasible due to DNA toxicity. Reducing the concentration of each injected plasmid to offset this toxicity also reduces the extent of gene overexpression. We were thus unable to observe ectopic Hcrt neurons when the hs:lhx9 plasmid was co-injected with a second plasmid, and the number of hypothalamic Hcrt neurons was unaffected (data not shown).

We detected widespread co-expression of lhx9 and hcrt 1 h after heat shock-induced lhx9 overexpression. Because ectopic hcrt expression was observed so soon after heat shock, it seemed likely that Lhx9 was directly inducing hcrt expression. Consistent with this hypothesis, we identified two putative Lhx9 binding sites in the zebrafish hcrt promoter and found that they are necessary for both endogenous and ectopic hcrt expression in vivo. We also found that one binding site can form a complex with the Lhx9 homeodomain.

Because the extent of both lhx9 overexpression and of ectopic hcrt expression is dramatically reduced by 24 h after heat shock (supplementary material Fig. S14), we propose that the ability of Lhx9 to drive hcrt expression may depend on Lhx9 levels. High Lhx9 levels might directly induce hcrt expression, whereas endogenous Lhx9 levels require one or more co-factors in the hypothalamus and medial hindbrain to promote and maintain hcrt expression.

Since the expression patterns of hcrt and lhx9 are conserved between zebrafish and mammals (Peukert et al., 2011; Shimogori et al., 2010), we tested whether Lhx9 overexpression could induce Hcrt neuron specification in mice. Indeed, Lhx9 overexpression by
Hcrt expression after viral transduction of an Hcrt-overexpression construct in the hypothalamus of adult mice suggests that the zone of cells competent to specify Hcrt neurons is more spatially restricted in mice than in zebrafish.

However, electroporation was sufficient to specify additional Hcrt neurons, although they were only observed in the endogenous Hcrt neuron domain. This result suggests that the zone of cells competent to specify Hcrt neurons is more spatially restricted in mice than in zebrafish. Our findings differ from a previous study that saw no change in Hcrt expression after viral transduction of an Lhx9-overexpression construct in the hypothalamus of adult mice (Dalal et al., 2013). These results suggest that Lhx9 can induce Hcrt neuron specification in the embryonic, but not adult, mouse hypothalamus, possibly because cells competent to specify Hcrt neurons are fully differentiated in adults.

Our study demonstrates the utility of zebrafish to identify and test genes that regulate vertebrate development. Furthermore, the ability of Lhx9 to induce Hcrt neuron specification suggests a therapeutic approach to compensate for the loss of Hcrt neurons that is thought to cause narcolepsy. This strategy would use Lhx9 to generate HCRT-expressing neurons from human pluripotent stem cells in vitro, followed by screening and selection of Hcrt neurons to be transplanted into the hypothalamus. The promise of this approach is highlighted by the recent demonstration that narcolepsy-like sleep induced by the lesion of Hcrt neurons in rats is diminished by the transplantation of Hcrt neurons into the lateral hypothalamus (Arias-Carrión and Murillo-Rodríguez, 2014).

**MATERIALS AND METHODS**

**Ethics statement**

Zebrafish experiments followed standard protocols (Westerfield, 1993) in accordance with Caltech Institutional Animal Care and Use Committee guidelines. Mouse procedures were approved by the RIKEN Institutional Animal Care Committee.

**Transgenic zebrafish**

A 1-kb fragment of zebrafish genomic DNA upstream of qrfp was cloned upstream of EGFP by using primers 5′-CTGACTCTCCATCAGTCT-3′ and 5′-CTGAAATTTAAGGAATAATTTAAAGTTG-3′. A 1-kb zebrafish genomic fragment upstream of hctr (Faraco et al., 2006) was subcloned upstream of mRFP and Kaede using primers 5′-ATAATAAAATACGTAGGGTTTT-3′ and 5′-GAGTGTACCTGTCCCTGTT-3′. The dhh:EGFP transgene was generated by cloning a 1.1-kb fragment upstream of zebrafish dhh using primers 5′-ACTTGACCGACCTTCT-3′ and 5′-GTTTGAAGGCTTCTAAGTTTT-3′. Transgenes were co-injected with tol2 transposase mRNA to generate stable transgenic lines. The Tg(hcrt:EGFP), Tg(isl1:Gal4VP16, UAS:EGFP), Tg(p2r3b:EGFP) and Tg(vglut2a:RFP) lines have been described (Prober et al., 2006; Park et al., 2000; Pan et al., 2012; Sagasti et al., 2005; Kucenas et al., 2006; Koyama et al., 2011). Hindbrain neurons were stochastically labeled by injecting a Tg(elavl3:Kaede) transgene (Sato et al., 2006).

**Microarray analysis**

We analyzed embryos co-expressing qrfp:EGFP and hctr:mRFP, as well as separate transgenic lines expressing EGFP in all neurons [Tg(isl1:Gal4VP16)], or in subsets of sensory neurons [Tg(topalb:EGFP), Tg(isl1:Gal4VP16, UAS:EGFP), Tg(p2r3b:EGFP)]. Dechorionated embryos were dissected in calcium-free tricine at 26 °C, incubated in stress conditions at 25 °C under isometric conditions for 1 h and fixed in 4% paraformaldehyde for 12-16 h at room temperature. ISH was performed using digoxigenin (DIG)-labeled antisense riboprobes (Thisse and Thisse, 2008). Images were acquired with a Zeiss Axiovert 200 M microscope equipped with Axioscam MRc5 camera.

Samples were fixed in 4% paraformaldehyde for 12-16 h at room temperature. ISH was performed using digoxigenin (DIG)-labeled antisense riboprobes (Thiis and Thisse, 2008). Images were acquired using a Zeiss Axioskop microscope.
using a Zeiss Axio ImagerM1 microscope. Fluorescent ISH used DIG- and 2,4-dinitrophenol (DNP)-labeled antisense riboprobes with the TSA Plus DNP System (PerkinElmer). Rabbit polyclonal anti-GFP (1:1000; MBL International, #598), rabbit polyclonal anti-orexin-A (Hcrt) (1:1000; Millipore, #AB3704) and goat anti-rabbit Alexa488 (1:500; Invitrogen, #A-11008) antibodies were used. Images were acquired using a Zeiss LSM 780 confocal microscope and analyzed using Fiji (Schindelin et al., 2012).

Candidate gene overexpression
The coding sequence of each candidate gene was amplified from 24 hpf zebrafish cDNA and cloned into pENTR-D-TOPO (Invitrogen). Gateway recombination (Invitrogen) was used to clone each gene downstream of a heat shock-inducible promoter (Halleran et al., 2000), and the entire cassette was flanked by Tol2 transposase sites. We co-injected individual overexpression plasmids with tol2 transposase mRNA into zebrafish embryos at the 1-cell stage. Gene overexpression was induced by incubating embryos in a 37°C water bath for 1 h.

Morpholino-mediated knockdown
Morpholinos (GeneTools) were injected into wild-type embryos at the 1-cell stage. We used a splice-blocking morpholino to knock down Lhx9 (5′-AGGCCTCAAGTATTACCTTCTG-3′). A morpholino with a 5 bp mismatch to the target was injected as a negative control (5′-AGGCGTAACACTTACCTTCTG-3′). Potential apoptosis was suppressed by co-injecting a p53 morpholino (5′-GCGCATCAAGACTTTTCGATACA-3′). To verify knockdown efficacy, we isolated RNA from pools of five injected embryos and used RT-PCR (Superscript III First-Strand Synthesis System, Invitrogen) to amplify a fragment of the lhx9 transcript that spans exon 2. To detect apoptosis, 24-hpf embryos were bathed in 1 μg/ml Acridine Orange for 1 h at room temperature, followed by three 10 min washes with E3 medium. Splice-blocking morpholinos were designed for lmx2 (5′-TGGGACCTGACTCAGGACAGCAGA-3′) and lmx3 (5′-TCGCTGCTCAATGAGGCCAAA-3′) to retain the first intron of each gene, resulting in an early stop codon.

Quantitative reverse-transcription PCR (qRT-PCR)
We isolated total RNA from three biological replicates (25 embryos each) of lhx9 morpholino-injected and control morpholino-injected embryos. TURBO DNase I was used to remove genomic DNA (TURBO DNA-free Kit, Invitrogen). We then generated cDNA (Superscript III First-Strand Synthesis System, Invitrogen) and amplified hert transcripts with primers 5′-ATGGAACTCAATCTACCGGAGCAGGAC-3′ and 5′-AGGCTCAAAAGTTAATGCTTACCTGT-3′. Transcripts of the rpl13a reference gene were amplified (Superscript III First-Strand Synthesis System, Invitrogen) to amplify a fragment of the rpl13a transcript that spans exon 2. To detect apoptosis, 24-hpf embryos were bathed in 1 μg/ml Acridine Orange for 1 h at room temperature, followed by three 10 min washes with E3 medium. Splice-blocking morpholinos were designed for lmx2 (5′-TGGGACCTGACTCAGGACAGCAGA-3′) and lmx3 (5′-TCGCTGCTCAATGAGGCCAAA-3′) to retain the first intron of each gene, resulting in an early stop codon.

CRISPR/Cas9
Ten target sites within the lhx9 open reading frame were chosen using CHOPCHOP (Montague et al., 2014) (supplementary material Table S2). Cas9 protein was mixed with all ten sgRNAs and injected into embryos at the 1-cell stage (Gagnon et al., 2014). At 24 hpf, deformed embryos were removed and the remainder were fixed in 4% paraformaldehyde for ISH. To control for potential sgRNA off-target effects, we co-injected Cas9 protein with independent subsets of the sgRNAs (subset 1 comprised sgRNAs 1, 3, 5, 7, 9; subset 2 comprised sgRNAs 2, 4, 6, 8, 10) (supplementary material Table S2).

hert enhancer constructs
We mutated one or both putative Lhx9 binding sites in the zebrafish hert promoter by PCR and Gibson assembly (Gibson et al., 2009). Enhancer fragments were placed upstream of the EGFP coding sequence. An hs:lhx9 cassette was placed downstream of hert:EGFP in a vector containing Tol2 transposase sites.

Electrophoretic mobility shift assay (EMSAs)
The following oligonucleotide probes (5′-3′) were used: site A wild-type probe, GTTGGTTATTGATGCAGGCTGTGCTCTGTTCA; site A scramble probe, GTTGGTATTGATGCAGGCTGTGCTCTGTTCA; site B wild-type probe, TGACAAAGATGCTAAACCCGGAAAAATTCTTGTG; site B scramble probe, TGACAAAGATGCTAAACCCGGAAAAATTCTTGTG. Probes were radiolabeled using [γ-32P]ATP and T4 polynucleotide kinase for 1 h at 37°C and column-purified (Illustra Microspin G-50, GE Healthcare). EMSAs were performed using a truncated form of zebrafish Lhx9 (Lhx9 HD) that contains the DNA-binding homeodomain but lacks both LIM domains (amino acids 224-396 of the 396 amino acid protein), as described (Wilhelm and Englert, 2002). Lhx9 HD was synthesized in vitro (TNT SP6 Quick Coupled Transcription/Translation System, Promega). 1 μl normalized radiolabeled oligo and 4.5 μl TNT lysate were added to a final volume of 30 μl binding buffer, containing 100 mM KCl, 1 mM MgCl2, 10 μM ZnSO4, 10 mM Tris pH 7.5, 4% glycerol, 1 mg/ml BSA, 200 ng poly(dIdC) and 0.5 mM DTT. After 1 h at room temperature, DNA-protein complexes were resolved by electrophoresis at 4°C on a 6% polyacrylamide DNA retardation gel (Invitrogen) at 100 V for 90 min in 0.5× TBE buffer. The gel was dried at 60°C for 2 h using a gel dryer (Bio-Rad) and analyzed by phosphorimaging (GE Healthcare).

Mouse experiments
Outbred ICR (CD-1) timed-pregnant mice were obtained from Japan SLC. Midday of the day of vaginal plug discovery was considered embryonic (E) day 0.5. Early postnatal mice were anesthetized with a lethal dose of pentobarbital (100 mg/kg) and, after three failed attempts to elicit a foot withdrawal reflex, the animals were transcardially perfused with 4% paraformaldehyde in PBS. For ISH, brains were fixed overnight in 30% sucrose/4% paraformaldehyde and sectioned in the coronal plane on a Leica sledge microtome at 28 μm. Sections were mounted on slides and processed for non-radioactive ISH as described (Grove et al., 1998). DNA for riboprobes and electroporation were obtained from FANTOM clones (Carninci et al., 2005). Mouse Lhx9 (GenBank NM_001025565) was subcloned into a CAG vector as described (Onishi et al., 2010) to generate an overexpression plasmid. In utero electroporation was performed as described (Matsui et al., 2011).

Acknowledgements
We thank Farhad Imam for assistance with microarray data analysis, Anthony Kirilushka for EMSA assistance and Daniel Lee for comments on the manuscript.

Competing interests
The authors declare no competing or financial interests.

Author contributions
D.A.P. and A.F.S. conceived and directed the project; J.L., F.T.M., J.A.G., T.S. and D.A.P. performed experiments; A.V.G. and C.N.C. generated and maintained strains; R.S. generated sgRNAs; L.A. performed microarray data analysis; T.S. performed microarray data analysis; A.F.S. and D.A.P. wrote the paper.

Funding
This work was supported by grants from the National Institutes of Health [F31NS77844 to J.L.; 1R21NS071598, 5K99NS083713 to F.T.M.; R01HL109525, R21NS071598 to A.F.S.; R00NS060996, R01NS070911, R01DA031367 to D.A.P.], the American Cancer Society [to J.A.G. and I.G.W.], the Jane Coffin Childs Memorial Fund for Medical Research [to F.T.M.], the RIKEN Brain Institute [to T.S.], the Mllincroft Foundation, the Rita Allen Foundation, and the Brain and Behavior Research Foundation [to D.A.P.]. Deposited in PMC for release after 12 months.

Supplementary material
Supplementary material available online at http://dev.biologists.org/lookup/doi/10.1242/dev.117424/-/DC1

References


