

# UCSF

## UC San Francisco Previously Published Works

### Title

Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic Lhx6-expressing GABAergic neurons.

### Permalink

<https://escholarship.org/uc/item/38x5q4hb>

### Journal

Communications biology, 4(1)

### ISSN

2399-3642

### Authors

Kim, Dong Won  
Liu, Kai  
Wang, Zoe Qianyi  
et al.

### Publication Date





2021

### DOI


10.1038/s42003-020-01616-7

Peer reviewed

## Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic Lhx6-expressing GABAergic neurons

Dong Won Kim<sup>1</sup>, Kai Liu<sup>1,10</sup>, Zoe Qianyi Wang<sup>1</sup>, Yi Stephanie Zhang<sup>1</sup>, Abhijith Bathini<sup>1</sup>, Matthew P. Brown<sup>1</sup> , Sonia Hao Lin<sup>1</sup>, Parris Whitney Washington<sup>1</sup>, Changyu Sun<sup>1</sup>, Susan Lindtner<sup>2</sup>, Bora Lee<sup>3</sup> , Hong Wang<sup>1</sup>, Tomomi Shimogori<sup>4</sup> , John L. R. Rubenstein<sup>2</sup> & Seth Blackshaw<sup>1,5,6,7,8,9</sup> 

GABAergic neurons of the hypothalamus regulate many innate behaviors, but little is known about the mechanisms that control their development. We previously identified hypothalamic neurons that express the LIM homeodomain transcription factor *Lhx6*, a master regulator of cortical interneuron development, as sleep-promoting. In contrast to telencephalic interneurons, hypothalamic *Lhx6* neurons do not undergo long-distance tangential migration and do not express cortical interneuronal markers such as *Pvalb*. Here, we show that *Lhx6* is necessary for the survival of hypothalamic neurons. *Dlx1/2*, *Nkx2-2*, and *Nkx2-1* are each required for specification of spatially distinct subsets of hypothalamic *Lhx6* neurons, and that *Nkx2-2*/*Lhx6* neurons of the zona incerta are responsive to sleep pressure. We further identify multiple neuropeptides that are enriched in spatially segregated subsets of hypothalamic *Lhx6* neurons, and that are distinct from those seen in cortical neurons. These findings identify common and divergent molecular mechanisms by which *Lhx6* controls the development of GABAergic neurons in the hypothalamus.

<sup>1</sup>Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>2</sup>Nina Ireland Laboratory of Developmental Neurobiology, Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, CA 94158, USA. <sup>3</sup>Center for Neuroscience, Korea Institute of Science and Technology (KIST), Seoul 02792, Korea. <sup>4</sup>RIKEN Center for Brain Science, Laboratory for Molecular Mechanisms of Brain Development, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. <sup>5</sup>Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>6</sup>Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>7</sup>Center for Human Systems Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>8</sup>Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>9</sup>Kavli Neuroscience Discovery Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>10</sup>Present address: Genentech, South San Francisco, CA 94080, USA. email: [sblack@jhmi.edu](mailto:sblack@jhmi.edu)

Although much is now known about both the diversity and development of GABAergic neurons of the telencephalon<sup>1,2</sup>, far less is known about their counterparts in the hypothalamus, where over 20% of neurons are GABAergic<sup>3</sup>. Previous work shows that hypothalamic GABAergic neuronal precursors first appear in a domain that separates the anterodorsal and posteroventral halves of the developing hypothalamus, and is delineated by expression of transcription factors that regulate the development of telencephalic GABAergic neurons, including *Dlx1/2* and *Arx*<sup>4–8</sup>. Within this structure, which has been termed the intrahypothalamic diagonal/tuberomammillary terminal (ID/TT), nested expression domains of LIM homeodomain family genes are observed, in which expression of *Lhx1*, *Lhx8*, and *Lhx6* delineates the anterior–posterior axis of the ID/TT<sup>4</sup>. *Lhx1* is essential for the terminal differentiation and function of neurons in the master circadian oscillator in the suprachiasmatic nucleus<sup>9–11</sup>. *Lhx6*-expressing neurons in the zona incerta (ZI) of the hypothalamus are sleep-promoting and activated by elevated sleep pressure, and hypothalamic-specific loss of function of *Lhx6* disrupts sleep homeostasis<sup>12</sup>.

*Lhx6* has been extensively studied in the developing telencephalon. It is essential for the specification, migration, and maturation of GABAergic neurons of the telencephalon—particularly the cortex and hippocampus<sup>13,14</sup>. *Lhx6* is expressed in the medial ganglionic eminence (MGE) of the embryonic telencephalon, where it is co-expressed with both *Nkx2-1* and *Dlx1/2*<sup>15–17</sup>. *Shh* induces expression of *Nkx2-1*<sup>18</sup>, which in turn directly activates *Lhx6* expression<sup>16,19</sup>. *Nkx2-1*, in turn, cooperates with *Lhx6* to directly activate the expression of multiple other genes that control cortical interneuron specification and differentiation, including *Sox6* and *Gsx2*<sup>20,21</sup>. Furthermore, *Lhx6* is both necessary and sufficient for the tangential migration of the great majority of interneuron precursors from the MGE to their final destinations in the cortex and hippocampus<sup>17,22,23</sup>. Finally, *Lhx6* expression persists in mature interneurons that express parvalbumin (*Pvalb*) and somatostatin (*Sst*), and is necessary for their expression<sup>22</sup>.

The functional role of *Lhx6* in hypothalamic development has not been previously investigated. However, previous studies imply this may differ in certain key ways from its function in the developing telencephalon. Notably, the hypothalamic domain of *Lhx6* expression only partially overlaps with that of *Nkx2-1*<sup>4</sup>. Furthermore, in sharp contrast to cortical interneurons, *Lhx6* is not co-expressed with either *Pvalb* or *Sst* in the ZI<sup>12</sup>. In this study, we sought to determine the extent to which gene regulatory networks controlling the development of hypothalamic *Lhx6* neurons diverge from those that control the development of telencephalic *Lhx6* neurons. We find that hypothalamic *Lhx6* regulates neuronal differentiation and survival. Further, we observe extensive molecular heterogeneity among mature hypothalamic *Lhx6* neurons and a lack of overlap with annotated subtypes of *Lhx6*-expressing cortical interneurons. Combinatorial patterns of transcription factor expression delineate spatial subdomains of *Lhx6* expression within the ID/TT, and we find that *Nkx2-1*, *Nkx2-2*, and *Dlx1/2* each regulate expression of *Lhx6* in largely nonoverlapping domains. Finally, *Lhx6* neurons derived from *Nkx2-2*-expressing precursors are activated by sleep pressure. These findings identify mechanisms by which *Lhx6* can regulate the development of hypothalamic GABAergic neurons, and more broadly, how diverse subtypes of hypothalamic neurons can be generated during development.

## Results

**Distribution of hypothalamic *Lhx6*-expressing neurons.** Our previous work has indicated that *Lhx6* is expressed in two

continuous yet distinct domains of the developing hypothalamus: the intrahypothalamic diagonal (ID) and the more posterior tuberomammillary terminal (TT)<sup>4,5</sup>. We next sought to more carefully determine the expression pattern of *Lhx6* and its putative regulators during early hypothalamic development. High-quality chromogenic *in situ* hybridization (ISH) detects both the ID and TT domain of *Lhx6* expression at E11.5, E12.5, and E14.5 (Fig. 1A–E). By E16.5, hypothalamic *Lhx6*-expressing neurons are observed in the ZI and dorsomedial hypothalamus (DMH), in a pattern that broadly corresponds to the earlier ID domain, while expression in the posterior hypothalamus (PH) in turn broadly corresponds to the TT domain (Fig. 1F, G). This closely matches the pattern of hypothalamic *Lhx6* expression previously reported in adults<sup>12</sup>. *Lhx6*-expressing neurons are only a small minority of hypothalamic GABAergic neurons<sup>12</sup>, with single-cell RNA-sequencing (scRNA-Seq) revealing that only ~2% of all hypothalamic GABAergic neuronal precursors (defined by *Gad1/2* and *Dlx1/2* expression) express *Lhx6* between E11 and E13 (Fig. 1H)<sup>24</sup>.

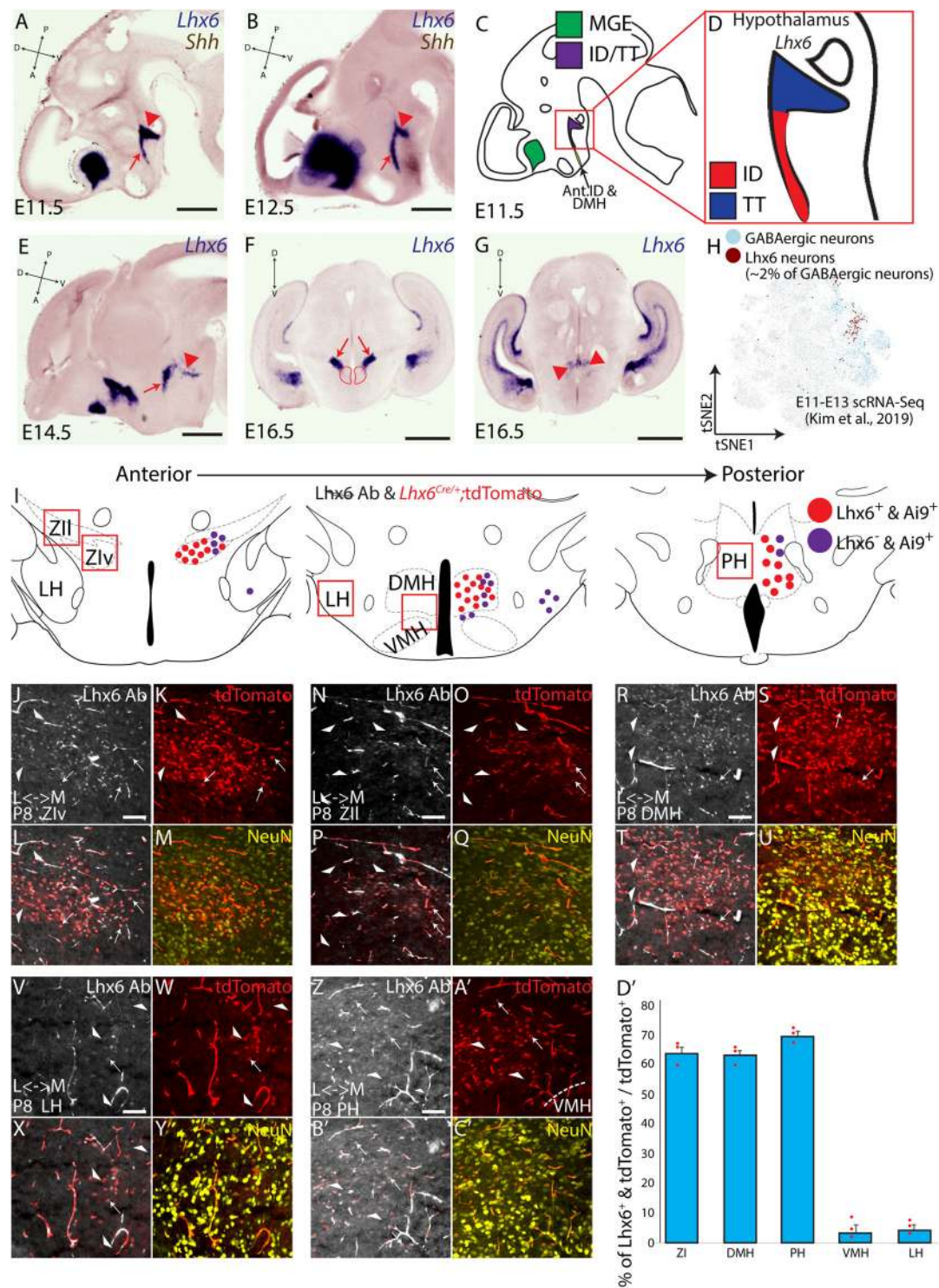
This regional pattern of hypothalamic *Lhx6* expression is broadly similar to that reported for *Lhx6*<sup>Cre/+</sup>;*Ai9* mice (Fig. 1J–D')<sup>12</sup>, with ~65–70% of tdTomato-expressing neurons in the ZI, DMH, and PH of *Lhx6*<sup>Cre/+</sup>;*Ai9* postnatal mice also continuing to express *Lhx6* (Fig. 1J–U, D'). Notably, we see a few tdTomato-expressing neurons in other hypothalamic regions, with the largest numbers found in adjacent structures such as the ventromedial hypothalamus (VMH) and lateral hypothalamus (LH), although only ~5% of these tdTomato-expressing neurons still express *Lhx6* (Fig. 1V–D'). This shows that, in contrast to telencephalic interneuron precursors, hypothalamic *Lhx6* cells do not appear to undergo long-distance tangential migration, and that hypothalamic *Lhx6*-expressing cells that do undergo short-range tangential dispersal during early development generally repress *Lhx6* expression as they mature.

## *Lhx6* is necessary for the survival of hypothalamic neurons.

These findings led us to investigate other potential differences in the *Lhx6* function in hypothalamic neurons relative to the telencephalon. While *Lhx6* does not regulate the survival of cortical interneuron precursors<sup>22</sup>, hypothalamic-specific loss of function of *Lhx6* leads to substantial changes in sleep patterns<sup>12</sup>, raising the possibility that *Lhx6* may be necessary for the viability or proper functions of these neurons.

To investigate this possibility, we tested P8 *Lhx6*<sup>CreER/CreER</sup> mice, in which a CreER cassette has been inserted in frame with the start codon to generate a null mutant of *Lhx6*, to determine if read-through transcription of endogenous *Lhx6* could be detected in the hypothalamus (Fig. 2A). Chromogenic ISH of telencephalic structures such as the amygdala and cortex reveals that *Lhx6*-expressing cells are still detected in both regions, although the number of *Lhx6*-expressing cells in the cortex is substantially reduced in *Lhx6*<sup>CreER/CreER</sup> mice relative to *Lhx6*<sup>CreER/+</sup> heterozygous controls (Fig. 2B–M). This is consistent with the severe reduction in tangential migration of cortical interneurons seen in *Lhx6*-deficient mice<sup>22,23</sup>. In the hypothalamus, however, no read-through transcription of *Lhx6* was detected (Fig. 2J, M). This implies that, in contrast to its role in the telencephalon where *Lhx6* is necessary for the tangential migration and proper laminar positioning<sup>21,23</sup>, hypothalamic *Lhx6* is required to promote neuronal survival and/or to activate its expression.

To distinguish between these possibilities, we sought to determine whether neonatal loss of function of *Lhx6* would lead to the death of *Lhx6*-expressing neurons. This was done using the genetic fate mapping of *Lhx6*-deficient neurons. Using a series of 4-Hydroxytamoxifen (4-OHT) injections between postnatal day (P) 1 and P5 in *Lhx6*<sup>CreER/+</sup>;*Ai9* and *Lhx6*<sup>CreER/lox</sup>;*Ai9* mice, we



labeled *Lhx6*-expressing cells with tdTomato while also simultaneously disrupting *Lhx6* function in a subset of *Lhx6*-expressing neurons in *Lhx6*<sup>CreER/lox</sup> mice (Fig. 2N). We then quantified the number of neurons that expressed both tdTomato and *Lhx6* protein at P45, as well as the number of neurons that only expressed tdTomato. Expression of the only tdTomato indicates that a cell has lost expression of *Lhx6*, either as a result of Cre-dependent disruption of the *Lhx6* locus or as a result of normal repression of expression during postnatal development (Fig. 2O). In both hypothalamic and telencephalic regions in *Lhx6*<sup>CreER/+</sup>; *Ai9* mice, we observed that the fraction of neurons that only

express tdTomato was only 10–15% of the number of neurons expressing both *Lhx6* and tdTomato (Fig. 2P–S, Supplementary Fig. 1). This indicates that the great majority of neurons in both regions that express *Lhx6* in neonates continue to do so at P45. However, when we performed this same analysis in *Lhx6*<sup>CreER/lox</sup>; *Ai9* mice, we found that while 75% of tdTomato-expressing neurons in the cortex and amygdala remain even in the absence of detectable *Lhx6* protein, a substantially smaller fraction of tdTomato-expressing neurons are detected in the absence of *Lhx6* protein in the ZI, DMH, and PH (Fig. 2P, T–V, Supplementary Fig. 1).

**Fig. 1 Distribution of hypothalamus *Lhx6*-expressing neurons.** **A, B, E–G** In situ hybridization showing *Lhx6* (blue) and *Shh* (brown) at E11.5 (**A**), E12.5 (**B**), E14.5 (**E**), shown in sagittal planes, and E16.5 (**F, G**) in coronal planes. Red arrows (**A, B, E**) indicate the intrahypothalamic diagonal (ID), red arrowheads (**A, B, E**) indicate the tuberomammillary terminal (TT), black arrows (**D**) indicate migrated telencephalic *Lhx6*-expressing cells (tangential migration from the medial ganglionic eminence to the cortex). Red arrows in (**F**) indicate the ZI and red circle in (**F**) indicates the DMH, and red arrows in (**G**) indicate the PH. **C, D** Schematics showing the distribution of telencephalic (green) and hypothalamic (purple) *Lhx6*-expressing cells at E11.5, with ID (red) and TT (blue) are highlighted in (**D**) (E11.5) Note anterior domains to the ID that shows a weak and transient *Lhx6* expression during development (Ant.ID anterior ID, DMH dorsomedial hypothalamus). **H** scRNA-Seq from E11–E13 hypothalamus scRNA-Seq from<sup>24</sup> showing the distribution of neurons that express GABAergic markers (blue, *Dlx1/2*, *Gad1/2*, and *Slc32a1*) and *Lhx6*-expressing GABAergic neurons (brown) that are ~2% of all hypothalamic GABAergic neurons during development. **I** Schematic distribution of *Lhx6*-expressing neurons across the dorsolateral hypothalamus (red = neurons that continue to express *Lhx6*, purple = neurons that transiently expressed *Lhx6*) across Zlv (zona incerta ventral), Zll (zona incerta lateral), LH (lateral hypothalamus), DMH (dorsomedial hypothalamus), VMH (ventromedial hypothalamus), and PH (posterior hypothalamus). **J–C'** *Lhx6*-antibody staining (gray), tdTomato expression from *Lhx6*<sup>Cre/+</sup>; *Ai9* line (red), NeuN (yellow) in Zlv (**J–M**), Zll (**N, Q**), DMH (**R–U**), LH (**V–Y**), PH (**Z–C'**). L = lateral, M = medial. White arrowheads show neurons that transiently expressed *Lhx6*, and white arrows show neurons continue to express *Lhx6*. **D'** A bar graph showing the percentage of tdTomato<sup>+</sup> and *Lhx6*-expressing neurons over the total number of tdTomato<sup>+</sup> neurons. Scale bar = 0.45 mm (**A**), 0.5 mm (**B, F, G**), 0.55 mm (**E**), 0.1 mm (**I–C'**). All bar graphs (**D'**) show mean and standard error of the mean (SEM), with individual data points plotted.

This is consistent with *Lhx6* playing a selective role in regulating the survival of *Lhx6*-expressing hypothalamic neurons. To directly address this hypothesis, we next generated *Lhx6*<sup>CreER/lox</sup>; *Bax*<sup>lox/lox</sup>; *Ai9* mice, with loss of function of *Bax* predicted to selectively prevent apoptosis in *Lhx6*-expressing neurons<sup>25</sup>. When Cre recombinase activity was induced using the same protocol, we observed that the fraction of tdTomato-expressing neurons that lacked *Lhx6* expression was indistinguishable from that seen in cortex and amygdala (Fig. 2P, W–Y, Supplementary Fig. 1).

These data indicate that *Lhx6* is selectively required for the survival of hypothalamic *Lhx6*-expressing neurons. To determine whether *Lhx6* is also required for normal differentiation of these cells, we next conducted RNA-Seq analysis on sorted tdTomato-expressing hypothalamic cells from P10 *Lhx6*<sup>CreER/+</sup>; *Ai9* and *Lhx6*<sup>CreER/lox</sup>; *Bax*<sup>lox/lox</sup>; *Ai9* mice (Fig. 2Z, Supplementary Fig. 2). We observe that *Lhx6*<sup>CreER/lox</sup>; *Bax*<sup>lox/lox</sup>; *Ai9* mice show no change in expression of markers of GABAergic neurons, including *Gad1*, *Gad2*, *Slc32a1*. However, substantially increased expression of genes expressed in mitotic neural progenitors, including *Ccna1*, *Aurka*, *Msx1*, and *Msx2* (Supplementary Fig. 2, Table S1) is observed, along with a decreased expression of axon guidance/growth factors such as *Sema3c*, *Sema4d*, and *Sema5a*. Notably, we also observe ectopic expression of genes that are not normally found in the brain but are expressed in germline stem cells (*Sycp1*), testes (*Ccdc144b*, *Samd15*, *Stag3*) mucosa (*Slc12a8*), colon (*Nlrp6*), liver (*Tfr2*), heart (*Popdc2*, *Spta1*), and cochlear hair cells (*Pdzd7*)<sup>26</sup>. This suggests that, as in telencephalic neurons, *Lhx6* is not required for expression of GABAergic markers<sup>21–23</sup>, but might be required to repress inappropriate expression of genes expressed both in neural progenitor and in nonneuronal cells. This does not, however, exclude the possibility that these may be in part induced as a result of the loss of function of *Bax*.

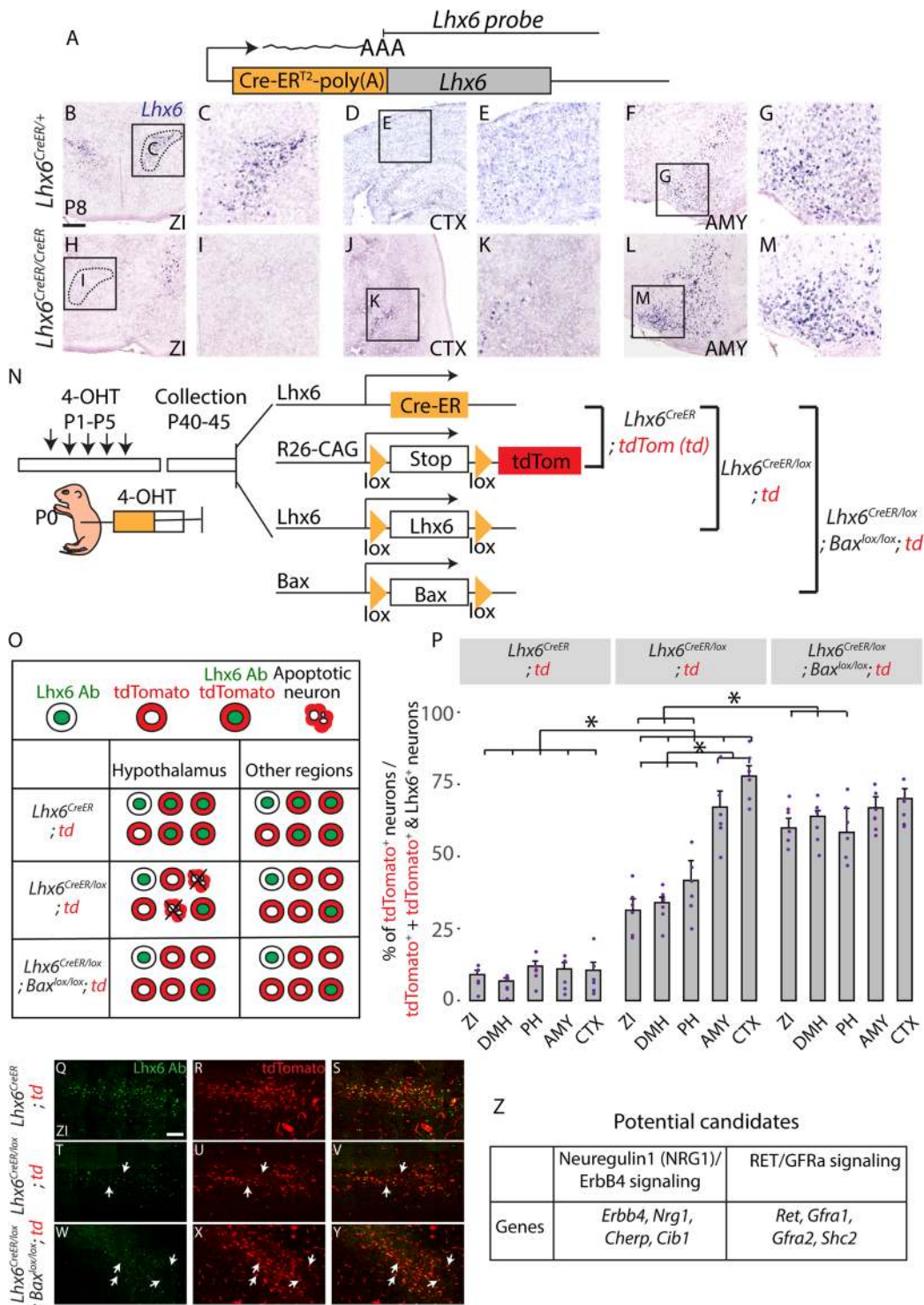
Genetic and biochemical analyses have identified several genes as direct or indirect *Lhx6* targets in the developing telencephalon<sup>18,19,21,23,27</sup>. These include *Shh*, the transcription factors *Arx*, *Cux2*, *Mafb*, and *Nkx2-1*; as well as *Sst* and chemokine receptors such as *Cxcr4*, *Cxcr7*, and *ErbB4*. To identify genes and signaling pathways that are strong candidates for selectively regulating survival of hypothalamic *Lhx6* neurons, the bulk RNA-Seq data from P10 *Lhx6*<sup>CreER/+</sup>; *Ai9* neurons were directly compared to profiles obtained from FACS-isolated *Lhx6*-GFP positive and negative hypothalamic and cortical neurons that were collected at E15.5, P8 (Fig. 2Z, Supplementary Fig. 2), since regulation of hypothalamic *Lhx6* in cell survival is evident during embryonic and early neonatal periods and we expected to

detect potential signaling pathways at both datasets. Genes found to be enriched in hypothalamic samples of bulk RNA-Seq data were then compared to scRNA-Seq datasets of hypothalamic *Lhx6*-expressing neurons collected at E15.5 and P8 (Fig. 2Z, Supplementary Fig. 2)<sup>24</sup>, and a core set of *Lhx6*-regulated genes that were selectively enriched in hypothalamic *Lhx6*-expressing neurons was thus identified.

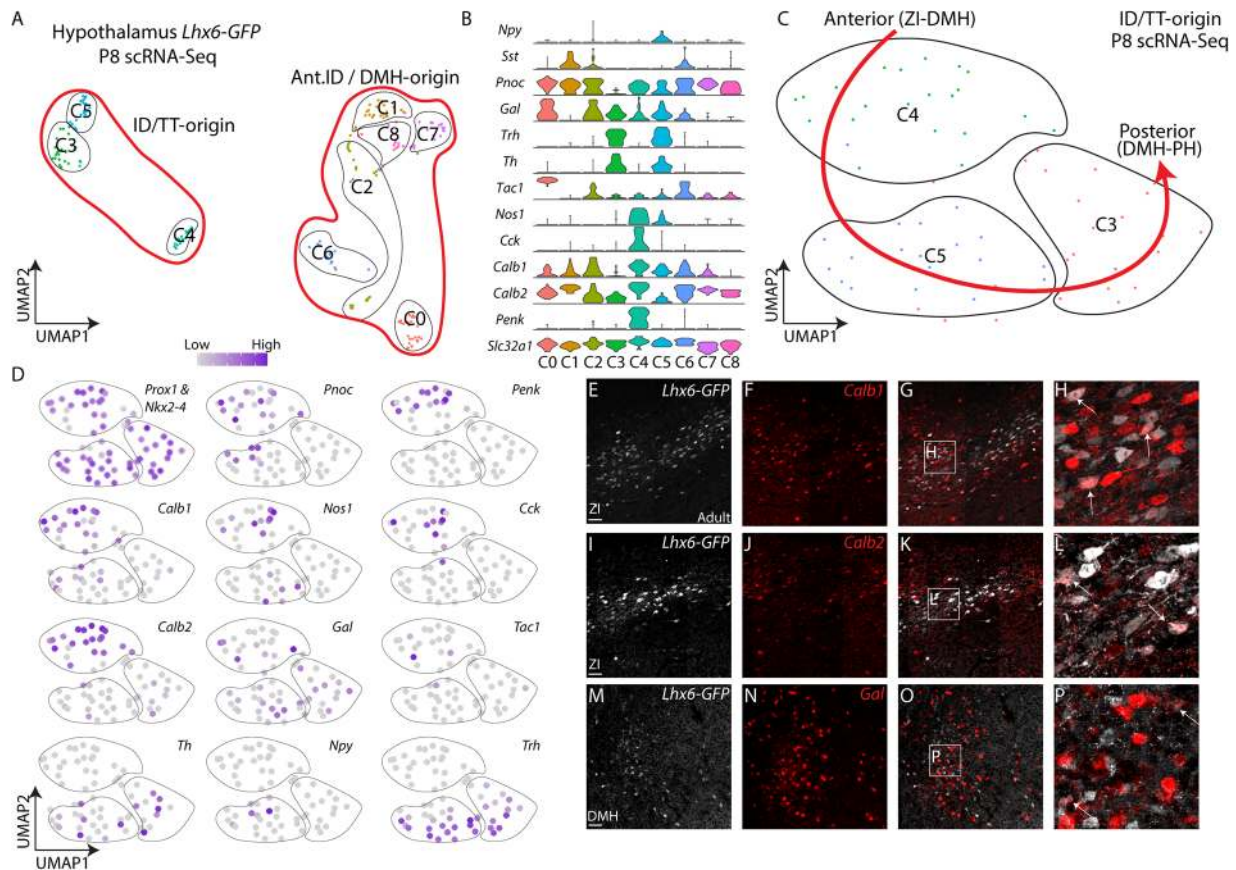
We observe that many previously identified *Lhx6* targets either show little detectable expression in wildtype hypothalamic *Lhx6* neurons, (*Cux2*, *Mafb*, *Sst*, and *Cxcr4/7*) or else showed no detectable change in expression following *Lhx6* loss of function (*Arx*, *Nkx2-1*). One notable exception is the Neuregulin receptor *ErbB4*, which has been shown to be necessary for tangential migration and differentiation of MGE-derived immature *Lhx6*-expressing cortical interneurons<sup>28–30</sup>. *ErbB4* is both highly expressed in hypothalamic *Lhx6* neurons, and its expression is strongly *Lhx6*-dependent (Fig. 2Z, Supplementary Fig. 2). Since Neuregulin signaling is also neurotrophic in many cell types<sup>31</sup>, this suggested that the loss of neuregulin signaling could be a potential mechanism behind the apoptotic death of *Lhx6*-deficient hypothalamic cells. Indeed, we observed that additional components of the both the Neuregulin (*Nrg1*) and Gdnf (*Ret*, *Gfra1*, *Gfra2*) neurotrophic signaling pathways were selectively enriched in hypothalamic *Lhx6* neurons (Fig. 2Z, Supplementary Fig. 2), a finding which was confirmed using fluorescent ISH and scRNA-Seq (Supplementary Fig. 3).

**Diverse subtypes of *Lhx6*-expressing neurons are found in the postnatal hypothalamus.** Our previous work<sup>12</sup> showed that adult ZI *Lhx6*-expressing neurons do not highly express traditional markers of MGE *Lhx6*<sup>+</sup> derived GABAergic neurons of the cortex. No ZI *Lhx6*-expressing neurons co-express *Pvalb* and *Sst*, with only a small subset expressing *Npy*<sup>12</sup>. We thus hypothesized that subtypes of *Lhx6* neurons in the postnatal hypothalamus might be diverged substantially from those present in the cortex<sup>32</sup>, and might be more molecularly heterogeneous.

scRNA-Seq analysis of P8 *Lhx6*-eGFP neurons from the hypothalamus that expressed high levels of *Lhx6* mRNA shows that these neurons express a diverse pool of neuropeptides and neurotransmitters that are not expressed in telencephalic *Lhx6*-expressing neurons, including *Gal*, and *Trh* (Fig. 3, Supplementary Fig. 4, Table S2). Other markers that are specific to distinct subsets of cortical *Lhx6* neurons were expressed in hypothalamic *Lhx6* neurons, such as *Pnoc*, *Tac1*, *Nos1*, and *Th*. Hypothalamic *Lhx6*-expressing neurons do not express *Pvalb*, but a small fraction expresses *Npy* and *Cck*. We also identified a rare subpopulation of hypothalamic *Lhx6*-expressing neurons in the



**Fig. 2** *Lhx6* in the hypothalamus is necessary for neuronal survival. **A** *Lhx6*<sup>CreER</sup> knock-in site (JAX #010776) and location of the *Lhx6* probe used to detect read-through transcription. **B–M** Coronal planes showing *Lhx6* mRNA expression in the zona incerta (ZI, **B, C, H, I**), amygdala (AMY, **F, G, L, M**), and cortex (CTX, **D, E, J, K**) in control (*Lhx6*<sup>CreER/+</sup>, **B–G**) and mutant (*Lhx6*<sup>CreER/CreER</sup>, **H–M**) at P8. Note *Lhx6* mRNA is not detected in the *Lhx6*-deficient hypothalamus but is detected in the telencephalon. **N** Schematic diagram showing the overall design of the experiment from three genotypes (1. *Lhx6*<sup>CreER/+</sup>;Ai9, 2. *Lhx6*<sup>CreER/lox</sup>;Ai9, 3. *Lhx6*<sup>CreER/lox</sup>;Bax<sup>lox/lox</sup>;Ai9). **O** Schematic diagram showing the overall outcome of the experiment. **P** A bar graph showing the percentage of tdTomato<sup>+</sup> / (tdTomato<sup>+</sup> + tdTomato<sup>+</sup>/Lhx6<sup>+</sup>) across three genotypes in five different brain regions. \*Indicates p < 0.05. DMH dorsomedial hypothalamus, PH posterior hypothalamus. **Q–Y** Representative images of three genotypes (1. *Lhx6*<sup>CreER/+</sup>;Ai9 (**Q–S**), 2. *Lhx6*<sup>CreER/lox</sup>;Ai9 (**T–V**), 3. *Lhx6*<sup>CreER/lox</sup>;Bax<sup>lox/lox</sup>;Ai9 (**W–Y**) in ZI. More images of different brain regions are available in Supplementary Fig. 1. **Z** Potential candidate genes from bulk RNA-Seq (Supplementary Fig. 2) controlling cell survival can be regulated by *Lhx6* in hypothalamic Lhx6<sup>+</sup> neurons: Neuregulin-ErbB4 signaling and Gdnf signaling pathways (Supplementary Figs. S3, S4). Scale bar = 0.6 mm (**B–M**), 100 μm (**Q–Y**). All bar graphs (**P**) show mean and standard error of the mean (SEM), with individual data points plotted.



**Fig. 3 Diverse subtypes of mature hypothalamic *Lhx6*-expressing neurons.** **A** Uniform Manifold Approximation and Projection (UMAP) plot showing hypothalamic *Lhx6*-expressing GABAergic neurons at P8 scRNA-Seq. **B** Violin plots showing key markers in individual clusters in **(A)**. **C** UMAP plot showing *Lhx6*-expressing neurons originated from ID and TT. **D** UMAP plots showing the distribution of diverse neuropeptides and neurotransmitters across ID and TT derived *Lhx6*-expressing neurons. **E–P** Fluorescent in situ hybridization showing *Lhx6*-GFP (gray) with *Calb1* (red, **E–H**), *Calb2* (red, **I–L**), and *Gal* (red, **M–P**). Scale bar = 50  $\mu$ m.

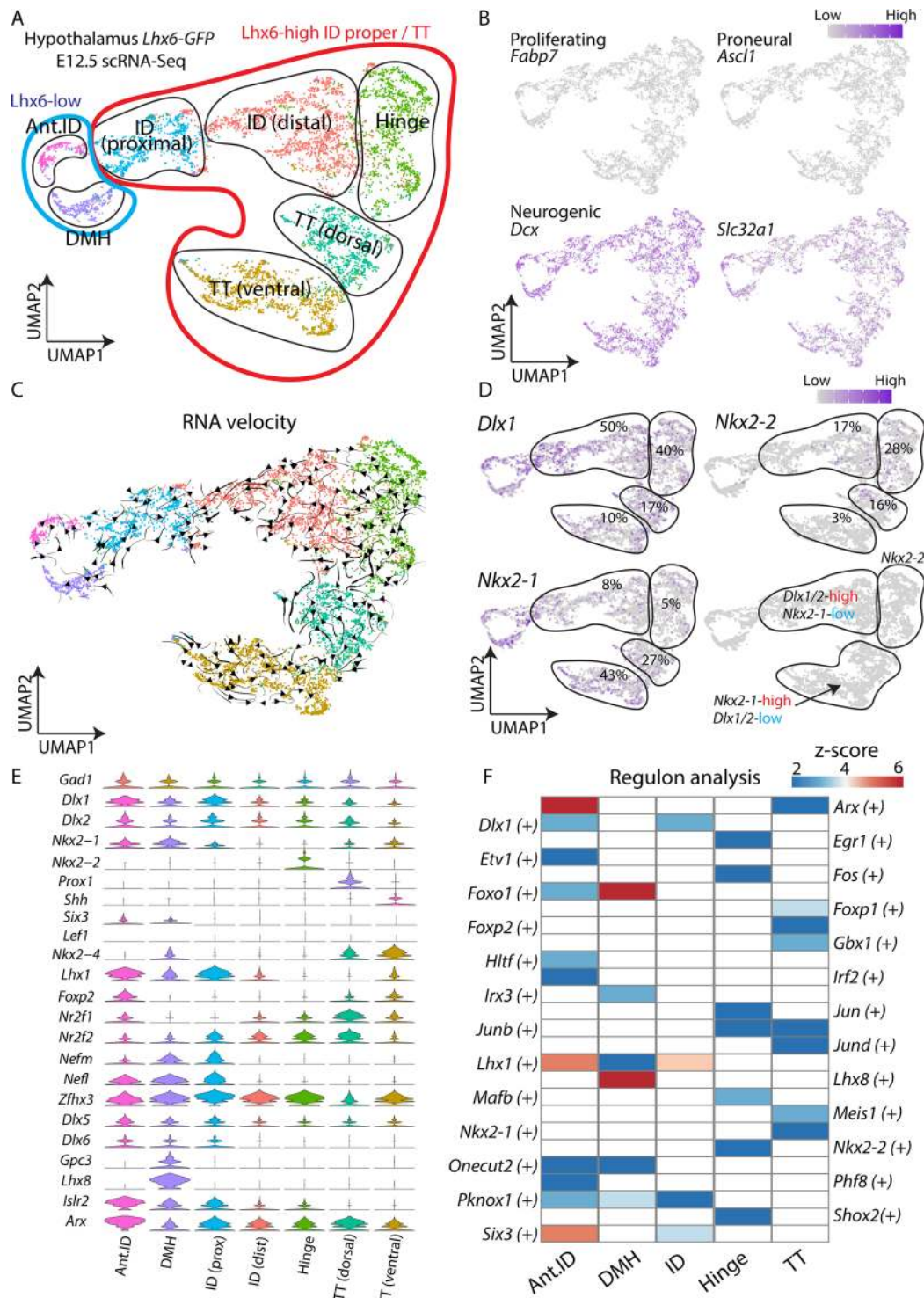
PH that co-express *Sst*, although these are absent in more anterior regions (Fig. 3, Supplementary Fig. 4). *Tac1* is expressed broadly in cortical and hypothalamic *Lhx6*-expressing neurons. Similar patterns of gene expression are observed in scRNA-Seq data obtained from *Lhx6* neurons in the adult hypothalamus of mice that are older than P30 (Supplementary Fig. 5)<sup>24,33,34</sup>. However, all these enriched markers (neuropeptides and neurotransmitters) are not specific to *Lhx6*-expressing neurons but rather expressed broadly in hypothalamic GABAergic neurons across nuclei (Supplementary Fig. 6).

Mature *Lhx6* hypothalamic neurons were organized into three major clusters that showed close similarity to the two subdomains of the ID and the main TT region observed at E12.5, and in turn appear to represent individual subtypes of *Lhx6* neurons that are differentially distributed along the anteroposterior axis of the hypothalamus, and which may correspond to *Lhx6* neurons of the ZI, DMH, and PH, respectively (Fig. 3, Supplementary Fig. 4). *Lhx6* neurons express a mixture of *Pnoc*, *Penk*, *Calb1*, *Calb2*, *Cck* in both the ZI and DMH, whereas *Tac1* is more restricted to the ZI. *Npy* and *Nos1* are enriched in DMH *Lhx6* neurons. *Th*, *Trh*, *Gal* are located in the region spanning the DMH and PH, while *Sst* is expressed only in a small subset of PH *Lhx6* neurons.

**scRNA-Seq identifies molecular markers of spatially distinct domains of hypothalamic *Lhx6* neurons.** *Lhx6*-expressing neurons of the postnatal hypothalamus are molecularly diverse and

distributed across a broad region of the dorsolateral hypothalamus. We hypothesized that this diversity is regulated by multiple transcription factors that control the specification of region-specific subtypes of *Lhx6*-expressing neurons.

To identify these anatomically and molecularly distinct *Lhx6*-expressing domains in the hypothalamus, we performed scRNA-Seq with the *Lhx6*-GFP line at E12.5 and E15.5. At E12.5 and E15.5, scRNA-Seq analysis readily distinguishes the ID, TT, and hinge domains (Fig. 4A, Supplementary Figs. 7, 8). By E12.5, all *Lhx6* cells in the hypothalamus express the early neuronal precursor marker *Dcx*, as well as the synaptic GABA transporter *Slc32a1*, but do not express progenitor markers (e.g., *Fabp7* and *Ascl1*). It is not immediately clear whether the molecular identities of anatomically and molecularly distinct clusters of *Lhx6*-expressing cells in the ID, TT and hinge clusters are already distinct at E12.5, we used RNA velocity analysis<sup>35</sup> to determine whether any cells appeared to be undergoing transition between individual clusters. RNA velocity analysis does not identify trajectories connecting individual clusters, indicating that their regional identity appears to be fixed by this age (Fig. 4B, C). In addition, weak *Lhx6* expression was observed in *Lhx1* and *Lhx8* co-expressing neurons of the anterior ID cluster, which are *Nkx2-1*<sup>+</sup> (Fig. 4D, E), and give rise to GABAergic neurons in the suprachiasmatic nucleus and DMH, although little or no *Lhx6* mRNA was detected in these neurons after E13.5 (Figs. 1, 6)<sup>4,10</sup>. We observed that *Dlx1/2*, *Nkx2-2*, and *Nkx2-1* are differentially expressed in the ID, hinge, and TT domains, respectively, at both



**Fig. 4** scRNA-Seq identifies molecular markers of spatially distinct domains of hypothalamic *Lhx6* neurons. **A** UMAP plot showing different *Lhx6*-expressing hypothalamic regions at E12.5. **B** UMAP plots showing a lack of expression of *Fabp7* (proliferating cells), *Ascl1* (proneural), whereas the neuronal precursor marker *Dcx* and the GABAergic neuronal marker *Slc32a1* is highly expressed. **C** UMAP plot with RNA velocity trajectories. **D** UMAP plots showing expression and percentage of ID, hinge, and TT *Lhx6* neurons expressing *Dlx1/2*, *Nkx2-2*, and *Nkx2-1*. **E** Violin plots showing expression of key transcription factors (and other genes) that are highly expressed in individual domains. **F** A heatmap showing z-scores of significantly differentially expressed key regulatory transcription factors among *Lhx6*<sup>+</sup> hypothalamic regions. Note activity of *Dlx1* in the ID, *Nkx2-2* in the hinge, and *Nkx2-1* in the TT. Ant.ID anterior ID, DMH dorsomedial hypothalamus.

ages (Fig. 4D, E, Supplementary Fig. 8). These three transcription factors are each shown as key putative regulatory transcription factors of the ID, hinge, and TT domains respectively (Fig. 4F). Furthermore, we observe several molecularly distinct cell clusters

that have not been previously described. The first cluster expresses low levels of *Nkx2-1*, but high levels of *Prox1* and *Sp9*, transcription factors that are highly expressed in the developing prethalamus. This may therefore correspond to a



dorsal subdomain of the TT located adjacent to the hinge domain (Fig. 4D, E). We also observe a distinction between more proximal and distal domains of the ID, based on the expression of *Nefl*, *Dlx6*, *Nefm*, *Lhx1*, and *Nr2f1*.

In all, five molecularly distinct clusters of neurons that strongly express *Lhx6* could be resolved in the embryonic hypothalamus (Fig. 4). These can be distinguished not only by the expression of different subsets of transcription factors at E12.5, but also by more conventional markers of cell identity such as neuropeptides and calcium-binding proteins such as *Sst*, *Tac1*, *Pnoc*, *Islr2*, *Gal*, and *Npy* at E15.5 (Supplementary Fig. 8, Tables S3, S4). We also observed clusters that were located in the hinge and TT region at E12.5 (Supplementary Fig. 8 cluster 4 and 7), but which postnatally expressed markers that are restricted to neurons at the most anterior domain of hypothalamic *Lhx6* neurons. These markers include *Nfix*, *Nfib*, and *Tcf4* (Supplementary Fig. 8, Fig. 3, Tables S2–S4).

These molecularly distinct domains of hypothalamic *Lhx6* neurons were also visualized using traditional two-color ISH with *Nkx2-1*, *Nkx2-2*, *Arx*, and *Prox1* probes (Supplementary Fig. 9). This also confirms that *Shh* is only expressed in dorsal TT *Lhx6* neurons, while *Six3* is expressed only in the weakly *Lhx6*-expressing neurons in the anterior ID. scRNA-Seq showed that *Lef1*, which is expressed broadly in the ID and TT region at E12.5, was expressed in only very few *Lhx6* neurons at both E12.5 and E15.5 (Fig. 4, Supplementary Fig. 9), indicating that *Lef1* and *Lhx6* are not extensively co-expressed.

***Dlx1/2*, *Nkx2-2*, and *Nkx2-1* mediate patterning of discrete spatial domains of hypothalamic *Lhx6* neurons.** *Dlx1/2*, *Nkx2-2*, and *Nkx2-1* are selectively expressed in the ID, hinge, and TT domains, respectively. Since these three transcription factors were also identified as putative key regulatory transcription factors from scRNA-Seq analysis, we sought to investigate their function in regulating *Lhx6* expression in more detail. Using *Lhx6-GFP* mice, which faithfully recapitulate the endogenous expression pattern of *Lhx6*<sup>12</sup>, we integrated bulk RNA-Seq analysis obtained at E15.5 and P0 from hypothalamus with age-matched ATAC-Seq data to cross-reference our scRNA-Seq result (Fig. 5A). We further sought to investigate similarities and differences in gene expression and chromatin accessibility in age-matched hypothalamic and telencephalic *Lhx6*-expressing neurons (Fig. 5), since the role of *Lhx6* in development of telencephalic interneurons is extensively studied, and it is therefore critically important to connect these findings to prior work characterizing *Lhx6* mechanisms of action in forebrain development.

At E15, many region-specific differences in gene expression were observed between hypothalamic and telencephalic *Lhx6*-expressing neurons, particularly for transcription factors. We observed enriched expression of *Six3*, *Nkx2-2*, and *Nkx2-4* in the hypothalamus. As predicted by earlier studies, we observed enriched expression of the telencephalic marker *Foxg1*, *Satb2*, and *Nr2e1*<sup>36,37</sup>, in the cortex (Fig. 5B, Table S5).

However, expression of genes broadly expressed in GABAergic neurons showed no significant differences, including *Nkx2-1*, and *Dlx1/2*. At P0, hypothalamic *Lhx6* neurons continued to show enriched expression for multiple transcription factors, including *Prox1*, *Foxp2*, and *Nhlh2*. Hypothalamic *Lhx6* neurons show little detectable expression of the cortical interneuron markers *Pvalb*, *Sst*, and *Npy*, but we observed a higher level of *Gal* and *Pnoc* at P0 in hypothalamic *Lhx6* neurons. Relative to *Lhx6*-negative hypothalamic neurons, we also observed a higher level of transcription factors such as *Dlx1*, *Onecut1*, *Pax5*, and *Nkx2-2* in hypothalamic *Lhx6* neurons compare to the rest of the hypothalamus at E15.5, as well as a higher level of *Tac1* and *Pnoc* at P0 (Supplementary Fig. 10, Table S6).

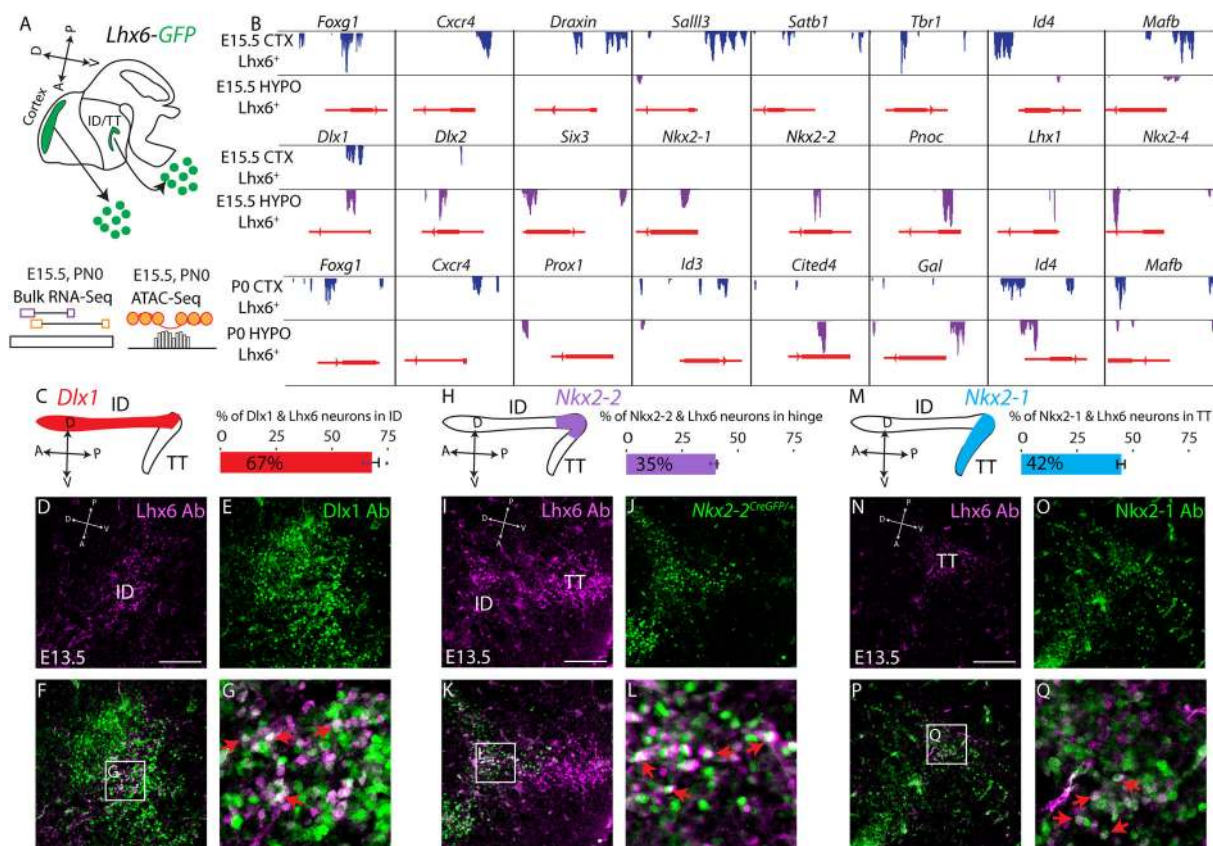
Regions of accessible chromatin identified by ATAC-Seq were, as expected, clustered in the proximal promoter and intronic regions of annotated genes in all samples profiled (Supplementary Fig. 10, Tables S7, S8). Region-specific differences in chromatin accessibility frequently corresponded to differences in mRNA expression. For instance, proximal promoter and/or intronic regions of *Foxg1*, *Npy*, *Pvalb*, and *Sst* were selectively accessible in cortical *Lhx6* neurons, while those of *Nkx2-2*, *Sall3*, and *Gal* were accessible only in the hypothalamus at both E15.5 and P0 (Fig. 5B, Tables S7, S8, Supplementary Fig. 10). However, substantial differences in chromatin accessibility were also observed for *Nkx2-1* and *Dlx1/2* at both E15.5 and P0, implying that different gene regulatory networks may control the expression of these genes in hypothalamus and cortex (Tables S7, S8).

To determine whether any of the spatial domains of *Lhx6* expression could closely resemble telencephalic *Lhx6* cells, we compared E12.5 hypothalamic scRNA-Seq results to data previously obtained from E13.5 MGE<sup>38</sup>. These data confirmed that, while transcription factors such as *Nkx2-1*, *Dlx1/2*, and *Lhx8* are broadly expressed in *Lhx6* MGE cells, they are not expressed (*Lhx8*) or expressed only in discrete subsets (*Nkx2-1*, and *Dlx1/2*) of hypothalamic *Lhx6* neurons. No identified subset of hypothalamic *Lhx6* neurons resembled MGE *Lhx6* cells (Supplementary Fig. 11A, B, Table S9).

With substantial differences between hypothalamic and telencephalic *Lhx6*-expressing neurons in both gene expression and chromatin accessibility, we reasoned that the transcriptional regulatory networks identified as controlling the development of telencephalic *Lhx6*-expressing neurons would not broadly apply in developing hypothalamus. Thus, based on both scRNA-Seq data and analysis of our ATAC-Seq data, as well as our previous work<sup>4,24</sup>, three previously mentioned transcription factors—*Nkx2-1*, *Dlx1/2*, and *Nkx2-2*—emerged as strong candidates for regulating specific domains of hypothalamic *Lhx6* neurons. *Nkx2-1* is required for *Lhx6* expression in the telencephalon<sup>16,19</sup> and is expressed in the TT, but not ID, domain in the hypothalamus<sup>4,24</sup>, while *Dlx1/2* are required for tangential migration of cortical interneurons and are also broadly expressed in both cortical and hypothalamic *Lhx6* neurons<sup>4,6,8,39</sup>. *Nkx2-2*, in contrast, is expressed only in the hypothalamus in a zone immediately dorsal to the region of *Nkx2-1* expression<sup>4,40</sup>.

Each of these transcription factors is expressed in discrete spatial domains that overlap with distinct subsets of hypothalamic *Lhx6* neurons at E13.5 (Fig. 5C–Q). *Dlx1* was strongly expressed in the ID (Fig. 5C–G, Supplementary Fig. 11C–H), but not the TT. *Nkx2-2*, in contrast, selectively demarcated the region joining the ID and TT (Fig. 5H–L), which we have termed the hinge domain. *Nkx2-1* was selectively expressed in the TT region, but essentially absent from the ID and hinge domain (Fig. 5M–Q, Supplementary Fig. 11I–K). These spatial differences in the expression of *Dlx1* and *Nkx2-1* in hypothalamic *Lhx6* neurons are preserved at E17.5, where *Dlx1* is enriched in the more anterior ZI and DMH (Supplementary Fig. 12), and *Nkx2-1* expression is enriched in the PH (Supplementary Fig. 12A–L). Furthermore, unlike the MGE *Lhx6*-expressing cells, *Dlx1* and *Nkx2-1* formed mutually exclusive expression domains in the ID and TT (Supplementary Fig. 11F–K). However, we observed a much more even distribution of *Nkx2-2*/*Lhx6* neurons across the ZI, DMH, and PH, which could indicate either short-range tangential dispersal of hinge neurons or widespread induction of *Nkx2-2* expression in *Lhx6* neurons at later ages (Supplementary Fig. 12M–X). These results indicate that distinct spatial domains of hypothalamic *Lhx6* expression can be delineated by combinatorial patterns of homeodomain transcription factor expression.

To determine the final location of *Nkx2-1* expressing *Lhx6* neurons, we next used fate-mapping analysis, in which *Nkx2-*



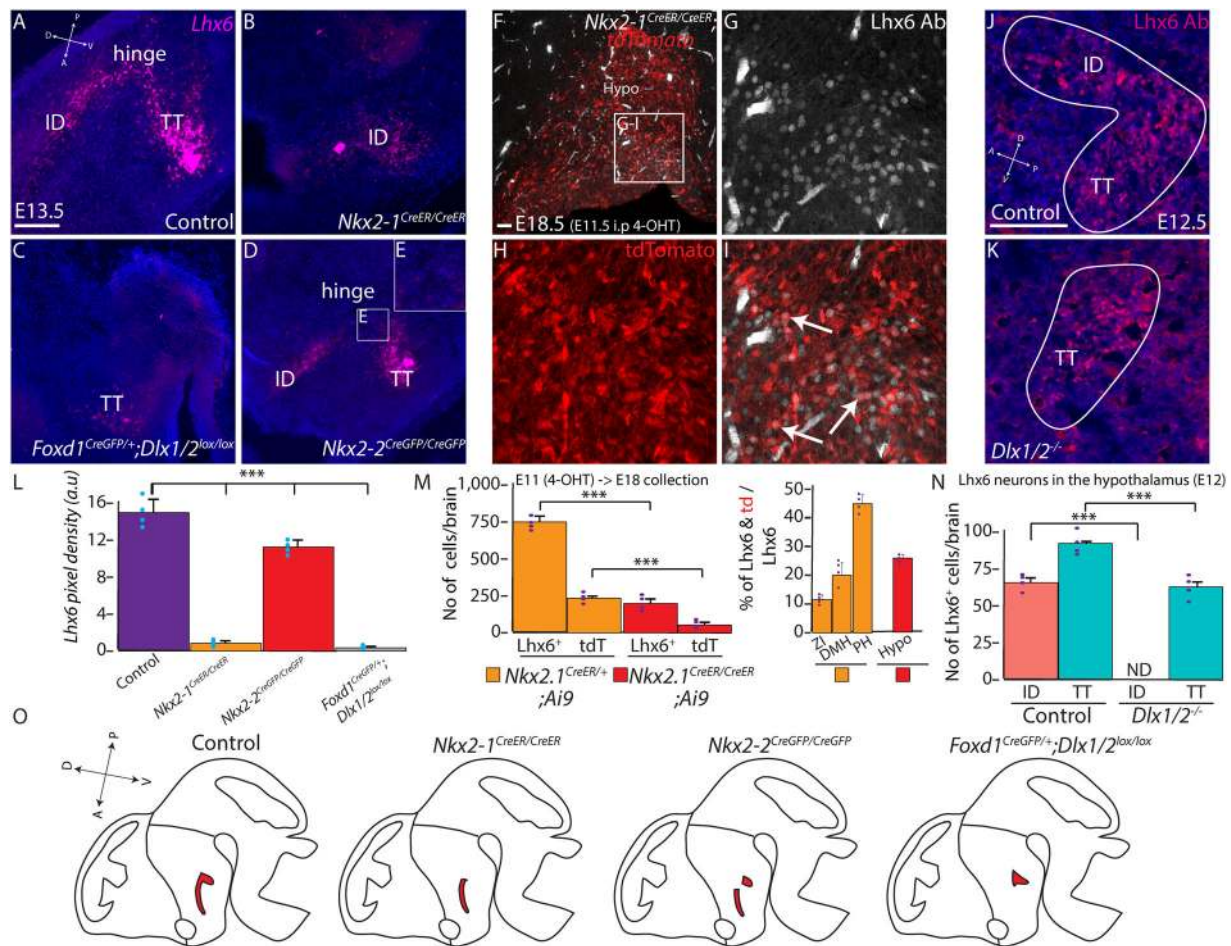
**Fig. 5** *Dlx1/2*, *Nkx2-2*, and *Nkx2-1* are expressed in distinct spatial domains of hypothalamic *Lhx6* neurons. **A** Schematic showing bulk RNA-Seq and bulk ATAC-Seq pipelines from flow-sorted *Lhx6-GFP*<sup>+</sup> neurons of the cortex and hypothalamus at E15.5 and P0. **B** Peaks in ATAC-Seq showing potential transcription factor binding sites near the promoter regions of differentially expressed genes from bulk RNA-Seq data in the cortex and hypothalamus at E15.5 and P0. **C** Schematic showing the ID/TT of the developing hypothalamus and expression of *Dlx1/2* (left), and the percentage of ID *Lhx6*-expressing neurons that co-express *Dlx1* (right). **D–G** Immunostaining with *Lhx6* (purple) and *Dlx1* (green) of E13.5 hypothalamus, showing co-localization of *Lhx6* and *Dlx1* in the ID of the hypothalamus (**G**, red arrows). **H** Schematic showing ID/TT of the developing hypothalamus and expression of *Nkx2-2* (left) and the percentage of hinge *Lhx6*-expressing neurons that co-express *Nkx2-2* (right). **I–L** Immunostaining with *Lhx6* (purple) and GFP in *Nkx2-2*<sup>CreGFP/+</sup> (green) of E13.5 hypothalamus show co-localization of *Lhx6* and *Nkx2-2*<sup>GFP</sup> between the ID and TT (hinge region) of the hypothalamus (**L**, red arrows). **M** Schematic showing the ID/TT of the developing hypothalamus and expression of *Nkx2-1* (left) and the percentage of TT *Lhx6*-expressing neurons that co-express *Nkx2-1* (right). **N–Q** Immunostaining with *Lhx6* (purple) and *Nkx2-1* (green) of E13.5 hypothalamus, showing co-localization of *Lhx6* and *Dlx1* in the ID of the hypothalamus (**Q**, red arrows). Scale bar = 50  $\mu$ m.

*1*<sup>CreER/+</sup>; *Ai9* mice<sup>41</sup> were labeled with 4-OHT at E11 (Supplementary Fig. 13). At E18, tdTomato expression was detected in the majority of *Lhx6*-expressing neurons in the amygdala and cortex (Supplementary Fig. 13) as expected<sup>16,19</sup>, but we observed anterior–posterior bias in the distribution of tdTomato-expressing neurons in the hypothalamus that closely matched the location of *Lhx6/Nkx2-1* expressing neurons at earlier ages. We observe that only a small fraction (~10%) of ZI *Lhx6*-expressing neurons, which correspond to the most anterior region of *Lhx6* expression at later developmental ages<sup>12</sup>, were labeled with tdTomato. In contrast, a much larger fraction of PH *Lhx6* neurons, corresponding to the most posterior domain of *Lhx6* expression, were tdTomato positive. This implies that *Nkx2-1/Lhx6*-expressing neurons of the TT primarily give rise to *Lhx6* neurons found in the PH, but that a small fraction may undergo tangential migration to more anterior structures such as the ZI. This was also shown with immunostaining of *Nkx2-1* and *Lhx6*-expressing neurons at E17.5 (Supplementary Fig. 12Y–J’).

We next investigated whether loss of function of *Nkx2-1*, *Nkx2-2*, and *Dlx1/2* led to the loss of spatially-restricted hypothalamic

expression of *Lhx6*. We first examined *Nkx2-1*<sup>CreER/CreER</sup> mice, in which targeted insertion of the CreER cassette generates a null mutation in *Nkx2-1*<sup>41</sup>. This leads to severe hypoplasia of the posteroventral hypothalamus, as previously reported for targeted *Nkx2-1* null mutants<sup>42</sup>. The ventrally-extending TT domain of *Lhx6* expression is not detected in *Nkx2-1*-deficient mice at E12.5, but the *Nkx2-1*-negative ID domain persists (Fig. 6, Supplementary Fig. 14). Fate-mapping analysis, in which *Nkx2-1*<sup>CreER/+</sup>; *Ai9* and *Nkx2-1*<sup>CreER/CreER</sup>; *Ai9* mice were injected with tamoxifen at E11 and analyzed at E18, indicate that surviving *Lhx6* neurons in the ID region represent a mixture of tdTomato-positive and -negative neurons, and confirm that a subset of these surviving neurons derived from *Nkx2-1*-expressing precursors. As previously reported, no *Lhx6*-expressing neurons are detected in the mutant cortex (Supplementary Fig. 14).

We next generated null mutants of *Nkx2-2* in the same manner, generating mice homozygous for a knock-in CreGFP cassette that disrupts expression of the endogenous *Nkx2-2* locus<sup>43</sup>. In this case, we observe a loss of *Lhx6* expression in the hinge region, located between the posterior ID and dorsal TT (Fig. 6, Supplementary Fig. 14). Finally, we examined the



**Fig. 6** *Dlx1/2*, *Nkx2-2*, and *Nkx2-1* mediate patterning of discrete spatial domains of hypothalamic *Lhx6* neurons. **A–E, L** RNA scope showing *Lhx6* expression (magenta) in control (**A**), *Nkx2-1<sup>CreER/CreER</sup>* (**B**), *Foxd1<sup>Cre/+</sup>;Dlx1/2<sup>lox/lox</sup>* (**C**), and *Nkx2-2<sup>CreGFP/CreGFP</sup>* (**D, E**). Pixel density of *Lhx6* is shown across all four groups in (**L**). **F–I, M** *Nkx2-1<sup>CreER/CreER</sup>;Ai9* (4-OHT treatment at E11.5, collection at E18.5) showing *Lhx6* expression (gray) and tdTomato (red) in the hypothalamus. Arrows in (**I**) indicate *Lhx6<sup>+</sup>* and tdTomato<sup>+</sup> neurons. Raw number of *Lhx6<sup>+</sup>* or tdTomato<sup>+</sup> (tdT) neurons (left) and percentage of *Lhx6<sup>+</sup>* and tdTomato<sup>+</sup> neurons (right) in *Nkx2-1<sup>CreER/+</sup>;Ai9* (Supplementary Fig. 5) and *Nkx2-1<sup>CreER/CreER</sup>;Ai9*. **J, K, N** *Lhx6* expression (magenta) in control and *Dlx1/2<sup>-/-</sup>* at E12 ID and TT. The number of *Lhx6*-expressing neurons in the ID and TT are shown in (**N**). **O** Schematic diagram showing distribution of *Lhx6* expression across four groups. ID intrahypothalamic diagonal, TT tuberomammillary terminal. Scale bar = 50 μm. \*\*\**p* < 0.05. All bar graphs (**L, M, N**) show mean and standard error of the mean (SEM), with individual data points plotted.

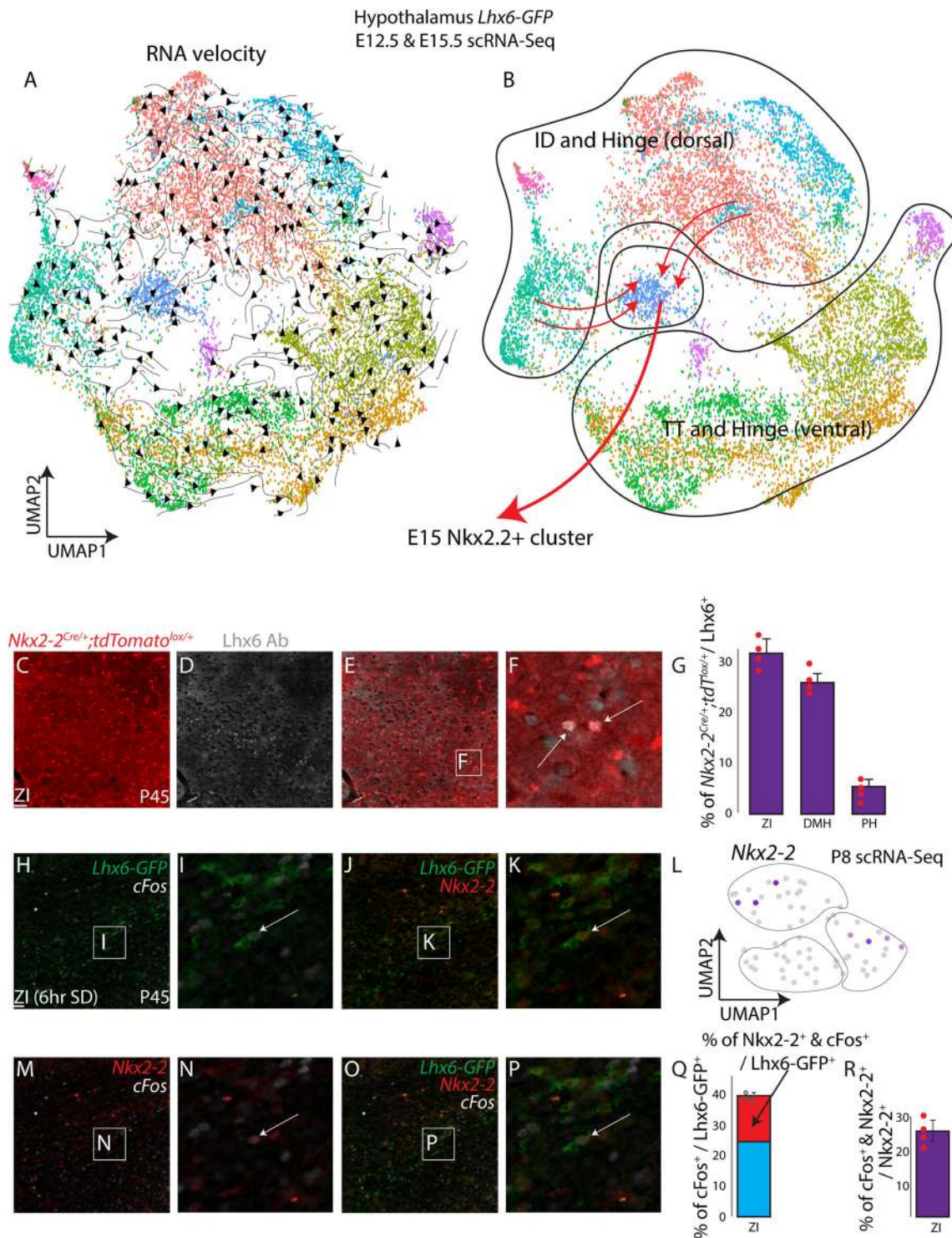
phenotype of mice deficient for *Dlx1/2*, examining both global knockouts<sup>44</sup> and *Foxd1<sup>Cre/+</sup>;Dlx1/2<sup>lox/lox</sup>* mutants<sup>45</sup>, in which *Dlx1/2* are selectively deleted in hypothalamic and prethalamic neuroepithelium<sup>12,46,47</sup>. In both global and diencephalic-specific *Dlx1/2* knockouts, the ID domain of *Lhx6* expression is absent at E12.5, whereas the TT domain is intact. At E17, we also observe a major reduction in the number of *Lhx6*-expressing neurons in the ZI (Supplementary Fig. 14). These results indicate that spatially discrete domains of hypothalamic *Lhx6* expression are controlled by the expression of different transcription factors (Supplementary Fig. 14).

***Nkx2.2*-derived *Lhx6*-expressing neurons in the ZI respond to sleep pressure.** Our previous work showed that around 40% of ZI *Lhx6*-expressing neurons respond to sleep pressure, and ZI *Lhx6* neurons promote both REM and NREM sleep<sup>12</sup>. We sought to identify whether *Lhx6* neurons derived from *Nkx2.2*-expressing precursors might selectively respond to sleep pressure. *Nkx2-2* is uniquely expressed in hypothalamic *Lhx6* neurons, but *Nkx2-2* is absent in cortical *Lhx6* neurons, unlike *Nkx2-1* and *Dlx1/2*. Our scRNA-Seq analysis and immunostaining indicate that a small

number of *Nkx2-2<sup>+</sup>* *Lhx6*-expressing neurons are located in the postnatal ZI (Supplementary Fig. 14). In addition, RNA velocity analysis on the combined E12.5 and E15.5 scRNA-Seq datasets to identify potential lineage relationships between individual clusters at E12.5 and E15.5 indicates that cells in the *Nkx2-2<sup>+</sup>* cluster in E15.5 are derived from both cells located in the ID and hinge region, indicating potential short-range tangential migration from the hinge region to the ID, which in turn leads to a subset of *Nkx2-2<sup>+</sup>* *Lhx6*-expressing neurons reaching the ZI (Fig. 7). This is supported by our observation that 28% of *Lhx6* ZI neurons express *Nkx2.2* at E17.5 (Supplementary Fig. 12).

To determine what fraction of *Lhx6*-expressing ZI neurons express *Nkx2.2* during development, we performed lineage-tracing analysis using *Nkx2-2<sup>Cre/+</sup>;Ai9* mice, analyzing the distribution of TdTom/*Lhx6*-expressing neurons at P45 (Fig. 7, Supplementary Fig. 14), and observed that ~30% of *Lhx6* ZI neurons co-labeled with tdTomato, along with a similar fraction of *Lhx6* DMH neurons. In contrast, only a small fraction (~5%) of PH *Lhx6*-expressing neurons were labeled with tdTomato.

scRNA-Seq analysis and immunostaining reveal that 30% of *Lhx6*-expressing ZI neurons continue to express *Nkx2.2* in



**Fig. 7** *Nkx2.2*-expressing *Lhx6* ZI neurons respond to increased sleep pressure. **A** UMAP plot with RNA velocity trajectories of E12.5 and E15.5 combined scRNA-Seq dataset. **B** UMAP plot shows distinct domains of hypothalamic *Lhx6* neurons. A specific population that continues to express the transcription factor *Nkx2-2* is derived from the ID and dorsal hinge region. **C–F** TdTomato expression in *Nkx2-2<sup>Cre/+</sup>;tdTomato<sup>lox/+</sup>* mice (red), and *Lhx6*-antibody staining (gray) identifies *Lhx6* neurons in the zona incerta (ZI) are derived from *Nkx2-2*-expressing precursors. **G** A bar graph showing the percentage of tdTomato<sup>+</sup> and *Lhx6*-expressing neurons relative to the total number of *Lhx6*-expressing neurons in the ZI, dorsomedial hypothalamus (DMH, Supplementary Fig. 15) and posterior hypothalamus (PH, Supplementary Fig. 15). **H–P** GFP expression from *Lhx6-GFP* (green, **H–P**), *cFos* antibody staining (gray) and *Nkx2-2* antibody staining (red) shows a specific population of *Nkx2-2*<sup>+</sup> *Lhx6* neurons in ZI that selectively responds to sleep pressure. **L** UMAP plot showing *Nkx2-2* expression in the anterior portion of *Lhx6* neurons. **Q** A bar graph showing the percentage of *cFos*<sup>+</sup> and *Lhx6-GFP*<sup>+</sup> neurons relative to the total number of *Lhx6-GFP*<sup>+</sup> neurons, and demonstrates that a subset of sleep pressure-responsive *Lhx6* neurons express *Nkx2-2*. **R** A bar graph showing the percentage of *cFos*<sup>+</sup> and *Nkx2-2*<sup>+</sup> neurons relative to the total number of *Nkx2-2*<sup>+</sup> neurons, and that a subset of sleep-pressure responding *Lhx6* neurons express *Nkx2-2*. Scale bar = 100 μm. All bar graphs (**G**, **Q**, **R**) show mean and standard error of the mean (SEM), with individual data points plotted.

adulthood, raising the question of their potential physiological function. We next then performed 6 h sleep-deprivation, a robust method to detect cells that respond to sleep pressure, on *Lhx6-GFP* mice<sup>12</sup> and stained with antibodies to *Nkx2-2* and *cFos*. As shown previously, around 40% of *Lhx6*-expressing neurons in the ZI responded to sleep pressure, and around 35% of sleep pressure-activated neurons (~15% of all *Lhx6*-expressing neurons in the ZI) were *Nkx2-2*<sup>+</sup> (Fig. 7). In total, 25% of *Nkx2-2*<sup>+</sup> ZI neurons express *cFos* in response to increased sleep pressure. This indicates that *Nkx2.2* may guide the differentiation of a distinct subset of sleep-promoting ZI neurons.

## Discussion

The LIM homeodomain factor *Lhx6* is a master regulator of the differentiation and migration of GABAergic neurons of the cortex and hippocampus, as well as many other subcortical telencephalic structures such as striatum and amygdala. Over 70% of cortical interneurons express *Lhx6* into adulthood, where it is required for expression of canonical markers of interneuron subtype identity such as *Sst* and *Pvalb*<sup>27,32</sup>. In contrast, *Lhx6* is expressed in only 1–2% of hypothalamic GABAergic neurons. *Lhx6* expression is confined to a broad domain in the dorsolateral hypothalamus, and *Lhx6*-expressing cells do not undergo widespread long-distance tangential migration. *Lhx6*-expressing hypothalamic neurons in the ZI play an essential role in promoting sleep<sup>12</sup>, but their function is otherwise uncharacterized. In this study, we seek to characterize the development and molecular identity of hypothalamic *Lhx6*-expressing neurons, using previous knowledge obtained from studying telencephalic *Lhx6*-expressing neurons.

In the hypothalamus, in sharp contrast to the telencephalon, *Lhx6* is required to prevent neuronal apoptosis (Supplementary Fig. 16). The fact that loss of function of hypothalamic *Lhx6* leads to death of sleep-promoting neurons in the ZI may account for the more severe changes in sleep pattern that is seen in the hypothalamic-specific loss of function of *Lhx6* than is observed following DREADD-based manipulation of the activity of these neurons<sup>12</sup>. Analysis of *Lhx6/Bax* double mutants identified both the Neuregulin and *Gdnf* signaling pathways as potential neurotrophic mechanisms that promote the survival of hypothalamic *Lhx6* neurons. Interestingly, *Nrg1/ErbB4*-dependent signaling acts as a chemorepellent signal, while *Gdnf* signaling acts as a chemoattractant, and both regulate the long-range tangential migration of cortical *Lhx6* neurons<sup>29,48</sup>. Both signaling pathways may therefore have been at least partially repurposed to regulate cell survival in hypothalamic *Lhx6* neurons. The more modest phenotype seen following postnatal loss of function of *Lhx6*, relative to the constitutive mutant, may indicate that the survival of a specific subset of *Lhx6*-expressing neurons is no longer *Lhx6*-dependent at later ages.

We observe extensive transcriptional divergence between developing telencephalic and hypothalamic *Lhx6* neurons. Notably, we observe clear spatial differences in gene expression among hypothalamic *Lhx6* neurons that are not detectable in the MGE. While MGE cells require *Nkx2-1* to activate *Lhx6* expression, *Nkx2-1* is expressed primarily in the TT, in the posterior domain of hypothalamic *Lhx6* expression. The TT domain also expresses *Shh* similar to MGE that may regulate *Nkx2-1* expression<sup>18,49</sup>, leading to activation of *Lhx6* expression. However, we fail to observe any upstream gene expression (*Shh* or *Nkx2-1*) in MGE scRNA-Seq clusters when the downstream gene is detected (*Nkx2-1* or *Lhx6*)<sup>38</sup>, indicating *Nkx2-1* and *Lhx6* activation could lead to a shutdown of *Shh* and *Nkx2-1* in the MGE. In our hypothalamic *Lhx6*-expressing neurons, all three genes (*Shh*, *Nkx2-1*, and *Lhx6*) are highly co-expressed in the TT domain, unlike in the MGE.

*Dlx1/2* are expressed in virtually all *Lhx6*-expressing MGE cells but are not required to maintain *Lhx6* expression<sup>19,21</sup>, while *Dlx1/2* is primarily expressed in the ID domain in the hypothalamus. Furthermore, *Nkx2-2* is not expressed in the telencephalon but is selectively expressed in a previously uncharacterized hinge domain that connects the ID and TT. We find that mutants in *Nkx2-1*, *Nkx2-2*, and *Dlx1/2* selectively eliminate hypothalamic *Lhx6* expression in the TT, hinge, and ID domains, respectively. This indicates a high level of spatial patterning and transcriptional diversity among developing hypothalamic *Lhx6* neurons. Although hypothalamic *Lhx6* neurons do not undergo extensive tangential dispersal, as observed in telencephalon, lineage analysis indicates that by E18, a subset of neurons that express the TT-specific marker *Nkx2-1* have migrated to anterior structures such as the ZI. Combined with the observation that *Nkx2-2*-derived *Lhx6* neurons progressively disperse from the hinge domain into the ID implies that subsets of hypothalamic *Lhx6* neurons may undergo short-range migration during development.

*Lhx6* neurons in the postnatal hypothalamus are likewise highly transcriptionally diverse and do not directly correspond to any of their telencephalic counterparts (Supplementary Fig. 16). No hypothalamic *Lhx6* neurons express *Pvalb*, and only a few selected subsets express either *Sst* or *Npy*. In the cortex, many genes are exclusively expressed in *Lhx6*-expressing neurons—including *Sst*, *Pvalb*, and *Npy*. In contrast, in the hypothalamus, no genes were identified that were exclusively expressed in *Lhx6* neurons, other than *Lhx6* itself. Neuropeptides such as *Pnoc*, which are expressed in large subsets of hypothalamic *Lhx6* neurons, are also widely expressed in many neurons that do not express *Lhx6*. Finally, molecularly distinct subtypes of *Lhx6* neurons are broadly and evenly distributed in the cortex, owing to the widespread tangential dispersal during development. In contrast, in the hypothalamus, we observe clear differences in the expression of neuropeptides and calcium-binding proteins in *Lhx6* neurons that broadly correspond to the spatial position of these neurons.

These results provide a starting point to not only better define the molecular mechanisms that control differentiation, survival, and diversification of hypothalamic *Lhx6* neurons, but also serve as a molecular toolbox for selectively targeting molecularly distinct neuronal subtypes. Previous studies identified *Lhx6* neurons of the ZI as being unique in promoting both NREM and REM sleep<sup>12</sup>. Identification of molecular markers that distinguish different subtypes of *Lhx6* neurons in this region can help determine whether this is produced by the activation of distinct neuronal subtypes. We demonstrate that not only are a substantial fraction of *Lhx6* ZI neurons derived from *Nkx2.2*-expressing precursors, but that many also continue to express *Nkx2.2* into adulthood (Supplementary Fig. 16). Indeed, *Nkx2-2*<sup>+</sup> *Lhx6*-expressing ZI neurons represent 25% of *Lhx6* ZI neurons that express *c-fos* in response to elevated sleep pressure. Hypothalamic *Lhx6* neurons also send and receive connections from many brain regions that regulate innate behaviors, including the amygdala, periaqueductal gray, and ventral tegmental area<sup>12</sup>. The function of these circuits is as yet unknown, and the molecular markers identified in this study can serve as a starting point for investigating their behavioral significance.

## Methods

**Mice.** All experimental animal procedures were approved by the Johns Hopkins University Institutional Animal Care and Use Committee. All mice were housed in a climate-controlled facility (14 h dark and 10 h light cycle) with ad libitum access to food and water.

*Lhx6-GFP* (*Tg(Lhx6-EGFP)BP221Gsat*)<sup>50</sup>, *Lhx6*<sup>CreER</sup> knock-in (B6(Cg)-*Lhx6tm1(cre/ERT2)Zjh*), JAX #010776<sup>41</sup>, *Lhx6*<sup>lox/lox</sup><sup>51</sup>, *Ai9* (B6.Cg-Gt(*ROSA*)26Sortm9(CAG-tdTomato)Hze/J, JAX #007909)<sup>52</sup>, *Bax*<sup>lox/lox</sup> (B6;129-Baxtm2Sjk Bak1tm1Thsn/J, JAX #006329)<sup>25</sup>, *Nkx2-2*<sup>CreGFP</sup> (B6.129S6(Cg)-*Nkx2-2tm4.1(cre/EGFP)Suss*/J, JAX #026880)<sup>43</sup>, *Nkx2-1*<sup>CreER</sup> (*Nkx2-1tm1.1(cre/ERT2)Zjh*), JAX

#014552)<sup>41</sup>, *Foxd1<sup>Cre</sup>* (B6;129S4-*Foxd1tm1(GFP/cre)Amc/f*), JAX #012463)<sup>53</sup>, *Dlx1/2<sup>lox/lox</sup>* (*Dlx1tm1Rth Dlx2tm1.1Rth/f*), JAX #025612)<sup>45</sup>, *Dlx1/2<sup>-/-</sup>* (gift from J.L.R.R.) were used. Mice were time-mated and embryos at various ages (embryonic day (E)11.5, E12.5, E13.5, E15.5, E16.5, E17.5, E18.5, and postnatal day (P) 8) were collected for high-throughput sequencing and histology. Day of birth was considered as P0.

### Tamoxifen injection

***Lhx6<sup>CreER</sup>* pulse-chase experiments.** Pups with parental crosses of *Lhx6<sup>CreER/+</sup>*; *Lhx6<sup>lox/+</sup>*; *Bax<sup>lox/+</sup>*; *Ai9* (*Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/+</sup>*; *Ai9*) × *Lhx6<sup>CreER/+</sup>*; *Lhx6<sup>lox/+</sup>*; *Bax<sup>lox/+</sup>*; *Ai9* (*Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/+</sup>*; *Ai9*) were treated with intraperitoneal 4-Hydroxytamoxifen injection (4-OHT, 0.5 mg/per day, in corn-oil) for 5 consecutive days between P1 and P5. Pups were genotyped on the day of the birth and three different genotypes (1. *Lhx6<sup>CreER/+</sup>*; *Ai9*, 2. *Lhx6<sup>CreER/lox</sup>*; *Ai9*, 3. *Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9*) were used. *Lhx6<sup>CreER/+</sup>* genotype dies soon after weaning<sup>41</sup>, *Lhx6<sup>CreER/+</sup>*; *Lhx6<sup>lox/lox</sup>* genotype is not possible to generate due to similar sites of *CreER* and *lox* insertion.

Treated pups were collected between P40 and P45 and processed as described below. Cell counting was conducted in all three genotypes in the ZI, DMH, PH, S1 somatosensory cortex (CTX), and amygdala (AMY) following the Mouse Brain Atlas<sup>54</sup>. Borders were drawn to separate individual regions, using DAPI counterstaining and the Mouse Brain Atlas as a guideline, and 6500 μm × 500 μm region-of-interest was used to count across cortical layers per section. Three sections (every second section to avoid counting the same cell) were used per region, and six brains that were collected from between two and three individual litters (different parents) were used. tdTomato expression was observed in blood vessels as previously described<sup>12</sup>.

Three different classes of neurons were counted. The first class consists of neurons that only express Lhx6 protein as detected by immunostaining (indicating that no 4-OHT-induced Cre recombination occurred at the *Lhx6* locus). The second class consists of neurons that expressed the only tdTomato but not Lhx6 (indicating Cre-mediated activation of tdTomato, and disruption of *Lhx6*). The third class consists of neurons that expressed both tdTomato and Lhx6 (indicating incomplete 4-OHT-induced Cre recombination, with the induction of tdTomato expression and failure to recombine the conditional allele of *Lhx6*). Only neurons that expressed tdTomato (with or without Lhx6 protein expression) were counted and the total counted the number of neurons used as a denominator. Neurons that only expressed tdTomato were used as a numerator to calculate cell survival rate, as we expect to observe a decrease in the ratio (tdTomato<sup>+</sup>/(tdTomato<sup>+</sup> and tdTomato<sup>+</sup>/Lhx6<sup>+</sup>)) if *Lhx6* is required for cell survival.

***Nkx2-1<sup>CreER</sup>* pulse-chase experiments.** *Nkx2-1<sup>CreER/+</sup>*; *Ai9* female mice were time-mated to the same genotype male mice, and 4-OHT was intraperitoneally injected (2 mg) at E11.5, and embryos were collected at E18.5.

**Sleep deprivation.** Six-hour sleep-deprivation experiments were performed on *Lhx6-GFP* male mice as previously described<sup>12</sup>.

**Tissue fixation.** Embryos and mice younger than weaning age (P21) were fixed in 4% paraformaldehyde (PFA) between 8 and 12 h at 4 °C, incubated in 30% sucrose overnight at 4 °C, and snap-frozen in OCT compound for histology analysis. Whole embryos were used for fixation until E14.5, and from E14.5, brains were dissected out for fixation. Mice older than weaning age were anesthetized by intraperitoneal injection of avertin and perfused with cold 4% PFA. Brains were post fixed for 2 h at 4 °C with 4% PFA and processed as described above.

**Cryosectioning.** Frozen brains were sectioned at 25 μm with a cryostat (Leica CM3050S) along either the coronal or sagittal plane, and transferred to Superfrost™ Plus slides.

**In situ hybridization (ISH).** Chromogenic and fluorescent ISH was performed as previously described to stain for *Lhx6* (BC065077), *Gfra1* (AW060572), *Gfra2* (BE994145), *Ret* (AW123296), *Dlx1* (BC079609), *Calb1* (AW489595), *Calb2* (A1836013), *Gal* (BC044055), *Penk* (A1836252), *Tac1* (BE954293), *Npy* (A1848386), *Sst* (BE984677), *Th* (BF449409), *Gad1* (AW121495), *Nkx2-1* (BC080868), *Nkx2-2* (BG110), *Shh* (BC063087), *Prox1* (BE982394), *Six3* (BE953775), *Lhx8* (BE448496), and *Lef1* (BC038305)<sup>4,55</sup>. RNAscope with probe targeting *Lhx6* was tested on E13.5 mice following the manufacturer's protocol. Images were taken under the Keyence BZ-X800 fluorescence microscope or Zeiss LSM 700 microscope, and processed with ImageJ<sup>56</sup>, and pixel density was measured as previously described<sup>57</sup>.

**Immunostaining.** Immunostaining was performed with mouse-anti-Lhx6-antibody (1:200, sc-271433, Santa Cruz), rat-anti-RFP (1:500, ABIN334653, antibodies-online), rabbit-anti-Dlx1 (1:500, a gift from Jay Lee), guinea-pig-anti-Dlx1 (1:500, a gift from Jay Lee), rabbit-anti-Nkx2-1 (1:500, EP1584Y, Abcam), mouse-anti-Nkx2-2 (1:100, 74.5A5, DSHB), mouse-anti-NeuN (1:2000, MAB377), and rabbit-anti-cFos (1:1000, 226003, Synaptic Systems) as previously described<sup>57</sup>,

except that M.O.M blocking reagent (MKB-2213) was used following manufacturer's instruction when mouse primary antibodies were used. Alexa Fluor™ 488, 594, 647 secondary antibodies were used in 1:500 dilutions. Sections were mounted with DAPI-Vectamount (Vectorlabs) and imaged under a Keyence BZ-X800 fluorescence microscope or Zeiss LSM 700 microscope. All cell counting was done with ImageJ. Cell counting was conducted in multiple brain areas across developmental ages using standard reference atlases for orientation<sup>4,58</sup>, using DAPI counterstaining or NeuN staining as a guideline. For identification of ID and TT, criteria described in our previous study were used<sup>4</sup>. Three sections (every second section to avoid counting the same cell, <E15.5 = 2 sections) were used per region, and 4–6 brains collected from between two and three individual litters were used. Cell counting was conducted blinded.

### Bulk RNA-sequencing

***Lhx6<sup>CreER/+</sup>*; *Ai9* and *Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9* P1 pups** were treated with 4-OHT as described above and collected at P10. Between 4 and 6 pups from 2 different litters were pooled per sample without regard to sex, and papain-based enzymatic dissociation was performed on the dissected hypothalamus as previously described<sup>24</sup>. Dissociated cells were flow-sorted for tdTomato signal, and between 25,000 and 30,000 cells were collected directly into TRIZOL™ LS reagent. RNA was extracted using Direct-zol RNA kits (Zymo Research) and RNA-Sequencing libraries were made using stranded Total RNA-Seq library prep. 2 libraries were made for *Lhx6<sup>CreER/+</sup>*; *Ai9*, and three libraries were made for *Lhx6<sup>CreER/+</sup>*; *Lhx6<sup>lox/+</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9*. Libraries were sequenced with Illumina NextSeq500, paired-end read of 75 bp, 50 million reads per library. Illumina adapters of sequenced libraries were trimmed using Cutadapt (v1.18)/TrimGalore (v0.5.0)<sup>59</sup> with default parameters, library qualities were assessed using FastQC (v0.11.7)/MultiQC<sup>60</sup>. Libraries were then aligned to mm10 using STAR (v2.5.4b)<sup>61</sup> with `-twopassMode Basic`. RSEM (v1.3.0) was used for quantification<sup>62</sup>, with `semi-calculate-expression` (`-forward-prob 0.5`). Expected counts value from RSEM was used to perform differential expression using edgeR (v3.24.3)<sup>63</sup> using default parameters except for `calcNormFactors` (method = "TMM").

*Lhx6<sup>CreER/+</sup>*; *Ai9* or *Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9* enriched genes (fold change > 2 consistent gene value across replicates), were used with EnrichR<sup>64</sup>. *Lhx6<sup>lox/+</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9* enriched genes were compared to the Mouse Cells and Tissues (MESA) dataset available [ascot.cs.jhu.edu](http://ascot.cs.jhu.edu)<sup>26</sup>, relying on robustness of expression (NAUC >20) and specificity, as many of the enriched genes detected in this analysis are not strongly expressed in the developing brain.

We reasoned that the genes showing enriched expression in *Lhx6<sup>CreER/+</sup>*; *Ai9* relative to *Lhx6<sup>CreER/lox</sup>*; *Bax<sup>lox/lox</sup>*; *Ai9* would be regulated by *Lhx6* and/or *Bax*. Furthermore, since tdTomato expression is detected in blood vessels due to weak Lhx6 expression in endothelial neurons during development<sup>12</sup>, we wanted to enrich expression from Lhx6-expressing neurons of the hypothalamus. P8 *Lhx6-GFP*, in which GFP expression is absent in endothelial cells, was used to generate bulk RNA-sequencing (bulk RNA-Seq) from the cortex and hypothalamus (method described below). Hypothalamus-enriched genes from P8 *Lhx6-GFP* bulk RNA-Seq data were used to enrich genes that are highly expressed in the hypothalamus Lhx6 neurons. After enrichment, the gene lists were compared to scRNA-Seq data from P8 *Lhx6-GFP* hypothalamus using the method described below, to further cross-check specificity of expression and to remove any possible contamination that may occur during flow sorting from bulk RNA-Seq. EnrichR was used to identify gene pathways, and pathways previously implicated in the regulation of neuronal survival were selected.

***Lhx6-GFP* bulk RNA-seq.** To identify differences between cortical and hypothalamic Lhx6 populations, RNA-Sequencing was performed on E15.5, P0, and P8 *Lhx6-GFP* mice, by collecting 8–10 pups from two different litters per library. Libraries were sequenced with Illumina HiSeq 2500, and processed as described in the pipeline described above.

**ATAC-sequencing.** Cortex and hypothalamus of E15.5 and P0 *Lhx6-GFP* mice were collected, dissociated with papain-based enzymatic reaction, and GFP neurons were flow-sorted. Between 60,000 and 70,000 neurons were collected. Flow-sorted neurons were prepared for ATAC libraries as previously described<sup>65,66</sup>. Libraries were sequenced with Illumina NextSeq500, paired-end read of 75 bp, 50 million reads per library. Each sample was run in duplicate.

Illumina adapters of sequenced libraries were trimmed using Cutadapt (v1.18)/TrimGalore (v0.5.0) and library qualities were assessed using FastQC (v0.11.7)/MultiQC. Libraries were aligned to mm10 using Bowtie 2 (v2.25)<sup>67</sup> using—very-sensitive parameter and Samtools (v1.9)<sup>68</sup> was used to check the percentage of mitochondria DNA reads. Picard (v2.18) was used to remove PCR duplicates, and MACS2 (v2.1.2)<sup>69</sup> was used to capture narrow peaks (open chromatin regions) with `-shift 100`, `-extsize 200`, `-nolambda`, `-nomodel` parameters. ENCODE blacklist regions of the genome were removed using Bedtools (v2.27) intersect function<sup>69–71</sup>. Bedtools intersect function was used to find matching peaks between replicates, in which the distance between peak ends was <10 base pairs. ChIPseeker (v1.18.0)<sup>72</sup> was then used to identify regions that were within 3 kb of the transcription start site (TSS). Peaks between groups were compared as previously described<sup>65,66</sup> to visualize changes in chromatin accessibility between different ages

and brain regions using DiffBind (v.2.10.0)<sup>73</sup> and edgeR using default parameters (FDR <0.05 and adjusted *p* value <0.05). Differential peaks were compared to bulk RNA-Seq, and open chromatin peaks in promoter regions that correspond to altered gene expression from bulk RNA-Seq were identified<sup>65,66</sup> to obtain a positive correlation between promoter accessibility and gene expression. Peaks and differential gene expression was then cross-matched to scRNA-Seq, to identify potential different regions within Lhx6 hypothalamic neurons that are demarcated by expression of specific transcription factors.

**Single-cell RNA-sequencing.** Time-mated E12.5, E15.5, and P8 *Lhx6-GFP* mice were collected, and dissection and dissociation were performed as described previously<sup>24</sup>. Between six and ten embryos/pups from two different litters were collected. Following dissociation, GFP<sup>+</sup> neurons were flow-sorted using Aria IIu Sorter (BD). Between 20,000 and 25,000 neurons were flow-sorted for E12.5 and E15.5, 2000 neurons were flow-sorted for P8. Flow-sorted neurons were used for the 10× Genomics Chromium Single Cell System (10× Genomics, CA, USA) using V3.0 chemistry per manufacturer’s instruction. Three libraries were sequenced on Illumina NextSeq 500 with ~200 million reads per library. Sequenced files were processed through the Cell Ranger pipeline (v3.1.0, 10× Genomics) using mm10 genome.

Seurat V3<sup>74</sup> was used to perform downstream analysis following the standard pipeline described previously<sup>75</sup>, analyzing neurons that express a high *Lhx6* transcript. Louvain algorithm was used to generate different clusters, and spatial information from individual clusters at E12.5 and E15.5 was identified by referring to our previous hypothalamus scRNA-Seq database HyDD<sup>24</sup>, as well as previous analysis of anatomical locations of transcription factors<sup>4</sup>. For P8 scRNA-Seq, region-specific transcription factors that are expressed were compared to E12.5 and E15.5 scRNA-Seq gene lists, as well as matching the identified gene lists to the Allen Brain Atlas ISH data<sup>58</sup>. Previously published scRNA-Seq from E13.5 MGE<sup>38</sup> was processed as described above, and the key markers that label individual clusters were compared to E12.5 *Lhx6*-expressing hypothalamic neurons.

*Lhx6*<sup>+</sup> neurons across multiple mutant groups (*Foxd1*<sup>Cre/+</sup>/*Dlx1*<sup>2lox/lox</sup>, *Nkx2-1*<sup>CreER</sup>/*CreER*, *Nkx2-2*<sup>CreGFP</sup>/*CreGFP*) from<sup>24</sup>, were used to compare the expression level of key transcription factors that define sub-regions of hypothalamic *Lhx6* expression domains.

Previously generated scRNA-Seq datasets from the preoptic region<sup>76</sup>, suprachiasmatic nucleus<sup>76,77</sup>, VMH<sup>78</sup>, and whole hypothalamus<sup>24,79–81</sup>, were analyzed as described above. GABAergic neurons (*Slc32a1*<sup>+</sup>) were first subsetted from the dataset, and the percentage of neurons expressing *Pnoc*, *Penk*, *Calb1*, *Cck*, *Calb2*, *Gal*, *Tac1*, *Th*, *Npy*, *Trh*, *Sst* was determined.

RNA velocity<sup>35</sup> was used to understand the dynamic state of *Lhx6* neuronal development, and RNA velocity was used to identify (1) how *Lhx6*-expressing domains are established during development and (2) the origin of E15.5 *Nkx2-2*<sup>+</sup> cluster. Kallisto and Bustools<sup>82,83</sup> was used to obtain spliced and unspliced transcripts using --lamanno with GRCm38 mouse genome. Scanpy<sup>84</sup> and scVelo<sup>85</sup> was used to process the Kallisto output with default parameters, based on UMAP coordinates obtained from Seurat.

To identify regulatory transcription factors controlling gene expression in different *Lhx6*-expressing domains, SCENIC<sup>86,87</sup> (python implemented pySCENIC (using --masks\_dropouts)), was used to calculate regulatory transcription factors using default parameters with mm10 feather files on scRNA-Seq dataset using raw count matrix. This workflow involves three steps. This workflow involves three steps. First, we identify potential transcription factor targets in each cluster based on the co-expression of genes. Second, we perform transcription factor motif enrichment analysis and identify potential key regulatory transcription factors. Finally, we score the activity of these regulatory transcription factors based on the network of co-expressed genes.

**Statistics.** Two-way ANOVA was used for the *Lhx6* pulse-chase experiments in Fig. 2 (genotype, brain region). Unpaired *t* test was used for all other cell counting studies. The Seurat “FindAllMarkers” function with “LR = logistic regression model” with default parameters was used for analyzing differential gene expression, using the number of total mRNAs and genes as a variable. All bar graphs show mean and standard error of the mean (SEM), with individual data points plotted.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All sequencing data are available on GEO, GSE150687. All differential gene lists of bulk/scRNA-Seq and differential peak lists of bulk ATAC-Seq are provided in Supplementary Data 1–9. All other data associated with the paper are available upon request.

Received: 13 August 2020; Accepted: 17 December 2020;

Published online: 21 January 2021

## References

- Lim, L., Mi, D., Llorca, A. & Marin, O. Development and functional diversification of cortical interneurons. *Neuron* **100**, 294–313 (2018).
- Huang, Z. J. & Paul, A. The diversity of GABAergic neurons and neural communication elements. *Nat. Rev. Neurosci.* **20**, 563–572 (2019).
- Erő, C., Gewaltig, M.-O., Keller, D. & Markram, H. A cell atlas for the mouse brain. *Front. Neuroinform.* **12**, 28 (2018).
- Shimogori, T. et al. A genomic atlas of mouse hypothalamic development. *Nat. Neurosci.* **13**, 767–775 (2010).
- Bedont, J. L., Newman, E. A. & Blackshaw, S. Patterning, specification, and differentiation in the developing hypothalamus. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 445–468 (2015).
- Colasante, G. et al. *Arx* is a direct target of *Dlx2* and thereby contributes to the tangential migration of GABAergic interneurons. *J. Neurosci.* **28**, 10674–10686 (2008).
- Cobos, I. et al. Mice lacking *Dlx1* show subtype-specific loss of interneurons, reduced inhibition and epilepsy. *Nat. Neurosci.* **8**, 1059–1068 (2005).
- Bulfone, A. et al. Spatially restricted expression of *Dlx-1*, *Dlx-2* (Tes-1), *Gbx-2*, and *Wnt-3* in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries. *J. Neurosci.* **13**, 3155–3172 (1993).
- Bedont, J. L. et al. An LHX1-regulated transcriptional network controls sleep/wake coupling and thermal resistance of the central circadian clockworks. *Curr. Biol.* **27**, 128–136 (2017).
- Bedont, J. L. et al. *Lhx1* controls terminal differentiation and circadian function of the suprachiasmatic nucleus. *Cell Rep.* **7**, 609–622 (2014).
- Hatori, M. et al. *Lhx1* maintains synchrony among circadian oscillator neurons of the SCN. *eLife* **3**, e03357 (2014).
- Liu, K. et al. *Lhx6*-positive GABA-releasing neurons of the zona incerta promote sleep. *Nature* **548**, 582–587 (2017).
- Maroof, A. M., Brown, K., Shi, S. H., Studer, L. & Anderson, S. A. Prospective isolation of cortical interneuron precursors from mouse embryonic stem cells. *J. Neurosci.* **30**, 4667–4675 (2010).
- Kessarlis, N., Magno, L., Rubin, A. N. & Oliveira, M. G. Genetic programs controlling cortical interneuron fate. *Curr. Opin. Neurobiol.* **26**, 79–87 (2014).
- Wang, Y. et al. *Dlx5* and *Dlx6* regulate the development of parvalbumin-expressing cortical interneurons. *J. Neurosci.* **30**, 5334–5345 (2010).
- Du, T., Xu, Q., Ocbina, P. J. & Anderson, S. A. NKX2.1 specifies cortical interneuron fate by activating *Lhx6*. *Development* **135**, 1559–1567 (2008).
- Fogarty, M. et al. Spatial genetic patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult cortex. *J. Neurosci.* **27**, 10935–10946 (2007).
- Flandin, P. et al. *Lhx6* and *Lhx8* coordinately induce neuronal expression of *Shh* that controls the generation of interneuron progenitors. *Neuron* **70**, 939–950 (2011).
- Sandberg, M. et al. Transcriptional networks controlled by NKX2-1 in the development of forebrain GABAergic neurons. *Neuron* **91**, 1260–1275 (2016).
- Batista-Brito, R. et al. The cell-intrinsic requirement of *Sox6* for cortical interneuron development. *Neuron* **63**, 466–481 (2009).
- Zhao, Y. et al. Distinct molecular pathways for development of telencephalic interneuron subtypes revealed through analysis of *Lhx6* mutants. *J. Comp. Neurol.* **510**, 79–99 (2008).
- Liodis, P. et al. *Lhx6* activity is required for the normal migration and specification of cortical interneuron subtypes. *J. Neurosci.* **27**, 3078–3089 (2007).
- Vogt, D. et al. *Lhx6* directly regulates *Arx* and *CXCR7* to determine cortical interneuron fate and laminar position. *Neuron* **82**, 350–364 (2014).
- Kim, D. W. et al. The cellular and molecular landscape of hypothalamic patterning and differentiation. *bioRxiv* <https://doi.org/10.1101/657148> (2019).
- Takeuchi, O. et al. Essential role of BAX, BAK in B cell homeostasis and prevention of autoimmune disease. *Proc. Natl Acad. Sci. U.S.A.* **102**, 11272–11277 (2005).
- Ling, J. P. et al. ASCOT identifies key regulators of neuronal subtype-specific splicing. *Nat. Commun.* **11**, 137 (2020).
- Denaxa, M. et al. Maturation-promoting activity of SATB1 in MGE-derived cortical interneurons. *Cell Rep.* **2**, 1351–1362 (2012).
- Flames, N. et al. Short- and long-range attraction of cortical GABAergic interneurons by neuregulin-1. *Neuron* **44**, 251–261 (2004).
- Li, H., Chou, S.-J., Hamasaki, T., Perez-Garcia, C. G. & O’Leary, D. D. M. Neuregulin repellent signaling via ErbB4 restricts GABAergic interneurons to migratory paths from ganglionic eminence to cortical destinations. *Neural Dev.* **7**, 1–17 (2012).
- Bartolini, G. et al. Neuregulin 3 mediates cortical plate invasion and laminar allocation of GABAergic interneurons. *Cell Rep.* **18**, 1157–1170 (2017).
- Mei, L. & Xiong, W.-C. Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. *Nat. Rev. Neurosci.* **9**, 437–452 (2008).

32. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
33. Mickelsen, L. E. 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).
34. Rossi, M. A. et al. Obesity remodels activity and transcriptional state of a lateral hypothalamic brake on feeding. *Science* **364**, 1271–1274 (2019).
35. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
36. Wonders, C. P. et al. A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. *Dev. Biol.* **314**, 127–136 (2008).
37. Stenman, J. M., Wang, B. & Campbell, K. Tlx controls proliferation and patterning of lateral telencephalic progenitor domains. *J. Neurosci.* **23**, 10568–10576 (2003).
38. Mayer, C. et al. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).
39. Anderson, S. A. et al. Mutations of the homeobox genes *Dlx-1* and *Dlx-2* disrupt the striatal subventricular zone and differentiation of late born striatal neurons. *Neuron* **19**, 27–37 (1997).
40. Shimamura, K., Hartigan, D. J., Martinez, S., Puelles, L. & Rubenstein, J. L. Longitudinal organization of the anterior neural plate and neural tube. *Development* **121**, 3923–3933 (1995).
41. Taniguchi, H. et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995–1013 (2011).
42. Marín, O., Baker, J., Puelles, L. & Rubenstein, J. L. R. Patterning of the basal telencephalon and hypothalamus is essential for guidance of cortical projections. *Development* **129**, 761–773 (2002).
43. Balderes, D. A., Magnuson, M. A. & Sussel, L. *Nkx2.2:Cre* knock-in mouse line: A novel tool for pancreas- and CNS-specific gene deletion. *Genesis* **51**, 844–851 (2013).
44. Qiu, M. et al. Role of the *Dlx* homeobox genes in proximodistal patterning of the branchial arches: mutations of *Dlx-1*, *Dlx-2*, and *Dlx-1* and *-2* alter morphogenesis of proximal skeletal and soft tissue structures derived from the first and second arches. *Dev. Biol.* **185**, 165–184 (1997).
45. Silbereis, J. C. et al. *Olig1* function is required to repress *dlx1/2* and interneuron production in Mammalian brain. *Neuron* **81**, 574–587 (2014).
46. Newman, E. A., Wu, D., Taketo, M. M., Zhang, J. & Blackshaw, S. Canonical Wnt signaling regulates patterning, differentiation and nucleogenesis in mouse hypothalamus and prethalamus. *Dev. Biol.* **442**, 236–248 (2018).
47. Salvatierra, J. et al. The LIM homeodomain factor *Lhx2* is required for hypothalamic tanyocyte specification and differentiation. *J. Neurosci.* **34**, 16809–16820 (2014).
48. Pozas, E. & Ibáñez, C. F. GDNF and GFR $\alpha$ 1 promote differentiation and tangential migration of cortical GABAergic neurons. *Neuron* **45**, 701–713 (2005).
49. Abecassis, Z. A. et al. *Npas1+–Nkx2.1+* neurons are an integral part of the cortico-pallido-cortical Loop. *J. Neurosci.* **40**, 743–768 (2020).
50. Gong, S. et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917–925 (2003).
51. Denaxa, M. et al. Modulation of apoptosis controls inhibitory interneuron number in the cortex. *Cell Rep.* **22**, 1710–1721 (2018).
52. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
53. Humphreys, B. D. et al. Fate tracing reveals the pericyte and not epithelial origin of myofibroblasts in kidney fibrosis. *Am. J. Pathol.* **176**, 85–97 (2010).
54. Watson, C. & Paxinos, G. *Chemoarchitectonic Atlas of the Mouse Brain*. (Academic Press, 2010).
55. Miranda-Angulo, A. L., Byerly, M. S., Mesa, J., Wang, H. & Blackshaw, S. Rax regulates hypothalamic tanyocyte differentiation and barrier function in mice. *J. Comp. Neurol.* **522**, 876–899 (2014).
56. Rueden, C. T. et al. ImageJ2: imageJ for the next generation of scientific image data. *BMC Bioinform.* **18**, 529 (2017).
57. Kim, D. W., Glendinning, K. A., Grattan, D. R. & Jasoni, C. L. Maternal obesity in the mouse compromises the blood-brain barrier in the arcuate nucleus of offspring. *Endocrinology* **157**, 2229–2242 (2016).
58. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
59. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10 (2011).
60. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
61. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
62. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
63. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
64. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
65. Wang, J. et al. ATAC-Seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. *Nat. Commun.* **9**, 1364 (2018).
66. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive multimodal analysis of chromatin architecture. *Biophys. J.* **106**, 77a (2014).
67. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
68. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
69. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
70. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
71. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).
72. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
73. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
74. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019). e21.
75. Bell, B. J. et al. A clock-driven neural network critical for arousal. *bioRxiv*. <https://doi.org/10.1101/2020.03.12.989921> (2020).
76. Moffitt, J. R. et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018).
77. Wen, S. 'ang et al. Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nat. Neurosci.* **23**, 456–467 (2020).
78. Kim, D.-W. et al. Multimodal analysis of cell types in a hypothalamic node controlling social behavior. *Cell* **179**, 713–728.e17 (2019).
79. Romanov, R. A. et al. Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
80. Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-cell RNA-seq reveals hypothalamic cell diversity. *Cell Rep.* **18**, 3227–3241 (2017).
81. Campbell, J. N. et al. A molecular census of arcuate hypothalamus and median eminence cell types. *Nat. Neurosci.* **20**, 484–496 (2017).
82. Melsted, P. et al. Modular and efficient pre-processing of single-cell RNA-seq. *bioRxiv*. <https://doi.org/10.1101/673285>.
83. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
84. Alexander Wolf, F., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
85. Bergen, V., Lange, M., Peidli, S., Alexander Wolf, F. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
86. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
87. Kim, D.W. et al. Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic *Lhx6*-expressing GABAergic neurons. *Gene Expression Omnibus. Series GSE150687*. (2020).

## Acknowledgements

This work was supported by a grant from the NIH (R01DK108230) to S.B., the Maryland Stem Cell Research Fund (2019-MSCRFF-5124) to D.W.K., and the Japan Society for the Promotion of Science to T.S. We thank Transcriptomics and Deep Sequencing Core (Johns Hopkins) for sequencing of bulk RNA-Seq, bulk ATAC-Seq and scRNA-Seq libraries, and Ross Flow Cytometry Core (Johns Hopkins) for flow sorting of *Lhx6-GFP* cells, and Microscope facility (Johns Hopkins MICFAC, supported by the award number S10OD018118). We thank Marysia Placzek and Wendy Yap for comments on the paper.

## Author contributions

S.B. conceived the study. D.W.K., K.L. and S.B. designed experiments. D.W.K., K.L., Z.Q.W., S.Z., A.B., M.P.B., S.H.L., P.W.W. and T.S. performed experiments. D.W.K., K.L., Z.Q.W., S.Z., P.W.W., C.S. and T.S. analyzed data. B.L. and J.L.R. provided reagents. All authors contributed to writing the paper.



**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s42003-020-01616-7>.

**Correspondence** and requests for materials should be addressed to S.B.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021