Article

Molecular design of hypothalamus development

Roman A. Romanov[1,2,12], Evgenii O. Tretiakov[1,12], Maria Eleni Kastriti[1,3], Maja Zupancic[1], Martin Häring[1], Solomiia Korchynska[1], Konstantin Popadin[4,5], Marco Benevento[1], Patrick Rebernik[1], Francois Lallemend[2], Katsuhiko Nishimori[6], Frédéric Clotman[7], William D. Andrews[8], John G. Parnavelas[8], Matthias Farlik[9,10], Christoph Bock[9,11], Igor Adameyko[1,3], Tomas Hökfelt[2], Erik Keimpema[1,13] & Tibor Harkany[1,2,13 ✉]

A wealth of specialized neuroendocrine command systems intercalated within the hypothalamus control the most fundamental physiological needs in vertebrates[1,2]. Nevertheless, we lack a developmental blueprint that integrates the molecular determinants of neuronal and glial diversity along temporal and spatial scales of hypothalamus development[3]. Here we combine single-cell RNA sequencing of 51,199 mouse cells of ectodermal origin, gene regulatory network (GRN) screens in conjunction with genome-wide association study-based disease phenotyping, and genetic lineage reconstruction to show that nine glial and thirty-three neuronal subtypes are generated by mid-gestation under the control of distinct GRNs. Combinatorial molecular codes that arise from neurotransmitters, neuropeptides and transcription factors are minimally required to decode the taxonomical hierarchy of hypothalamic neurons. The differentiation of γ-aminobutyric acid (GABA) and dopamine neurons, but not glutamate neurons, relies on quasi-stable intermediate states, with a pool of GABA progenitors giving rise to dopamine cells[4]. We found an unexpected abundance of chemotropic proliferation and guidance cues that are commonly implicated in dorsal (cortical) patterning[5] in the hypothalamus. In particular, loss of SLIT–ROBO signalling impaired both the production and positioning of periventricular dopamine neurons. Overall, we identify molecular principles that shape the developmental architecture of the hypothalamus and show how neuronal heterogeneity is transformed into a multimodal neural unit to provide virtually infinite adaptive potential throughout life.

A kaleidoscope of neuroendocrine cell modalities is concentrated into a minimal brain volume within the hypothalamus by using sometimes only thousands of neurons to encode essential hormonal output. Therefore, diversification of neuronal subtypes, rather than the numerical expansion of single progenies[6,7], might underpin the success of vertebrate evolution in refining metabolic and adaptive capacity. Functional versatility at the level of individual neuroendocrine output neurons is coded by combinations of neurotransmitters and neuropeptides[1]. Therefore, interrogation of the molecular and positional diversity of hypothalamic neurons by morphological, circuit and endocrine analyses continues to mount a substantial challenge. The introduction of single-cell RNA sequencing (scRNA-seq)[6,8,9] has provided precise molecular insights into the existence of glutamate, GABA, dopamine and even 'mixed' neuronal phenotypes[4]. However, the question of

how cellular subtypes emerge, migrate, and differentiate during the development of the hypothalamus in order to achieve neuroendocrine readiness by birth remains relatively unexplored (but see refs. [3,10,11]).

Whereas a handful of transcription factors (TFs) are sufficient to mark anatomical footprints in cortical structures with a layered organization[6,8], the intercalated nature of nuclei makes it more challenging to establish an anatomical template within the hypothalamus. In particular, the breadth of endocrine command neurons and their ability to rapidly switch cell state (that is, to upregulate specific hormones or neuropeptides in an 'on-demand' fashion) suggest that what is considered terminally differentiated in the adult brain is in fact a neuronal 'anagram' that is primarily dictated by the neuronal circuit that orchestrates a specific endocrine modality. Therefore, we sought to identify the molecular determinants of ectodermal progenies as they advance

[1]Department of Molecular Neurosciences, Center for Brain Research, Medical University of Vienna, Vienna, Austria. [2]Department of Neuroscience, Biomedicum D7, Karolinska Institutet, Solna, Sweden. [3]Department of Physiology and Pharmacology, Biomedicum D6, Karolinska Institutet, Solna, Sweden. [4]Human Genomics of Infection and Immunity, School of Life Sciences, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland. [5]Center for Mitochondrial Functional Genomics, Institute of Living Systems, Immanuel Kant Baltic Federal University, Kaliningrad, Russia. [6]Deptartment of Obesity and Internal Inflammation, Fukushima Medical University, Fukushima City, Japan. [7]Laboratory of Neural Differentiation, Institute of Neuroscience, Université Catholique de Louvain, Brussels, Belgium. [8]Department of Cell and Developmental Biology, University College London, London, UK. [9]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [10]Department of Dermatology, Medical University of Vienna, Vienna, Austria. [11]Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria. [12]These authors contributed equally: Roman A. Romanov, Evgenii O. Tretiakov. [13]These authors jointly supervised the work: Erik Keimpema, Tibor Harkany. ✉e-mail: tibor.harkany@meduniwien.ac.at
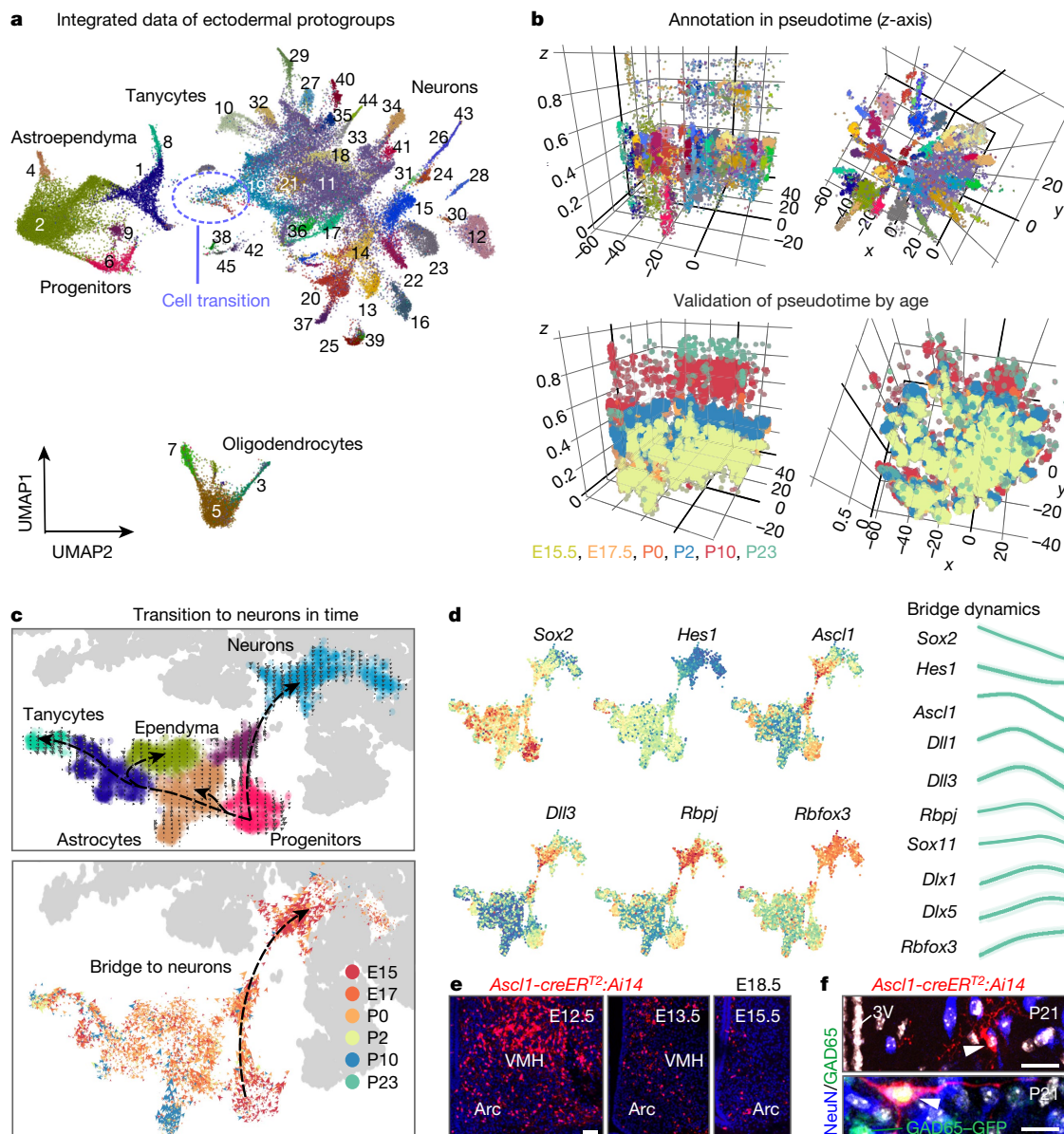
Fig. 1 | Developmental diversification of hypothalamic cell lineages.
a, UMAP plot of 51,199 cells of ectodermal origin, integrated by canonical correlation analysis (CCA) to achieve a hypothetical continuum that reflects the progressive attainment of cell identities. Walktrap in iGRAPH distinguished non-mature cells (clusters 11, 19) and neurons (31 proto-groups) at the end of each developmental trajectory. **b**, Schemes illustrating the conformity of alignment and clustering in pseudotime (z-axis, calculated independently, top) and biological age (bottom). **c**, UMAP on PAGA representation of progenitors, glia and immature neurons (bridge cells). RNA velocity[12,13] transformed multi-dimensional PAGA data into developmental trajectories. The colours of junctions accord with groups (top) and age (bottom). Note that a 'cell bridge' that links progenitors and immature neurons encompasses cells of early developmental stages, even though all time points are minimally represented therein. **d**, Imputed expression for the cell groups shown in **c** and pseudotime trajectories of differentiation into neurons starting from a $Sox2^+$ state. **e**, Genetic tracing of $Ascl1^+$ progenitors (induction at successive time points) in VMH and Arc. **f**, $Ascl^+$ progenitor-derived neurons (arrowheads) generated postnatally. Scale bars, 65 μm (**e**), 20 μm (**f**).

towards terminal neuroendocrine differentiation. By using a time series of scRNA-seq data across critical periods of intrauterine and postnatal hypothalamus development in mouse, we read out combinatorial codes for GABA, GABA-derived dopamine and glutamate neurons, catalogued GRNs (regulons) and their dynamic transitions during neurogenesis, directional migration and morphogenesis, and identified local chemotropic cues that define the anatomical constraints of the hypothalamus.

Emergence of ectoderm-derived cell pools

We addressed the differentiation programs for hypothalamic cell pools by parallel scRNA-seq on 51,199 dissociated cells at embryonic days (E)15.5 (n = 8,290 cells) and E17.5 (11,213), at birth (7,492), and at postnatal days (P)2 (12,824), P10 (8,965) and P23 (2,415; Online Methods and Supplementary Note). Overall, proto-groups of progenitors (2), tanycytes (2), astroependymal cells (2), immature oligodendrocytes (3), cells of the pars tuberalis (3) and neurons (33; Fig. 1a), reflecting diversity in the adult hypothalamus[4,8,9], were specified by differentially expressed TFs (Extended Data Fig. 1) during development (Fig. 1b).

We then investigated when and by which progenitors the various cell types are generated. The dynamics of gene expression in hypothalamic progenitors (Fig. 1c) to produce astrocytes, ependyma, tanycytes and neurons fit a pseudotime scale on a multidimensional integrated dataset[12] (Fig. 1c, d), including a bifurcation in cell transition towards
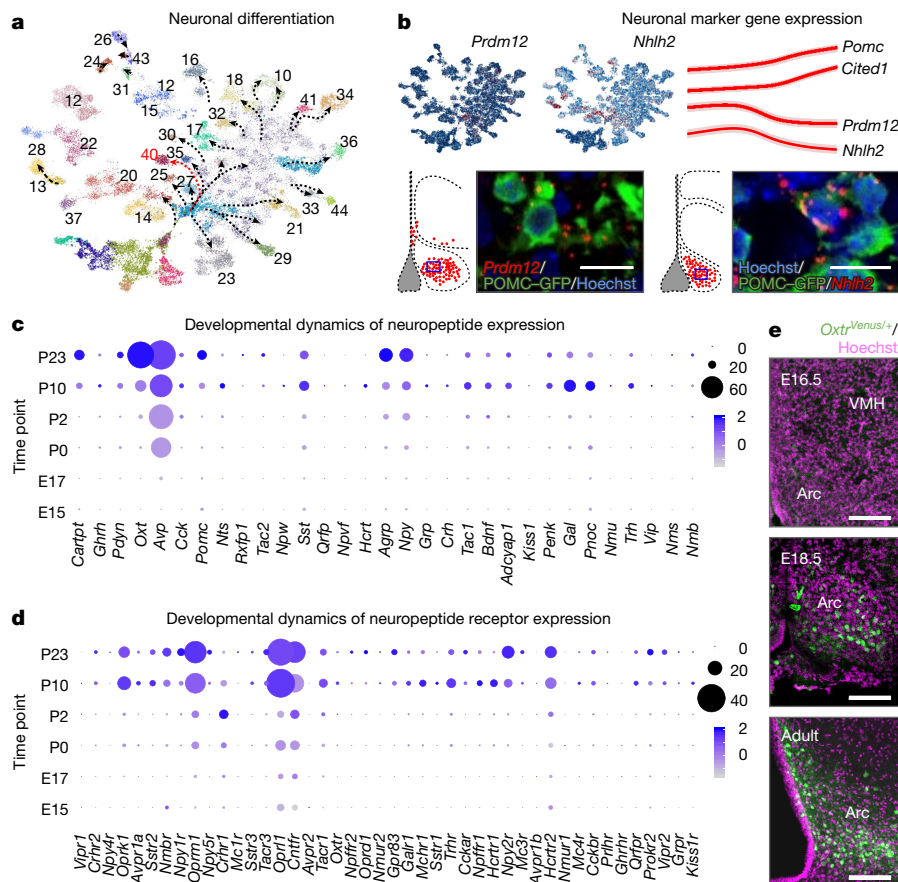
Fig. 2 | Neuronal differentiation in the hypothalamus. a, Cellular clusters from Fig. 1a (without clusters 38, 42, 45 and oligodendrocytes) represented as a graph-like map upon transforming UMAP embedding with the PAGA method[12] to assess cell differentiation trajectories. Red dotted line specifies the trajectory for *Pomc* (cluster 40) neurons. **b**, *Prdm12* and *Nhlh2* expression (top left) and their developmental dynamics (in pseudotime) relative to *Pomc* and *Cited1*, a transcriptional co-activator that specifies neurons of Arc (top right). Data in pseudotime were scaled[6]. Expression of *Prdm12* and *Nhlh2* in *Pomc*+ neurons was validated by in situ hybridization in *Pomc-Gfp* mice (bottom). Blue rectangles in topographical maps show the locations of images at single-cell resolution. **c**, **d**, Dynamics of gene expression for neuropeptides (**c**) and their receptors (**d**) during hypothalamus development. Data shown as dot plots. **e**, Developmental mapping of hypothalamic *Oxtr* expression in *Oxtr*Venus/+ mice. Scale bars, 12 μm (**b**), 200 μm (**e**).

glial subtypes or neuronal fates (Fig. 1a, c, Extended Data Fig. 2a, b) that peaked between E15.5 and E17.5 (Fig. 1c, bottom). RNA velocity[13] (as well as PAGA[12]; see Supplementary Note) demonstrated that the number of neuroblasts ('bridge cells') tailed off as a factor of age with an appreciable rupture of this cell continuum by birth (Extended Data Fig. 2b). Semi-supervised analysis of 327 genes for enrichment highlighted that the progression of bridge cells relied on the dominance of genes related to the regulation of pluripotency (*Sox2*), neural stem cell differentiation (*Hes1*, *Ascl1*, *Rbpj*, *Dll1* and *Dll3* for Notch signalling)[14,15], Erk signalling (*Sox11*), neuronal migration (*Dlx1*, *Dlx2*, *Dlx5*, *Dlx6*; subsequently referred to as *Dlx1/2/5/6*) and morphogenesis (*Rbfox3*; Fig. 1d)[16,17]. Additionally, scRNA-seq suggested the existence of an alternative and embryonically restricted (<E15.5) pathway of hypothalamic neurogenesis that centred on *Tbr1*+/*Eomes*+ progenitors that reside at the thalamus–hypothalamus dorsal boundary (an *Ascl1*− territory) and contribute multiple diencephalic neuronal subtypes (Extended Data Fig. 2d, e). The proposed waves of neurogenesis by self-renewing progenitors and early neuroblasts that ubiquitously express *Ascl1* along the third ventricle (Fig. 1d) were shown in *Ascl1-creER*T2::*Ai14* mice at E18.5 with recombination induced during the E12.5–E16.5 period (Fig. 1e) and validated in *Ascl1*−/− mice presenting a restricted cohort of *Sox2*+ immature precursors (Extended Data Fig. 2f). *tdTomato*+ progenies in the many hypothalamic subregions confirmed neurogenesis during mid-gestation with a gradual decline after E16.5 (Fig. 1e, Extended Data Fig. 2d). In support of postnatal neurogenesis, *Sox2*+ precursors persisted in the wall of the third ventricle and generated progenies that progressed through *Ascl1*+ and *Rbfox3*+ (NeuN) stages (Fig. 1f, Extended Data Fig. 2g, h).

Intermediate states for GABA neurons

Within our integrated dataset, about 47% of all cells committed to the neuronal lineage were in immature states (clusters 11, 19; Figs. 1a, 2a) before progressing towards final differentiation, as suggested by the expression of homeobox genes that are thought to determine GABA identities (cluster 11, *Foxg1* and *Nkx2-3*; number 19, *Sox2*, *Sox11*, *Gsx1*, *Gsx2*, *E2f1*, *Arx* and *Pbx3*; Extended Data Fig. 1). Specifically, cluster 19 contains a continuum of neuroblasts (bridge cells) with a normalized contribution of 36.5% (E15.5), 34.6% (E17.5), 13.2% (birth), 9.2% (P2), 4.8% (P10) and 1.6% (P23), and remains separated from other clusters despite our re-partitioning efforts (Supplementary Note). Cluster 11 is composed of immature neurons that are largely homogeneous with low-level differential gene expression, but express rate-limiting enzymes and transporters for GABA neurotransmission (Extended Data Figs. 1a, 3a). When re-partitioning these data, immature GABA neurons were re-assigned to phenotypically stable groups, emphasizing the intermediate nature of cluster 11. These findings contrast glutamate neurons, which immediately appear as differentiated and
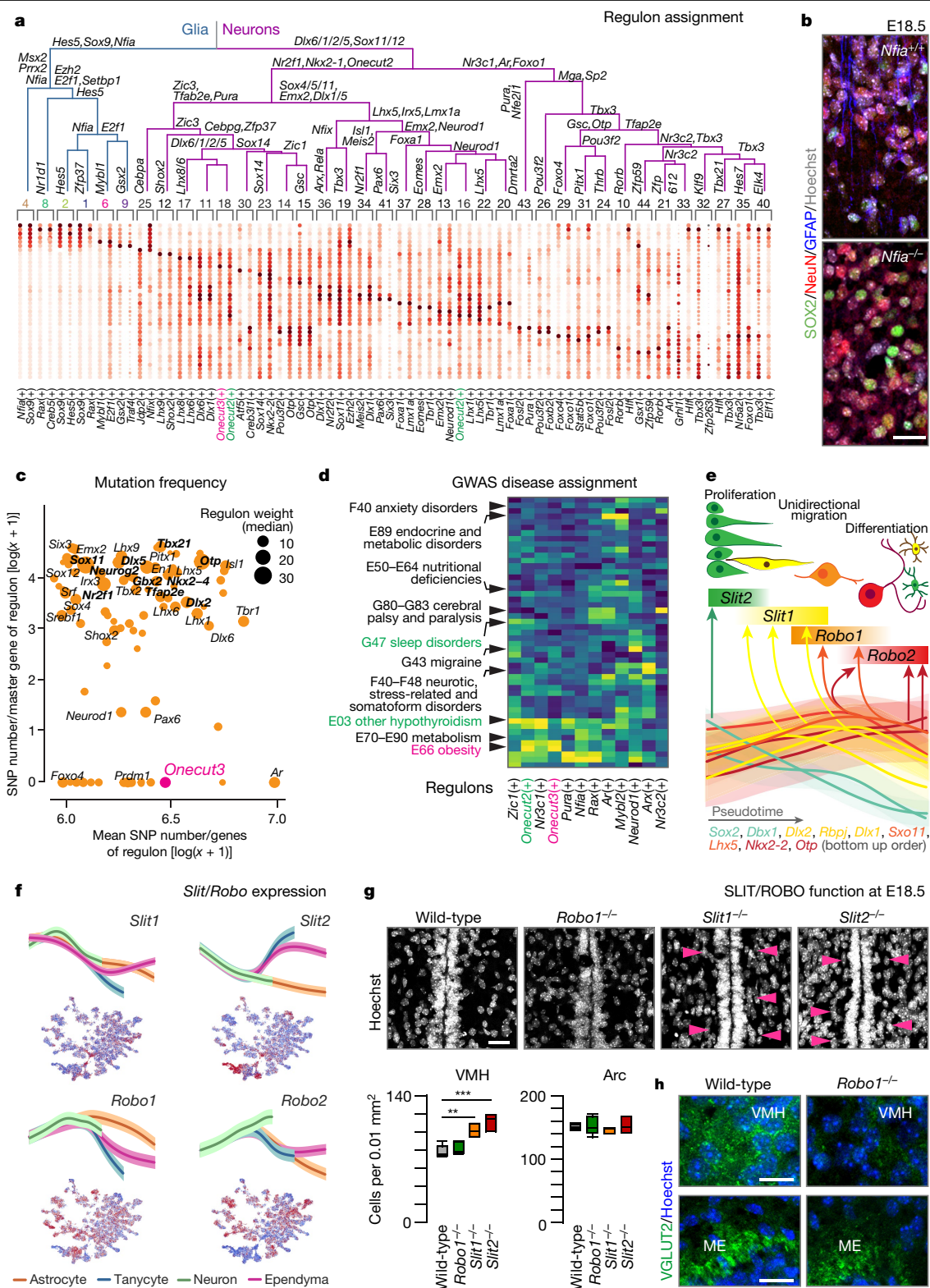
Fig. 3 | GRNs (regulons), including chemotropic guidance cues, in ectoderm-derived hypothalamic cells. a, A dendrogram of regulons for each cell cluster estimated in Fig. 1b. TFs at each branching point of the dendrogram are representative for subjacent groups of regulons. *Onecut* TFs are colour-coded. **b**, GFAP, SOX2 and RBFOX3 (NeuN) in the anterior Arc of wild-type and *Nfia*[−/−] mice at E18.5. **c**, Ratio between mutability for master genes and their downstream targets in regulons. A quadrant highlighting the *Onecut3* regulon is shown (see also Extended Data Fig. 5). **d**, Heat-map of associations between selected regulons and clinical phenotypes (see also Extended Data Fig. 5). Yellow shows stronger associations. **e**, Illustration of how

regulons that chiefly control SLIT–ROBO signalling contribute to neuronal differentiation in the hypothalamus. **f**, Expression of *Slit1/2* and *Robo1/2* in pseudotime (top; mean ± s.d.) and on an integrated dataset (bottom). Blue–red scale denotes low–high mRNA expression. **g**, *Slit1*[−/−] and *Slit2*[−/−] mice show increased cell density at the level of the VMH (arrowheads) but not the Arc relative to *Robo1*[−/−] and wild-type littermates. **h**, In turn, glutamatergic (VGLUT2) synaptogenesis is reduced in the VMH of *Robo1*[−/−] mice by E18.5. The median eminence (ME), where *Slit2* is not expressed, lacked any phenotype. Scale bars, 20 μm (**b**, **h**, **i**).
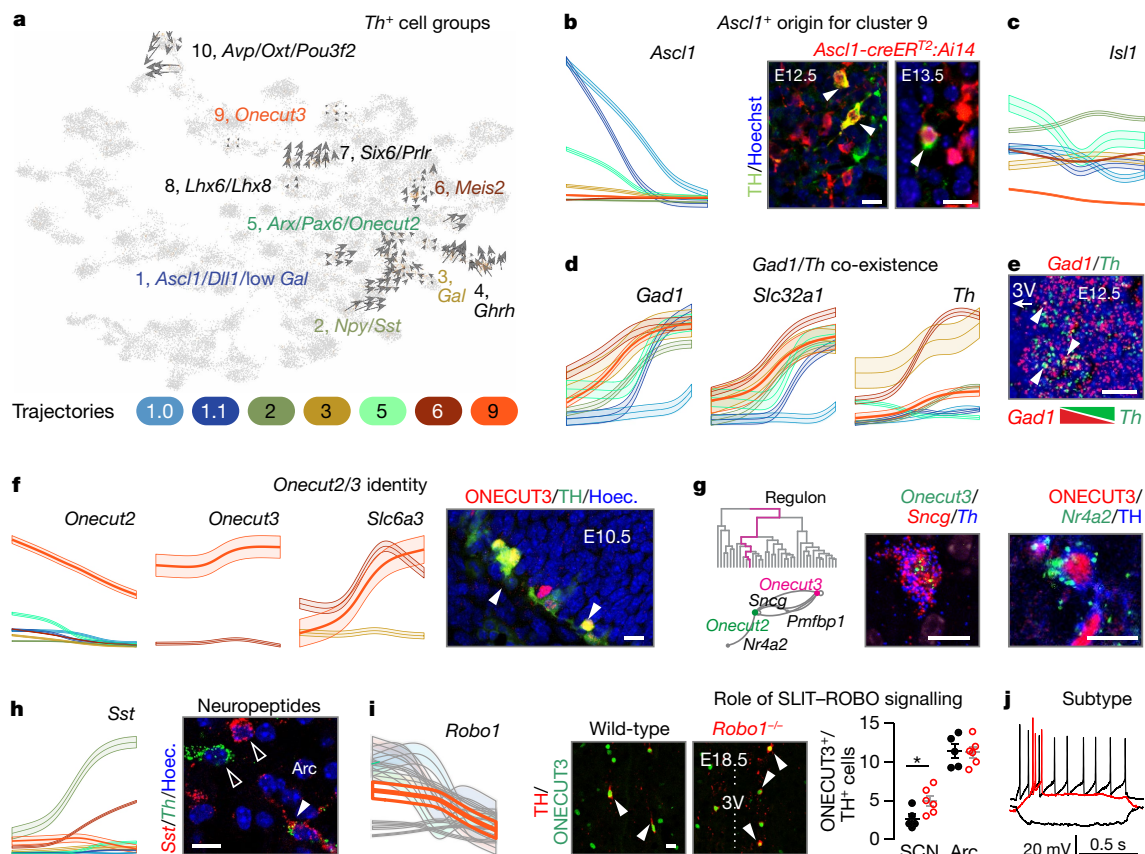
Fig. 4 | Molecular configuration of hypothalamic dopamine systems. a, RNA velocity vector embedding of *Th*⁺ neurons. Phenotypic convergence even for molecularly distant neurons is by uniform expression of *Th* and other enzymes of dopamine synthesis. **b**, *Th*⁺ neurons invariably rely on *Ascl1* as revealed in pseudotime (left). Genetic tracing in *Ascl1-creER^T2^::Ai14* mice showed the production of *Th*⁺/*Ascl1*⁺ progeny during the E12.5–E16.5 period. **c**, *Isl1* is expressed by all *Th*⁺ dopamine subgroups (pseudotime; Extended Data Fig. 9d). **d**, Pseudotime trajectories show that *Gad1*, *Slc32a1* and *Th* are co-expressed from the early fetal period. **e**, Antiparallel expression of *Gad1* and *Th* as a factor of medial-to-lateral positioning (Extended Data Fig. 9e). **f**, Pseudotime trajectories for *Onecut2*, *Onecut3* and *Slc6a3* (left) and earliest

positions of *Th*⁺/*Onecut3*⁺ neurons (arrowheads, right) in prospective PeVN dopamine neurons (cluster 9; right). **g**, Validation of target genes for the *Onecut3* regulon in PeVN dopamine neurons. Left, scheme identifies hierarchical relationships for *Onecut2*, *Nr4a2*, *Pmfbp1a* and *Sncg*. **h**, *Sst* was enriched in dopamine neurons in the Arc. **i**, Pseudotime trajectory for *Robo1* in dopamine neurons (left). Middle, *Th*⁺/*Onecut3*⁺ neurons in wild-type and *Robo1*⁻/⁻ mice. 3V, third ventricle. Right, quantification of cell numbers at the suprachiasmatic (SCN) and Arc levels (*P < 0.05; Student's *t*-test for independent groups). **j**, Patch-clamp electrophysiology classifies A14 *Onecut3*⁺ dopamine neurons as uniform 'type C' cells (Extended Data Fig. 8). Scale bars, 10 μm.

spatially segregated groups (Fig. 1a) without an intermediate cell pool being detected. Thus, GABA and glutamate neurons seem to adopt principally different developmental programs with immature GABA cells, rather than pre-formed GABA lineages[18], serving as precursors for terminal differentiation.

TF and neuropeptide codes of diversity

TF-mediated cell-autonomous differentiation programs are key to neuronal specification[6,18]. Hence, we screened TFs that distinguished hypothalamic cell clusters. We applied a supervised approach that sampled stationary states (genes that are spatially restricted in both pre- and postnatal brains; Fig. 1d, Extended Data Fig. 1) and integrated stages of fate transition and branching-off of differentiated neurons. *Avp*⁺ (cluster 26) and *Oxt*⁺ (cluster 43) magnocellular and parvocellular neuroendocrine clusters that are destined to the paraventricular nucleus (PVN; including *Trh*⁺ and *Crh*⁺ cells (cluster 24)) exhibited spatial convergence (Fig. 2a) and were separated by differentially expressed genes from their non-PVN counterparts. For example, *Mbnl3*, *Pgf*, *Irs4*, *Gpr101*, *Nr3c2* and *Agtr1* demarcated *Trh*⁺ neurons in the PVN, whereas *Trh*⁺ neurons that prospectively populate the dorsomedial hypothalamus (DMH, cluster 15) were labelled with *Onecut2*, *Onecut3* and *Cartpt*, and mapped distantly.

Subsequently, we selected *Pomc*⁺ neurons to test whether scRNA-seq-based temporal profiling of gene expression allowed us to reconstruct neuronal differentiation. Besides cataloguing *Pomc*-specific TFs (Fig. 2a), we show that, for example, *Prdm12* and *Nhlh2*, both placed by in situ hybridization into *Pomc−Gfp*⁺ neurons (Fig. 2b), are transiently expressed at early developmental stages, followed by gradual decay of expression (pseudotime; Fig. 2b). By contrast, *Cited1* expression was restricted to late gestation, when neuronal morphogenesis commences[19]. Cumulatively, our scRNA-seq data reliably resolved neuronal fate progression along both pseudotime and real-time scales.

Beside fast neurotransmission by GABA and glutamate, dopamine and neuropeptides are chief signalling units in the adult hypothalamus[20]. We assigned 27 neuropeptides specifically to GABA, glutamate and dopamine neuronal subtypes (Fig. 2c, d and Extended Data Fig. 3b–e). Our data demonstrate a transient increase in expression of, for example, *Sst*, *Tac1*, *Bdnf*, *Adcyap1*, *Pnoc*, *Nmu* and *Trh* in juvenile mice (P10)[21] (Fig. 2c) along with the rapid induction of their cognate receptors during the early postnatal window (Fig. 2d). Finally, we used *Oxtr^Venus/+^* mice to show that the onset of *Venus* expression (used as a surrogate for *Oxtr*) in, for example, *Pomc*⁺ neurons in the arcuate nucleus (Arc)[22], DMH and ventromedial hypothalamus (VMH) occurs at about E18.5 with

Article

a gradual increase postnatally (Fig. 2e, Extended Data Fig. 3f). These findings substantiate the precision of scRNA-seq in resolving hormone receptor expression in even the smallest neuronal contingents.

Regulons typify cell-type specificity

Cellular identities are shaped by developmentally timed GRNs ('regulons') that are centred on a 'master' TF that activates its targets through DNA-binding and transcriptional induction[23]. Therefore, we assigned regulons to each ectodermal cluster by combining 1,962 TF chromatin immunoprecipitation and sequencing (ChIP–seq) assays for gene interactions and our scRNA-seq data[23] (Fig. 3a, Extended Data Fig. 4), enriched positive interactions within 395 active regulons and estimated their prevalence per cell. Pleiotropic regulons with the highest representation defined the major cell lineage to which a progenitor was destined: for example, *Hes5*, *Sox9* and *Nfia* for prospective astroglia versus *Dlx1/2/5/6* and *Sox11/12* for neurons (Fig. 3a). Accordingly, *Nfia*$^{-/-}$ mice showed impaired formation of hypothalamic tanycytes and astrocytes but not neurons at E18.5 (Fig. 3b). Subordinate cell group-specific regulons defined select cell clusters (Fig. 3a). Co-existent regulons at all levels (for example, the *Pura* and *Fosl2* regulons in *Oxt*$^+$ neurons) produced combinatorial codes for cellular fate decisions (Fig. 3a shows at least three, from general to particular developmental processes).

We then evaluated the robustness and penetrance of hypothalamic regulons by testing whether their mutations (at all gene levels) manifest as clinical perturbations by focusing on metabolic and psychiatric diseases[9,24] in the genome-wide association study (GWAS) of the UK biobank (738 phenotypes; Extended Data Fig. 5a). We adapted existing methods in adults[9] by replacing stationary cell identities with regulons. By selecting multiple genes that co-define particular regulons as input, we substantially reduced selection bias due to strongly deleterious mutations ('survivorship bias'[25]). Regulons driven by pro-neurogenic genes were characterized by the lowest rate of mutated master genes (Extended Data Fig. 5b) with the *Foxo4* (cluster 29) and *Onecut3* (cluster 18) clusters completely depleted of mutations (Fig. 3a, c). Meanwhile, the *Onecut2*/*Onecut3* regulons correlated positively with the incidence of obesity (Fig. 3d).

Next, we confirmed that *Nr4a2*, *Ptfmb1a*, *Sncg*, *Lancl3* and *Zic5* (genes in the mutual *Onecut2*/*Onecut3* regulon) co-existed with *Onecut3* in differentiated neurons but were restricted to periventricular nucleus (PeVN) cell groups (Extended Data Fig. 6a–c). In addition, overexpression of *Onecut3* in Neuro2A cells in vitro stopped cell proliferation (Extended Data Fig. 6d, e), substantiating its role in neuronal specification. Cumulatively, these data assign regulon screens as a prime strategy to functionally annotate hypothalamic neurons and predict their linkages to metabolic (and psychiatric) disorders. The identification of a spatially restricted *Onecut3*$^+$ regulon to the PeVN suggests that neurons specified by the *Onecut3* regulon could be sensitive to developmental signalling cues that shape midline structures.

Regulons instruct chemotropic signalling

Within laminated structures, transitions from progenitor to committed progeny occur in a sequential unidirectional order[7,26]. We investigated whether similar gene sets[14,16,17], intercellular interactions and spatial arrangements could apply to the non-laminar hypothalamus. Early-expressed glial genes (*Hes1*, *Fabp7*, *Slc1a3*, *Vim*) marked progenitors (cells expressing, for example, *Sox2*, *Dll1/3*) in the innermost ('ventricular') zone of the third ventricle at E14.5–E15.5 (Fig. 1c, Extended Data Fig. 7a, b). Committed progeny then unidirectionally distanced themselves laterally (Extended Data Fig. 2d–g) and expressed protogenes for neuronal migration (*Dlx1/2/5/6*, *Rbfox3*; Fig. 1d). Plotting regulons along developmental age assigned *Sox2* to progenitors (clusters 6, 9), *Sox11* to bridge neurons (cluster 19) and *Dlx1/2* to both bridge and immature neurons (clusters 11, 19; Fig. 3e), confirming determination of function

by regulons[16]. These data suggest that regulons show temporal and spatial segregation in an onion skin-like layered configuration (Fig. 3e).

Next, we investigated the complexity of chemotropic signalling systems[27] that facilitate neuronal positioning and differentiation, with evidence for the expression of Ephrin–ErB, *Cbln1*/*Cbln2*, semaphorin–plexin–neuropilin, neurotrophin (*Bdnf*, *Gdnf*, *Cntf*), draxin, netrin (*Ntng1*/*Ntng2*) and endocannabinoids. Unexpectedly, we noted widespread expression of the *Slit1/2*–*Robo1/2* signalling cassette (Fig. 3e, f), which dictates direct (as opposed to indirect) neurogenesis in antagonism with *Dll*[26], and controls long-range axonal patterning in dorsal (cortical) structures[5]. Reconstruction of mRNA expression placed *Slit2* into neural progenitors at early developmental stages (Fig. 3f). Conversely, pseudotime analysis suggested that *Slit1* expression dominated early in postmitotic neuroblasts (Fig. 3f). Coincidently, *Robo2* expression defined a developmental trajectory specific to neurons (Fig. 3f). To genetically tie temporal variations in *Slit*–*Robo* signalling to neuronal differentiation, we show that major regulons include SLIT2 ligand synthesis for cell proliferation and gliogenesis (in the *Sox2*/*Dbx1*/*Rfx2*/*Rfx3*/*Myc* regulons that are specific to glia and progenitors) and SLIT1 for neuronal migration and morphogenesis (the *Dlx1*/*Dlx2*/*Rbpj* regulon). Meanwhile, the *Sox11* regulon is a chief determinant of both *Robo1* and *Robo2* expression as early as in bridge neurons (clusters 11, 19). At the level of terminal differentiation, *Lhx5*/*Emx2*/*Lhx1* and *Nkx2-1*/*Otp*/*Isl1* controlled expression of *Robo1* and *Robo2*, respectively (Fig. 3e). In situ hybridization confirmed the reciprocal distribution of *Slit2* and *Slit1*, with the former being restricted to ventricular progenitors (Extended Data Fig. 7b). Moreover, expression of *Slit2* (and to a lesser extent *Slit1*) was concentrated in the VMH by E17.5 or later (Extended Data Fig. 7b, bottom right). Indeed, both *Slit1*$^{-/-}$ and *Slit2*$^{-/-}$ mice showed increased cellular density in the subventricular zone at the level of the VMH, whereas *Robo1*$^{-/-}$ mice did not (Fig. 3g). Instead, *Robo1*$^{-/-}$ mice, in which SLIT ligands no longer act as repulsive axon guidance cues[5], showed a reduced density of *Slc17a6*$^+$/*Vglut2*$^+$ synapses in the VMH relative to wild-type controls, but unchanged levels at the median eminence, an area devoid of *Slit* gene expression (Fig. 3h). In sum, our data suggest that SLIT–ROBO signalling is involved in hypothalamic neurodevelopment, pointing to conserved SLIT–ROBO functions in ventral brain areas.

Molecular identity of dopamine neurons

Finally, we investigated how molecularly distinct subtypes of phenotypically uniform neurons arise during development of the hypothalamus. We took advantage of the at least nine morphologically and electrophysiologically distinct subtypes of parvocellular dopamine neurons in the A12 (Arc; three subtypes), A13 (zona incerta; two subtypes) and A14 territories (PeVN, four subtypes; Extended Data Fig. 8) of *Th*^{gfp} and *Slc6a3-Ires-cre::Ai14* mice, which are segregated from midbrain dopamine neurons that are chiefly regulated by *Lmx1a/b* and *Nr4a2*[28] (Fig. 3a).

First, we tested whether hypothalamic dopamine neurons that co-express tyrosine hydroxylase (*Th*), dopa decarboxylase (*Ddc*) and vesicular monoamine transporter 2 (*Slc18a2*) share a developmental trajectory. RNA velocity vector embedding for all *Th*$^+$ cells unequivocally identified ten molecularly distantly segregated neuronal clusters, of which clusters 4, 7 and 8 differentiated before E15.5 (Fig. 4a, Supplementary Note).

Considering that both pleiotropic and specific genetic programs contribute to molecular diversity among hypothalamic dopamine neurons, we addressed the earliest and uniform genetic codes in putative progenitors. Cascading *Ascl1* and *Isl1* expression was present in all dopamine neurons (Fig. 4b, c, Extended Data Fig. 9a–d), assigning these TFs to defining the entire dopamine class. Indeed, both *Ascl1-creER*^{T2}::*Ai14* and *Isl1-cre::Ai14* mice produced *tdTomato*$^+$ dopamine cells, particularly in the PeVN (Fig. 4b, Extended Data Fig. 9a, b, d), when induced at E12.5–E15.5.

The lack of Th^+ neurons in the hypothalamus but not midbrain of $Ascl1^{-/-}$ mice confirmed that the hypothalamus depends on an ASCL1-driven transcriptional pathway (Fig. 4b, c, Extended Data Fig. 9a, b).

Second, we investigated whether the dopamine phenotype evolves from the GABA lineage (Figs. 2d, 4a, d), a hypothesis that is consistent with data from adult mice[4,29]. Th^+/Ddc^+/$Slc18a2^+$ dopamine neurons arise from seven spatially segregated groups of GABA cells (clusters 1, 3, 4, 6, 7, 8 and 9; Fig. 4a). Our hypothesis was corroborated by the approximately 90% co-expression of Th and $Gad1$ in immature neurons (Fig. 4d, e, Extended Data Fig. 9e), including in (BAC)GAD65–eGFP and GAD67–GFP mice (Extended Data Fig. 10a). To identify genes that promote GABA-to-dopamine phenotypic transitions, we screened hypothalamic regulons for Th as target and found that the expression patterns of master genes for the $Meis2$, $Nfe2l1$, $Dlx1$ and $Pbx3$ regulons cover the broad initiation of Th expression at embryonic time points (Extended Data Fig. 10b).

Third, we searched for TFs that segregate dopamine subclasses. We focused on $Onecut3$, which specifies dopamine neurons in the PeVN[4]. Developmentally, $Onecut3$ serves as the master gene of the regulon that typifies Th/$Slc6a3$ neurons (cluster 9), and is detectable in the preoptic progenitor area by E10.5 (Fig. 4f, g). Histochemistry specifically tied the co-existence of $Onecut2$/$Onecut3$ and $Sncg$, $Pmfbp1a$ and $Nr4a2$ to PeVN dopamine neurons (Fig. 4g, Extended Data Fig. 10c, d). To further resolve the segregation of A14 neurons, we identified substantial Sst expression prenatally (with a gradual decay after birth) in $Onecut3^+$ dopamine neurons (Fig. 4h, Extended Data Fig. 10e). On the basis of $Robo1$ expression in the pseudotime scale, we integrated chemotropic cues for the final positioning of $Onecut3^+$/Th^+ neurons by showing that there were many more of these cells in $Robo1^{-/-}$ mice than in their wild-type littermates (Fig. 4i). Finally, $Onecut3$ expression distinguishes PeVN Th^+ neurons that produce a uniform electrophysiological signature that sets them apart amongst the nine dopamine subtypes tested (Fig. 4j, Extended Data Figs. 8, 10f), thus completing a differentiation trajectory that segregates PeVN dopamine neurons from all other dopamine subtypes.

Discussion

Our study provides an overview of ectodermal cell identities in the developing hypothalamus during pre- and postnatal periods. We show that a constellation of and temporal dependence on regulon activity, neurotransmitters and neuropeptides shapes ectodermal clusters. Large-scale GWAS-based disease assignment linked GRN activity to the life-long determination of neuronal functionality and consequently to predisposition to metabolic illnesses. In addition, transient waves of neuropeptide expression were synchronous with critical junctions of neuronal fate progression, thus generating long-lasting imprints on neuronal circuit complexity.

We found that a periventricular cellular reserve persists throughout life to generate hypothalamic neurons, with a contingent of GABA progenitors acting as a source for dopamine subtypes. We suggest that the existence of quasi-stable immature intermediates for GABA neurons, their provisional positioning for protracted periods, and sequential depletion until after birth are poised to assure flexibility in expanding functionally distinct neurocircuits by the insertion of neurochemically specialized cellular subtypes. Thus, the fundamental rules of neuronal specification in the hypothalamus could differ substantially from those found in laminated structures[6,7,26]. Nevertheless, we found that chemogenetic cues that have been classically viewed as dominating in cortical areas, particularly SLIT–ROBO signalling, also dictate neurogenesis, cell migration[5] and synaptogenesis (even if at the microscale) during development of the hypothalamus. Overall, combining differential gene expression analysis, screens for spatially restricted genes and GRN profiling into a discovery pipeline showcases the level of precision achievable to disentangle developmental processes that shape neuroendocrine centres and provides a template for studying both the origins of hypothalamic circuit operations and the molecular underpinnings of congenital and acquired metabolic disorders.

Online content

1. Saper, C. B. & Lowell, B. B. The hypothalamus. *Curr. Biol.* **24**, R1111–R1116 (2014).
2. Dulac, C., O'Connell, L. A. & Wu, Z. Neural control of maternal and paternal behaviors. *Science* **345**, 765–770 (2014).
3. Xie, Y. & Dorsky, R. I. Development of the hypothalamus: conservation, modification and innovation. *Development* **144**, 1588–1599 (2017).
4. Romanov, R. A. et al. Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
5. Andrews, W. et al. The role of Slit-Robo signaling in the generation, migration and morphological differentiation of cortical interneurons. *Dev. Biol.* **313**, 648–658 (2008).
6. Mayer, C. et al. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).
7. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10**, 724–735 (2009).
8. Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014. e1022 (2018).
9. Campbell, J. N. et al. A molecular census of arcuate hypothalamus and median eminence cell types. *Nat. Neurosci.* **20**, 484–496 (2017).
10. Toda, C., Santoro, A., Kim, J. D. & Diano, S. POMC neurons: from birth to death. *Annu. Rev. Physiol.* **79**, 209–236 (2017).
11. Burbridge, S., Stewart, I. & Placzek, M. Development of the neuroendocrine hypothalamus. *Compr. Physiol.* **6**, 623–643 (2016).
12. Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
13. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
14. Aujla, P. K., Naratadam, G. T., Xu, L. & Raetzman, L. T. Notch/Rbpjk signaling regulates progenitor maintenance and differentiation of hypothalamic arcuate neurons. *Development* **140**, 3511–3521 (2013).
15. McNay, D. E., Pelling, M., Claxton, S., Guillemot, F. & Ang, S.-L. Mash1 is required for generic and subtype differentiation of hypothalamic neuroendocrine cells. *Mol. Endocrinol.* **20**, 1623–1632 (2006).
16. De Marco García, N. V., Karayannis, T. & Fishell, G. Neuronal activity is required for the development of specific cortical interneuron subtypes. *Nature* **472**, 351–355 (2011).
17. Cobos, I., Borello, U. & Rubenstein, J. L. Dlx transcription factors promote migration through repression of axon and dendrite growth. *Neuron* **54**, 873–888 (2007).
18. Wamsley, B. & Fishell, G. Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nat. Rev. Neurosci.* **18**, 299–309 (2017).
19. Gerstner, J. R. & Landry, C. F. Expression of the transcriptional coactivator CITED1 in the adult and developing murine brain. *Dev. Neurosci.* **29**, 203–212 (2007).
20. Hökfelt, T., Meister, B., Melander, T. & Everitt, B. Coexistence of classical transmitters and peptides with special reference to the arcuate nucleus–median eminence complex. *Adv. Biochem. Psychopharmacol.* **43**, 21–34 (1987).
21. Motoike, T. et al. Transient expression of neuropeptide W in postnatal mouse hypothalamus—a putative regulator of energy homeostasis. *Neuroscience* **301**, 323–337 (2015).
22. Maejima, Y. et al. Oxytocinergic circuit from paraventricular and supraoptic nuclei to arcuate POMC neurons in hypothalamus. *FEBS Lett.* **588**, 4404–4412 (2014).
23. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
24. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).
25. Popadin, K. et al. Slightly deleterious genomic variants and transcriptome perturbations in Down syndrome embryonic selection. *Genome Res.* **28**, 1–10 (2018).
26. Cardenas, A. *et al.* Evolution of cortical neurogenesis in amniotes controlled by Robo signaling levels. *Cell* **174**, 590–606.e521 (2018).
27. Chédotal, A. & Richards, L. J. Wiring the brain: the biology of neuronal guidance. *Cold Spring Harb. Perspect. Biol.* **2**, a001917 (2010).
28. Arenas, E., Denham, M. & Villaescusa, J. C. How to make a midbrain dopaminergic neuron. *Development* **142**, 1918–1936 (2015).
29. Tritsch, N. X., Ding, J. B. & Sabatini, B. L. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature* **490**, 262–266 (2012).

Article

Methods

Mouse strains

All mice were housed in groups in clear plastic cages on a 12 h–12 h light–dark cycle (lights on at 08:00 h) and in a temperature (22 ± 2 °C) and humidity (50 ± 10%) controlled environment. Food and water were available ad libitum. Embryos and tissues were obtained from timed matings with the day of vaginal plug considered as embryonic day (E) 0.5. The day of birth was always registered as postnatal day (P) 0. Postnatal animals were weaned on P21. Commercial mouse lines were: C57Bl/6J wild-type (RRID:IMSR_JAX:000664), Ai14 (RRID:IMSR_JAX:007914), *Ascl1-creER^T2* (RRID:IMSR_JAX:012882), *Th–Gfp* (RRID:IMSR_RBRC03162), (BAC)GAD65–eGFP (RRID:MMRRC_011849-UCD), GAD67 (RRID:IMSR_RBRC03674), *Pomc–Gfp* (RRID:IMSR_JAX:009593), *Slc6a3*-Ires-cre (RRID:IMSR_JAX:006660), *Nfia^−/−* (RRID:MMRRC_010318-UNC), *Robo1^−/−* (RRID:IMSR_APB:5320), *Slit1^−/−* (RRID:MMRRC_030404-MU), *Slit2^−/−* (RRID:MMRRC_030405-MU), *Isl1-cre* (RRID:IMSR_JAX:024242) and *Oxtr^Venus/+* (MGI:3838764)[30–41]. *Ascl1-creER^T2* knock-in mice were used as heterozygotes when performing lineage tracing and as homozygotes to study developmental consequences of the lack of *Ascl1* since both copies of the gene were replaced by the Cre coding region (referred to as *Ascl1 ko*). *Nfia^−/−* mice were provided by J. Bunt and L. J. Richards as a mechanism to re-use tissue (QB/356/17). *Nfia^−/−* mice were bred for work conducted under National Health and Medical Research Council project grant GNT1100443 and Principal Research Fellowship GNT1120615. Tracing experiments for all other Cre lines were performed using heterozygotes. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Tissue collection and fixation

Whole heads of embryos (E10.5–E15.5) or dissected brains (E16.5 and older) were collected and fixed in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS, 0.05 M, pH 7.4) at 4 °C for 4 h for E13.5 and 16–24 h for E16.5 or older. For postnatal stages and adult brain samples, animals were transcardially perfused with 4% PFA in 0.1 M phosphate buffer (PB; pH 7.4) and dissected brains post-fixed overnight. Samples were then washed in PB and cryoprotected by incubating in 30% sucrose in distilled water at 4 °C overnight.

Ethical approval of animal studies

Experiments on live animals conformed to the 2010/63/EU European Communities Council Directive and were approved by the Austrian Ministry of Science and Research (66.009/0145-WF/II/3b/2014, and 66.009/0277-WF/V/3b/2017). Particular effort was directed towards minimizing the number of animals used and their suffering during experiments.

Tamoxifen injection and tissue processing

Ascl1-creER^T2::Ai14 dams were injected with tamoxifen (150 mg/kg) on one of the days of E11.5–E16.5 to induce Cre-mediated recombination. The brains of the embryos were collected and immersion fixed in 4% PFA in PB (pH 7.4) for 12–24 h before being immersed into 30% sucrose for cryoprotection (48 h). Embryonic brain tissues were cut at 16 μm thickness and mounted on fluorescence-free glasses. Postnatal animals were perfusion-fixed with 50–100 ml of 4% PFA in PB, followed by cryoprotection as above. Brains were then cut on a cryostat as 50-μm-thick serial free-floating coronal sections.

Cell capture, lysis and RNA-seq

C57Bl/6NRj mice (E15.5–P23, Janvier Labs Cat# SC-C57N-F) of both sexes were used for cell collection. Embryos were removed by Caesarean section and immersed in ice-cold pre-oxygenated (95% O$_2$/5% CO$_2$) cutting solution containing (in mM): 90 NaCl, 26 NaHCO$_3$, 2.5 KCl, 1.2 NaH$_2$PO$_4$, 10 HEPES-NaOH, 5 Na-ascorbate, 5 Na-pyruvate, 0.5 CaCl$_2$, 8 MgSO$_4$ and 20 glucose. Postnatal animals were deeply anaesthetized (5% isoflurane) and transcardially perfused with 40 ml of the same solution. Entire hypothalami were isolated manually under microscopy guidance from serial 300-μm-thick coronal slices and then dissociated using the Papain Dissociation System (Worthington) according to the manufacturer's recommendations with additional mechanical dissociation using Pasteur pipettes with 600-, 300- and 150-μm open tips. After the cells were re-suspended in sterile cutting solution supplemented with 0.1% BSA, they were fixed in ice-cold methanol for 10 min and stored at −80 °C until library preparation.

For the preparation of cDNA libraries, cells were re-suspended in PBS (0.01 M, pH 7.4) and concentrated to a range of 105–700 cells per μl. Thirty-three microlitres of the cell suspension together with 1 μl cellular spike-ins (lymphocytes) were added to the reverse transcription mix. cDNA synthesis, library preparation and sequencing were performed according to the instructions for the 10x Genomics Chromium Single Cell Kit (version 2). High-throughput RNA sequencing was on an Illumina HiSeq3000 instrument.

10x Genomics data pre-processing

Data derived at each time point were processed independently (Supplementary Fig. 1, Supplementary Note). Raw files were processed with Cell Ranger[42] (version 2.2.0) following default arguments for velocyto.py[13]. Reads were mapped to the Cell Ranger mm10-1.2.0 genome and counted with complimentary annotation (Supplementary Figs. 2, 3, Supplementary Note). To derive unique molecule count (UMI) expression matrices, we additionally compared two advanced computational approaches. First, we read raw matrices from the Cell Ranger pipeline (Supplementary Fig. 3, Supplementary Note) into emptyDrops[43] implemented in the DropletUtils R package. We used a false discovery rate (FDR) of 0.01 with 2×10^5 permutations (Supplementary Fig. 4, Supplementary Note). Second, we pre-processed raw fastq-files using the dropEst pipeline[44] with the UCSC mm10 mouse genome and default dropEst parameters for 10X (Supplementary Fig. 5, Supplementary Note). In brief, dropEst utilizes Bayesian correction of cell barcodes and UMIs, taking into account Hamming distance distribution for cell barcodes and probability distribution by sequential estimation of errors with maximal likelihood between different barcodes within each gene on multiple metadata sources. These include sequencing quality of nucleotide in position (Phred score) and the number of reads for each barcode (coverage) as the most critical parameters. When collision targets are merged, the pipeline estimates damaged and low-quality cells in two steps. First, it automatically assigns cells based on cell size (Supplementary Fig. 5, Supplementary Note). Cell labels were marked with two estimated thresholds: lower than first for 'low-quality' (red), then, 75% of cells higher than second as 'high-quality' (green) with the remaining cells considered as 'unknown' (grey). Second, initial labels together with sets of biological and technical factors (mitochondrial fraction, mean number of reads per UMI, mean number of UMIs per gene, fraction of drop-out genes, fraction of intergenic reads, fraction of not-aligned reads) were deciphered by the kernel density estimate (KDE) classifier to endow each cell with a quality score (0–1 range).

emptyDrops and dropEst algorithms hold substantially more cells than the default Cell Ranger approach without a crucial difference between them. Thus, we used the dropEst pipeline[44] throughout, which additionally provides quality control metrics for the cells albeit at the cost of high computational load. As a result, we used a corrected matrix with cells that passed filters of both emptyDrops and Cell Ranger and possessing dropEst's 'high-quality' label (upper quartile) together with all cells above the 90th percentile of quality score (Supplementary Fig. 5, Supplementary Note).

Expression matrix filtering

We performed an exploratory analysis of dropEst output matrices in sequential steps of annotation and filtering. First, we checked

known genes, which indicate diverse sources of bias, such as ribosomal, immediate-early stress-responsive and gender-specific genes (*Gm42418, Rpl26, Gstp1, Rpl35a, Erh, Slc25a5, Pgk1, Eno1, Tubb2a, Emc4, Scg5, Ehd2, Espl1, Jarid1d, Pnpla4, Rps4y1, Xist, Tsix, Eif2s3y, Ddx3y, Uty, Kdm5d*)[8,45]. Furthermore, we assessed the expression level of *HuR* (also known as *Elavl1*) to distinguish damaged neurons[46,47]. Thus, the above gene profiles were indicative of low-quality (potentially apoptotic) cells as well as identified cells of blood origin. Second, after removing biasing genes, we manually explored and annotated cell clusters with pagoda2[13] as described previously[44] using known marker genes[8,9,48–50] (https://doi.org/10.6084/m9.figshare.11867889). Third, we defined differentially expressed genes (DEG) for 'ribo-rubbish', 'excluded', 'duplets', 'endodermal and mesenchymal-related clusters' by comparison against putative ectodermal cell types using the model-based analysis of single-cell transcriptomics (MAST) test[51,52]. We repeated normalization, negative-binomial scaling, PCA dimension reduction, 'Jackstraw' pc-selection and kNN-graph construction steps after every cell–molecule matrix subsetting.

Integration of expression profiles against a time factor
We followed the lead design initially implemented for paired CCA-based integration of data on embryonic and adult cortical interneurons[6]. The crucial differences were: 1) intercalated nuclei in hypothalamus versus laminar cortical architectures; 2) higher adult stage transcriptional heterogeneity of neurons (for example, mixed GABA/glutamate phenotypes and magnocellular/parvocellular neurons)[4,53]; 3) distant volume transmission as additional factor[54,55]; 4) six developmental stages. We aimed to derive a manifold according to known adult cell types and lineages. We integrated datasets from successive time points (E15.5–P23) to discover the succession of developing cell lineages. This allowed us to apply retrospective analysis to distinguish 'ancestor' cells. We removed mature and myelin-forming oligodendrocytes, which existed only in late postnatal stages, from our analysis. We additionally filtered cells with <300 genes or 2.5×10^4 RNA molecules (taking into account only ectoderm-related genes). Finally, we used a variance-stabilizing transformation of the SCTransform method to find anchor-candidate genes[6,51].

Comparison of integration algorithms
Despite a recently published approach[6] for the integration of embryonic and adult scRNA-seq data, we additionally tested all presently available algorithms. To match our criteria, the algorithm should 1) provide a mixture of time and batch factor and 2) at the same time keep a defined cell type local structure. Therefore, we benchmarked 11 different solutions of integration using their default parameters (Supplementary Fig. 6, Supplementary Note): default balanced batch *k*-nearest neighbours (BBKNN)[56]; BBKNN with a neighbours trimming procedure based on their connectivity scores, which were derived by the UMAP algorithm[57] (BBKNN_TRM); BBKNN with exact neighbour identification via faiss[58] (BBKNN_FAISS); BBKNN based on *k*-dimensional-tree (cKDTree)[59]; Scanorama[60]; LIGER[61]; Harmony[62]; mutual nearest neighbour (MNN)[63]; CONOS[64]; Seurat 3 CCA-based integration of negative binomial fit scaled matrixes (SeuratCCA)[65]; and Seurat 3 CCA-based integration of scTransform derived matrixes of Pearsons residuals (SeuratCCAonSCT)[51]. We found that the updated version of the conventional Seurat approach performed similarly to CONOS, Harmony and LIGER, overperformed as compared to other algorithms in terms of the MixingMetric and underachieved compared to 1) Harmony (with Harmony over-fitting the batch factor, which we could not optimize against time and batch factors at the same time) and 2) its own more advanced version (Seurat 3.1[65]) in terms of the local structure metric, leading us to a more conservative way of integrating adult and embryonic stages.

CCA-based integration with Seurat 3.1
We integrated cells from different time points into a single manifold using the latest version of the Seurat 3.1 CCA-based integration pipeline[65]. The new version allowed us not to set the order of integration explicitly and thus determined the optimal order automatically. Thereby, we obtained the united manifold of 50 CC components from filtered and weighed integration anchors and linearly scaled genes used for integration into the CCA space. Subsequently, we performed cell-cycle difference regression: S-phase score minus G2M-phase score, according to the alternative Seurat workflow to remove differences in cell cycle phases amongst proliferating cells, following the available gene set annotation[66]. Next, we derived 50 principal component (PC) matrices from the corrected matrix. From this PC matrix, we selected PCs having >25 percentile of s.d. and learned UMAP embedding[57,67] (Fig. 1a, Extended Data Figs. 3a, 7a, d, e, Supplementary Note). As a result, we obtained a 37-dimensional manifold of all cells and its two-dimensional embedding.

Annotation of the collective manifold for hypothalamus development
Our next aim was to annotate putative cell lineages that pivot the hypothalamus through development to adulthood. Therefore, we used MetaNeighbour[68] in supervised mode to evaluate area under the recovery curve (AUC) scores of cross-explanation for different cellular annotations[4,8,9,69,70] that exist for adult hypothalamic scRNA-seq data (Supplementary Fig. 8, Supplementary Note). Thus, we expected to find putative replicative subtypes, but found an unexpected lack of consensus for cellular annotations, meaning that top calls were assigned to the same studies.

Therefore, we have chosen a recent dataset[8] as reference because of its completeness of anatomical sampling (another anatomically complete study of hypothalamus[69] was inferior in terms of cell numbers and their variability). Additionally, the label transferring approach of Seurat was used to verify the absence of possible contamination with cells from the thalamus in the filtered dataset. To this end, we integrated the expression profiles with signatures derived from the mouse brain atlas spanning the diencephalon[8].

Integration of juvenile data with default parameters for Seurat CCA revealed comprehensive coverage of the diversity of terminally differentiated progenies of reference[8] (Supplementary Fig. 9a, Supplementary Note). Next, we attempted to incorporate the adult reference to the entire developmental dataset (Supplementary Fig. 9b, Supplementary Note). We observed substantial disparities that suggest the existence of convergence processes during hypothalamus development, which necessitated an unbiased strategy of annotation. For this reason, we used clustering factors from the integration manifold by all field prevalent algorithms (graph-based algorithms mostly from the igraph package used in Pagoda2 with default parameters on 37 PCs and separately on the corrected UMI counts matrix, and a small local movement algorithm with different resolutions via Seurat's FindClusters function) for comparison (Supplementary Fig. 10, Supplementary Note). We observed robustness in the inner structure of the integrated manifold for all used algorithms, except infomap, by estimating information metrics in a cross-annotation manner (due to lacking a correct one) and silhouette scores using the R packages aricode[71] and clues[72]. As most algorithms produced similar results (with only minor differences between them), and our approach proved accurate (except for the infomap), we decided to proceed with walktrap[73].

Finally, we decided to split cells by major lineages: an exploratory analysis was performed in Pagoda2[13] after transferring our gene–cell integrated matrix and UMAP embedding. We used the standard Pagoda2 pipeline with the integrated manifold matrix clustered by walktrap[73]. We annotated the final 45 clusters by DEG testing with MAST[52] (Figs. 1a, 2a and Extended Data Fig. 1) on the scTransform-corrected[6,51] log-normalized UMI matrix (data slot in a Seurat object).

Dendrogram construction (gene-based)
A dendrogram was constructed for neuronal and glial cells separately with Seurat 3 with different feature sets as input: for the neuronal tree

Article

we used a list of genes from ref. [74]. The glial tree was based on genes from ref. [75]. For both trees, we used gene sets defined in ref. [76], which were filtered by taking only the upper 95th percentile for the corresponding cell types (neurons, astrocytes, oligodendrocytes), the upper 95th percentile for cell types in diencephalon and the upper 50th percentile of MeanExpression. We excluded housekeeping genes[77] for both trees. We used dendrograms to order our dot plots (Extended Data Fig. 1).

Neuropeptide and neurotransmitter assignment

Density plots on UMAP embedding for signalling molecules were assembled by ggalt[78] (Extended Data Fig. 3a, d, e). We manually split neuropeptides to prevalent GABAergic or glutamatergic co-existence. Top-ranked neuropeptides of the corrected UMI matrix were plotted with a threshold of ten molecules in colour and shape-coded manner for two groups separately.

Abstraction graph on repartitioned integrated data

We redistributed bridge cell (cluster 19) and immature neuron (cluster 11) populations by applying a combination of two methods: the Leiden network clustering algorithm[79] and PAGA[12]. Repartitioning immature-to-adult allowed us to negotiate selection bias[80]. We excluded clusters 3, 5, 7, 38, 39, 42 and 45 as oligodendrocytes and pars tuberalis from the integrated manifold matrix and performed clustering with Scanpy[81] wrapper with iterations until full optimization. These partitions were used for the construction of an abstraction graph (non-directed) with a threshold of 0.365. Nodes were not established as common clusters. Instead, we derived them to obtain the topological structure of cell ensembles. We analysed the projection of our preliminary clusters of each developmental stage on partitions of the abstraction graph and transformed it by Leiden algorithm (using time_slices_to_layers() function and then optimise_partition_multiplex() for class la.Optimiser()) to a connection-based annotation of cell lineages with consistent colours. Next, we used abstraction graph nodes as starting points for the UMAP algorithm with maxiter = 1,000, negative_sample_rate = 20, min_dist = 1, spread = 2 parameters to prepare embedding corresponding to cell lineage relations (Figs. 1b–d, 2a, b, 3f, 4a, Supplementary Fig. 11, Supplementary Note).

RNA velocity analysis 1

We performed RNA velocity analysis of time points separately, following the original deterministic approach of the velocyto.R/py packages[13]. Currently, it is not possible to split an UMI-matrix obtained by Bayesian correction of dropestR[44]. Thus, we exported metadata for filtering of a default matrix and our cellular annotations (walktrap algorithm derived 45 cell groups). They were then sub-grouped to relocate cells, which passed our filters and removed 'rubbish' and mesenchymal-related genes. Last, we applied the original RNA velocity method[13] with a few modifications: 1) we filtered out all non-DEGs, which were present in <20 spliced/unspliced molecules, and 2) to obtain a velocyto grid we exported the UMAP embeddings[67,73] of our high-quality cells from the previous step for visualization purposes (Extended Data Fig. 2b, c).

RNA velocity analysis 2

To perform RNA velocity analysis of the integrated dataset, we applied the scvelo python package using a generalized dynamical model[82]. As input, we used filtered to ectodermal cells and gene loom files, which were merged using loompy (version 2.0.17) and filtered to cells used for PAGA construction (see above). In brief, spliced and unspliced reads were separately size-normalized to the median of total molecules across cells. Additional gene filtering comprised those that passed a minimum threshold of ten expressed counts for spliced and two for unspliced mRNA. We quantified a 30-nearest neighbour graph based on Euclidean distances in 30-PC space (PCA performed on logarithmic spliced counts). Therefore, for each cell across its neighbours, we obtained first and second-order moments (means and uncentred variances),

then estimated RNA velocity with the explicit fitting of inferred splicing reaction rates. As a result, transition probabilities were estimated to form a velocity graph. Thus, we plotted individual cell velocities embedded in UMAP space (Figs. 2a, 3f, Extended Data Fig. 7a). Subgraph analysis of both the glial lineage (clusters 1, 10, 18, 24, 34 and 51) and bridge cells (we subset only the first entering node of the abstraction graph) was performed as described previously[12,82]. To this end, we subset cells of interest and transformed PAGA to a directed abstraction graph using a default constructed velocity graph as described above. Finally, we allocated root cells by using the backward Markov process on the transition probability matrix to define excessive density area, estimated the latent time on a learned transcription dynamic model and plotted a velocity grid and individual cell velocities embedded in UMAP space (Fig. 1c). We completed all steps using built-in functions with default parameters.

Estimation of developmental regulons

We prepared a subset of putative glial (astroependymal, tanycyte and progenitor) and neuronal clusters (as described above under 'Abstraction graph on repartitioned integrated data'). A spliced UMI count matrix of the integrated dataset was input into the pySCENIC[23,83] pipeline with default settings to infer active TFs and their target genes. In brief, the pipeline was implemented in three steps. First, we identified gene co-expression modules of TFs[83]. Second, we pruned each module based on a regulatory motif near a transcription start site (TSS). Cis-regulatory footprints could be obtained with positional sequencing methods (for example, from ChIP–seq motif calling with an antibody against a TF). Binding motifs of TFs across multiple species were then used to build an RCisTarget database. Precisely, modules were retained if the TF-binding motif was enriched among its targets, while target genes without direct TF-binding motifs were removed. Third, we scored the impact of each regulon for each single-cell transcriptome using the AUC score as a metric. Each step of this pipeline used rank statistics, and the last classification step ran independently for each cell, avoiding a batch effect.

Moreover, regulons tended to highlight higher-order similarities across cells. Thus, we determined whether the target genes in each regulon were enriched in each cell using the distribution of regulon activity across all cells in the dataset. The input list of TFs was downloaded from the RIKEN database[84–86]. As a result, we derived the AUC score matrix (AUCell) to validate our clustering and prepared annotation by 395 identified regulons. Inferred regulons and their activity across cells in the integrated dataset are reported at https://doi.org/10.6084/m9.figshare.11867889. Wilcoxon test, Logreg test[87] and dot-plot visualization of differentially recruited regulons across 45 identified cell groups (described above) were done in the Scanpy package[81] (Fig. 3a, Extended Data Fig. 4). GRN plots of the Onecut2/3 regulons were done using the Gephi software package[88] (Extended Data Fig. 6a).

Dendrogram construction (regulon-based)

Next, a dendrogram was constructed for neuronal and glial cells together using Seurat 3 on the AUCell matrix of 395 regulons to observe subtler changes. We deciphered the diverging composite rules of a regulon-based dendrogram by testing each branching node for differential regulon importance. Therefore, we performed the Wilcoxon test and Logreg test[87] of every node with min.pct = 0.01, logfc.threshold = 0.01 of Seurat's function FindAllMarkers to derive the action propagation program of the regulons (https://doi.org/10.6084/m9.figshare.11867889). We used dendrograms to order dot-plots (Fig. 3a, Extended Data Fig. 4).

Regulon assignment to clinical phenotypes

To understand the potential involvement of regulons in human disease phenotypes, we analysed the properties of human polymorphic variants (SNPs) located within regulon genes. Recently, a robust

correspondence between human and mouse regulons was reported[23]. To uncover associations between regulon-specific variants and human phenotypes we used Gene Atlas: a database with summary statistics of GWAS between millions of variants and hundreds of traits in the UK Biobank cohort (http://geneatlas.roslin.ed.ac.uk/)[89]. First, we converted a subset of mouse regulon genes with one-to-one orthologues into human Ensembl gene IDs. Here, we used mouse notation for regulons (only the first capital letter and the remaining lowercase letters) to clarify their source. Second, we extracted all SNPs belonging to regulon genes and analysed the distribution of their P values. Third, we characterized regulons in terms of the total number of SNPs that affected their master gene, as well as SNPs that affected other regulon-recruited genes. We observed that master genes tended to have a deficit in SNPs as compared to the downstream general representatives of these regulons (median number of SNPs that affect master genes across all regulons is less than the median number of SNPs that affect general representative genes). Some regulons showed zero SNPs (at least in the UK Biobank cohort) in their master genes but a substantial, close to the median, number of SNPs in regulon-recruited genes. Using the number of SNPs that affected master genes and the number of SNPs that affected other regulon-recruited genes as a metric of evolutionary constraint, we split all regulons into four quadrants that reflected the ratio of constraints between masters and members of regulons (Extended Data Fig. 5b). Next, we asked where hypothalamus-related regulons appeared in this plot. Thus, to focus on the hypothalamus, we used weights of genes in regulons output by the SCENIC workflow, which were interpreted as the importance of a given gene in a given regulon. To highlight the most important regulons, we chose ones with a median weight higher than the median weights of all individual genes from all regulons.

Henceforth, the logic behind our analysis is similar to pi1 statistics, estimating an excess of low P values[90]. To make a fast approximation of the excess of low P values for 778 phenotypes and each gene of 395 regulons, we estimated the fraction of P values that were <0.1. Limited testing demonstrated a strong correlation between our excess of low P values and pi1 estimated in the qvalue R package[91]. Also, limited testing demonstrated similar trends, observed when we changed the threshold from 0.1 to 0.01. Therefore, we manually selected clinical phenotypes related to the hypothalamus, subset to regulons with importance (weight) higher than the median and aggregate weighted pi1 statistics to the regulon level. For visualization purposes, we selected three contrast regulons from each quadrant by choosing distant regulons in the two-dimensional PCA space of phenotypes (red labels in Extended Data Fig. 5a, black spheres in Extended Data Fig. 5b) and plotted their normalized values via the heatmaply R package[92] (Fig. 3d).

Estimation of developmental genes and regulon dynamics

Next, we took the spliced molecule count matrix of the same subset of cells as for the PAGA subgraph analysis of progenitors and their nearest offspring. First, we size-normalized to the median of total molecules across cells. Second, a logarithmic matrix was used to estimate pseudo-time order and probabilities for cells to propagate through the subgraph of glial lineages or bridge cells. For this purpose, we used a probabilistic approach, Palantir[93], which we implemented as an external module to the Scanpy Python package[81] (Figs. 1d, 4b, Extended Data Fig. 7a, Supplementary Fig. 12 in Supplementary Note). Similarly, we applied this method to the *Pomc* cell group (Fig. 2b) and every distinct *Th* containing group (Fig. 4b–d, f, h, i, Extended Data Fig. 9c, Supplementary Fig. 13 in Supplementary Note) guided by PAGA topology. In all cases, we selected early cells by taking the upper 99th percentile of the *Sox2* regulon of AUC scores distribution and used default parameters for estimations with the exception of the waypoints parameter: for glia and bridge neurons, 500; for the *Pomc*+ cell lineage, 1,200; and for *Th*+ trajectories: cluster 1, 750; cluster 2, 500; cluster 3, 1,000; cluster 4, 500; cluster 5, 350; cluster 6, 1,000; cluster 7, 200; cluster 8, 200; cluster 9, 1,000; cluster 10, 800. Finally, we estimated the impact of regulon dynamics along the identified trajectories using the AUCell matrix as input for Palantir's function compute_gene_trends, which uses a generalized additive model. Both the trends of genes and regulon actions were clustered for each trajectory with default parameters using the Phenograph python package[94].

Tissue preparation and immunohistochemistry

After rinsing in 0.1 M PB, specimens were exposed to a blocking solution composed of 0.1 M PB, 10% normal donkey serum, 5% BSA and 0.3% TX-100 for 3 h followed by 48 h incubation with select combinations of primary antibodies: rabbit anti-TH (1:500; Millipore AB152, lot 2593900, 3199177), sheep anti-TH (1:1,000, Novus Biologicals, #NB300-110, lot ajo1217p), sheep anti-ONECUT2 (1:250; R&D Systems, AF6294, lot CDKS0116081), guinea pig anti-ONECUT3 (1:5,000)[95], rabbit anti-VGLUT2 (1:800; a gift from M. Watanabe)[96], goat anti-GFP (1:1,000; Abcam, #ab6662, lot GR311622-15, GR311622-7), chicken anti-GFP (1:500, Aves Labs Inc., #GFP-1020, lot GFP697986), rabbit anti-SOX2 (1:500, Abcam, #ab97959, lot GR3244885-1), chicken anti-mCherry (1:1,000; EnCor Biotech, #CPCA-mCHERRY, lot 7670-4), mouse anti-MASH1 (1:100, BD Pharmingen, 556604, clone: 24B72D11.1), guinea pig anti-GFAP (1:500, Synaptic Systems, 173004, lot 2-15, 2-17), rabbit anti-phospho-histone H3 (1:500; Cell Signaling Technology, 9701, lot 7), chicken anti-NeuN (1:500, Merck Millipore, ABN91, lot 3132967), mouse anti-FLAG-tag (1:1,000; Sigma, F1804, lot SLBR7936V), mouse anti-HA-tag (1:600; Cell Signaling Technology, mAb2367, lot 1). Secondary antibodies were from Jackson ImmunoResearch, including Alexa Fluor 488-AffiniPure donkey anti-goat (705-545-147, lot 131669), Alexa Fluor 488 donkey anti-mouse (715-545-151, lot 127820), Alexa Fluor 488-AffiniPure donkey anti-guinea pig (706-545-148, lot 138058), Alexa Fluor 647-AffiniPure donkey anti-guinea pig (706-605-148, lot 135631), Alexa Fluor 647-AffiniPure donkey anti-rabbit (711-605-152, lot 127614), carbocyanine (Cy)2-AffiniPure donkey anti-rabbit (711-225-152, lot 139999), Cy3-AffiniPure donkey anti-chicken (703-165-155, lot 142225), Cy3-AffiniPure donkey anti-goat (705-165-147, lot 134527), Cy3-AffiniPure donkey anti-guinea pig (706-165-148, lot 134844), Cy3-AffiniPure donkey anti-mouse (715-165-150, lot 116881), and Cy3-AffiniPure donkey anti-rabbit (711-165-152, lot 141941) and applied at a dilution of 1:300 in 0.1 M PB supplemented with 2% BSA (20–22 °C, 2 h). Nuclei were routinely counterstained with Hoechst 33,342 (1:10,000; Sigma). Tissues were photographed on a Zeiss LSM880 laser-scanning microscope. Images were acquired in the ZEN2010 software package. Multi-panel images were assembled in CorelDraw X7 (Corel Corp.).

RNA-scope in situ hybridization

C57Bl6/J mice were used to verify scRNA-seq candidate gene expression as described[97]. Dissected embryonic mouse heads were fixed in 4% PFA (pH 7.4) overnight. RNAscope 2.0 was performed according to the manufacturer's instructions (ISH, RNAscope, Advanced Cell Diagnostics)[98]. RNAscope probes for detection of *Slc1a3*, *Rax*, *Dll1*, *Dll3*, *Neurod1*, *Slit1*, *Slit2*, *Draxin*, *Prdm12*, *Nhlh2*, *Sox10*, *Ddc*, *Lancl3*, *Pmfbp1*, *Sncg*, *Sst*, *Th* and *Trh* were designed commercially by the manufacturer and are available from Advanced Cell Diagnostics. Imaging was performed using an LSM880 Zeiss confocal microscope equipped with a 40× objective.

Fluorescent in situ hybridization (HCR 3.0)

Staining was performed on fresh-frozen tissue sectioned at 16 μm following the HCR v3.0 protocol for 'generic sample on the slide' (Molecular Instruments)[99]. The pre-treatment of tissue sections included fixation with 4% PFA for 15 min, two washing steps with PBS and dehydration using an ascending EtOH gradient (25%, 50%, 75% and 100%, each step for 5 min with subsequent drying for 15 min). The tissue used for these experiments was obtained from E12.5, E15.5, E16.5 and E18.5 embryos or P2 and P7 pups. The probes used (*Ddc*, *Gad1*, *Meis2*,

Article

Onecut3, *Sncg*, *Th*, *Trh* and *Zic5*) were designed and purchased from Molecular Instruments.

In vitro overexpression of *Onecut3*

Neuro2A cells (ATCC; negative for mycoplasma upon visual inspection) were propagated in DMEM containing 4.5g/l glucose, glutamax, 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin (all from Gibco). Prior to transfection, cells were plated on glass coverslips (coated with poly-D-lysine (Sigma) at 37 °C overnight) at a density of 75,000 cells/well in a 24-well format. Cells were transfected with 500 ng of either OC3 or ABCD2 (an ATP-binding cassette transporter located on peroxisomes as CMV control) using the jetPRIME transfection system. The medium was replaced after 30 min with growth medium containing 2% FBS (to limit excessive proliferation) and cells were either immersion fixed for immunocytochemistry in 4% PFA in PBS (pH 7.4) for 20 min or lysed for qPCR after 3 days in vitro. Note that no cell death was observed due either to overexpression or the transfection reagent.

For immunocytochemistry, cells were permeabilized and blocked before adding a primary antibody cocktail overnight at 4 °C. Staining was performed with phospho-histone H3 (rabbit host; 1:500; Cell Signaling Technology), FLAG-tag (mouse host; 1:1,000; Sigma) or HA-tag (mouse host; 1:600; Cell Signaling Technology) and counterstained with Hoechst 33,342 (1:10,000; Sigma). After mounting with glycerol gelatine (Sigma), random overview images (20× magnification) were taken on an Zeiss LSM880 confocal microscope. Hoechst$^+$ and phospho-histone H3$^+$ cells were counted using ImageJ1.52a (cell counter plugin) and normalized to *Abcd2*.

For qPCR quantification, RNA was extracted with the Aurum Total RNA Kit (BioRad). A cDNA library was prepared by transcribing 2 µg RNA with the High Capacity RNA-to-cDNA Kit (Applied Biosystems). For qPCR reactions, 20 ng cDNA was amplified using SYBR green (BioRad) on a CFX Real Time Amplifier (BioRad) with [*Th* forward: TGTTTCAGTG CACACAGTAC]; [*Th* reverse: CCAATGTCCTGGGAGAACTG], [*Cxxc5* forward: AGTGGACAAAAGCAACCCTA]; [*Cxxc5* reverse: TTAGCATCTCTG TGGACTGT], [*Tmprss9* forward: GCTTGGTGCGACCCATCT]; [*Tmprss9* reverse: CATGGAGCCTCCCTCGC] and [*Tbp* for: CCTTGTACCCTTCACC AATGAC]; [*Tbp* rev: ACAGCCAAGATTCACGGTAGA] primers as loading control.

Preparation of acute brain slices

All experiments were performed in 300-µm-thick coronal slices prepared on a VT1200S vibratome (Leica) using a protective recovery method for slice preparation[100]. All constituents were from Sigma-Aldrich. Solutions were aerated with carbogen (5% CO_2/95% O_2).

Patch-clamp electrophysiology

Whole-cell recordings were carried out using borosilicate glass electrodes (Hilgenberg, Germany) of 3–4 MΩ pulled on a P-1000 instrument (Sutter). Electrodes were filled with an intracellular solution containing (in mM): 130 K-gluconate, 3 KCl, 4 ATP-Na_2, 0.35 GTP-Na_2, 8 phosphocreatine-Na_2, 10 HEPES, 0.5 ethyleneglycol-bis (2-aminoethylether)-*N,N,N',N'*-tetraacetate (EGTA), (pH 7.2 set with KOH) and 0.5% biocytin (Sigma) for post hoc cell identification. After recordings, brain slices were immersion fixed in 4% PFA in 0.1 M PB at 4 °C overnight. Recordings were carried out on an EPC-10 triple amplifier (HEKA) controlled by PatchMaster 2.80.

Sample sizes, statistics and reproducibility

Sample sizes for scRNA-seq experiments: $n = 8$ (E15.5), $n = 7$ (E17.5), $n = 4$ (P0), $n = 4$ (P2), $n = 3$ (P10), $n = 3$ (P23). SCTransform corrected UMI-count matrices were statistically tested to obtain DEGs using log-normalized values with pseudocount = 1 for 45 identified cell groups as previously described[6,51] using MAST test[52]. We used the Wilcoxon test and Logreg test[87] to define upregulated regulons. Results of the DGE tests are specified at https://doi.org/10.6084/m9.figshare.11867889.

For Fig. 1a, a UMAP plot was built for $n = 51,199$ cells of ectodermal origin integrated by canonical correlation analysis (CCA). For Fig. 1c, $n = 5,070$ cells were used for UMAP. For Fig. 1d (right), we sampled cell-ordering ($n = 5,070$ cells) to 500 bins to compute gene expression trends. For each bin, mean ± s.e.m. expression was estimated by generalized additive models. For Fig. 1e, $n = 2$ animals for E12.5, 3 $n =$ animals for E13.5, and $n = 3$ animals for E15.5 were used for embryonal tracing of *Ascl1* at corresponding time points. For Fig. 1f, $n = 2$ animals were used to trace *Ascl1*$^+$ cells in (BAC)GAD65−eGFP:*Ascl1-creER^{T2}::Ai14* mice (tamoxifen injection at P14, sample collection at P21). For Fig. 2b (top right), to compute gene trends we sampled the trajectory for POMC neurons ($n = 1,643$ cells from progenitors to mature POMC cells) to 500 bins. Each single trajectory shows mean ± s.e.m. estimated by generalized additive models. For Fig. 2b (bottom), RNA scope in situ hybridization was performed on samples from POMC−GFP mice ($n = 4$ for *Prdm12* and $n = 3$ for *Nlh2*). For Fig. 2e, experiment was performed in triplicate with $n = 2$ animals per time point. For Fig. 3b, images are representative of wild-type ($n = 3$) and *Nfia*$^{-/-}$ ($n = 3$) mice. For Fig. 3f, $n = 5,070$ cells belong to glia (excluding oligodendrocytes) or bridge neurons. To compute gene expression trends, we sampled cell-ordering to 500 bins. For each bin, mean ± s.e.m. was estimated by generalized additive models. For Fig. 3g, $n = 5$ (wild-type), $n = 5$ (*Robo1*$^{-/-}$), $n = 3$ (*Slit1*$^{-/-}$) and $n = 3$ (*Slit2*$^{-/-}$) mice were used. Data are presented as percentile box-whisker plots (10th, 25th, 50th, 75th and 90th percentiles). Data were statistically analysed using one-tailed Student's t-test of raw data: $P = 0.0019$ for wild-type versus *Slit1*$^{-/-}$; $P = 0.0006$ for wild-type versus *Slit2*$^{-/-}$. For Fig. 3h, the experiment was performed in duplicate with wild-type ($n = 3$) and *Robo1*$^{-/-}$ ($n = 4$) mice. For Fig. 4b–d, f, h, i, to compute gene trends we independently sampled seven differentiation trajectories containing neurons (for group 1: 1,506 cells, for group 2: 997 cells, for group 3: 1,453 cells, for group 5: 948 cells, for group 6: 1,779 cells and for group 9: 1,181 cells) to 500 bins. For each bin, mean ± s.e.m. expression was estimated by generalized additive models. For Fig. 4b (right), $n = 2$ animals for E12.5, $n = 3$ animals for E13.5. The experiment was performed in duplicate. For Fig. 4e, $n = 2$ animals were used. The experiment was reproduced twice. For Fig. 4f (right), $n = 3$ animals were used in the experiment. For Fig. 4g (right), $n = 2$ animals were tested in each independent experiment. For Fig. 4h (right), $n = 2$ animals were tested in the experiment. For Fig. 4i, $n = 5$ (wild-type) and $n = 6$ (*Robo1*$^{-/-}$) mice were tested. Data are visualized as error bar plots with individual data point information. Data were statistically evaluated using one-tailed Student's t-test on raw data: $P = 0.015$ for wild-type versus *Robo1*$^{-/-}$ at the anatomical (rostral-to-caudal) level of the SCN. For Fig. 4k, post hoc neuroanatomical reconstruction identified that all $n = 9$ *Onecut3*$^+$ neurons were of 'type C' in A14 ($n = 15$ cells were characterized as 'type C' within 62 cells recorded in total from A14).

For Extended Data Fig. 1, we used the dataset of $n = 51,199$ ectodermal cells for dot-plot representation (Extended Data Fig. 1a, b) and UMAP visualizations (Extended Data Fig. 1b, c). For Extended Data Fig. 2a, we used the same dataset of $n = 51,199$ ectodermal cells for each alignment algorithm. For Extended Data Fig. 2c, $n = 6,314$ cells from E15.5 passed filters of original RNA velocity analysis and are presented in the figure panels for each analysis of gene expression. For Extended Data Fig. 2d, numbers of animals used for embryonal tracing are: $n = 2$ for E12.5, $n = 3$ for E13.5, $n = 3$ for E15.5 and $n = 2$ for E16.5. For Extended Data Fig. 2d, freely available data from Allen's Mouse Developmental Brain Atlas were used. For Extended Data Fig. 2f, images are representative of *Ascl1*$^{+/-}$ ($n = 3$) and *Ascl1*$^{-/-}$ ($n = 3$) mice. For Extended Data Fig. 2g, three animals were checked for each developmental stage. For Extended Data Fig. 2h, two *Ascl1*$^+$ cells were traced in (BAC)GAD65−eGFP:*Ascl1-creER^{T2}::Ai14* animals for the time-point shown (tamoxifen injection at P14, sample collection at P21). For Extended Data Fig. 3d, e, analysis was done based on $n = 33,893$ cells of neuronal and glial origin (excluding oligodendrocytes). For Extended Data Fig. 3f, the experiment was reproduced three times with $n = 2$ animals per time-point. For Extended Data

Fig. 4, n = 33,893 cells in total were tested with two-tailed Wilcoxon rank sum test. For Extended Data Fig. 6b, images are representative of samples from n = 2 animals. For Extended Data Fig. 6c, images are representative of n = 2 animals per experiment. For Extended Data Fig. 6d (top), images are representative of the experiment reproduced twice with n = 6 coverslips. For Extended Data Fig. 6d (bottom), the sample size was n = 12 for each group. Data were statistically evaluated using two-tailed Student's t-test on raw data. P = 0.0000292 for the density of Hoechst-positive cells; P = 0,0000000103 for the density of pHH3$^+$ cells. For Extended Data Fig. 6e, there were three biological and three technical replicates for each probe. Data are visualized as error bar plots (mean ± s.e.m.) with individual data point information. Data were statistically evaluated using two-tailed Student's t-test on raw data. P = 0.0210 for Th expression; P = 0.000153 for $Cxxc5$ expression. For Extended Data Fig. 7a (UMAP), n = 33,893 cells were used for graph construction. For Extended Data Fig. 7a (bottom), to compute gene expression trends we sampled n = 5,070 cells to 500 bins. For each bin, mean ± s.e.m. expression was estimated by generalized additive models. For Extended Data Fig. 7a (right), b, all images represent results from n = 3 animals for each experiment. For Extended Data Fig. 8a, we recorded n = 20 cells from A12 (including n = 12 of type A, n = 3 of type B, n = 5 of type C), n = 8 cells from A13 (n = 5 cells of type A and n = 3 cells of type B) and n = 62 cells from A14. Specifically for A14, we distinguished four electrophysiological profiles: n = 32 cells of type A, n = 9 cells of type B, n = 15 cells of type C and n = 6 cells of type D. Post hoc neuroanatomical reconstruction identified that all n = 9 $Onecut3$$^+$ neurons were of type C in A14. For Extended Data Fig. 8b, images are representative of the experiment reproduced nine times. For Extended Data Fig. 9a, images are representative for $Ascl$$^{+/-}$ (n = 3) and $Ascl1$$^{-/-}$ (n = 3) mice. The immunohistochemical experiment was reproduced twice. For Extended Data Fig. 9b, n = 3 animals from E13.5 tamoxifen injection, n = 3 animals from E15.5 tamoxifen injection, and n = 2 mice from E16.5 tamoxifen injection were tested for embryonal tracing of $Ascl1$ cells at corresponding time points at E18. For Extended Data Fig. 9c, to compute gene trends we independently sampled seven differentiation trajectories containing neurons (for group 1: 1,506 cells, for group 2: 997 cells, for group 3: 1,453 cells, for group 5: 948 cells, for group 6: 1,779 cells and for group 9: 1,181 cells) to 500 bins. For each bin mean ± s.e.m. expression was estimated by generalized additive models. For Extended Data Fig. 9d, n = 3 animals were tested. For Extended Data Fig. 9e (images), n = 2 animals were used for each developmental time-point. The experiment was reproduced twice. For Extended Data Fig. 9e (scatter plot), for both ages n = 2 animals were tested. In total 86 Th-containing (minimum >2 mRNA molecules) cells were randomly analysed. Data are visualized as error bar plots (mean ± s.e.m.) with individual data point information. For Extended Data Fig. 9f, images are representative of n = 2 animals from each time-point. For Extended Data Fig. 10a, images are representative for the experiment performed in duplicate on n = 4 for each developmental stage of both mouse lines. For Extended Data Fig. 10b, n = 26,316 cells from the neuronal lineage (including progenitors) were used. For Extended Data Fig. 10c, n = 3 animals were tested for each developmental time-point. For Extended Data Fig. 10d, n = 2 animals were tested for each independent experiment. For Extended Data Fig. 10e, n = 2 animals were tested for the experiment. For Extended Data Fig. 10f, n = 9 $Onecut3$$^+$ neurons were reconstructed.

Reporting summary

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

Data availability

Raw, processed and supplementary datasets have been deposited in GEO (accession number: GSE132730). GEO files include: 1) raw fastq files for every sequencing run; 2) filtered matrices for every sample in RDS file format including Seurat 3 objects with all processed cells; 3) original integrated dataset in RDS file format including Seurat 3 objects with all processed cells as well as all used commands; 4) integrated dataset used for dynamics analysis (which passed filtering of RNA velocity analysis); 5) AUCell matrices from pySCENIC pipeline; 6) full regulon hypothalamic network in GraphML file format; 7) metadata protocol describing all experimental, computational procedures and quality control. An interactive view of the integrated dataset (for processing in Pagoda2) can be accessed at https://doi.org/10.6084/m9.figshare.11867889 (~1.1 GB). All data presented (for example, imaging) will be made available by T. Harkany (tibor.harkany@ki.se or tibor.harkany@meduniwien.ac.at) upon reasonable request.

Code availability

The code used is available at https://doi.org/10.6084/m9.figshare.11867889.

30. Matsushita, N. et al. Dynamics of tyrosine hydroxylase promoter activity during midbrain dopaminergic neuron development. *J. Neurochem.* **82**, 295–304 (2002).
31. Cowley, M. A. et al. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* **411**, 480–484 (2001).
32. Kim, E. J., Ables, J. L., Dickel, L. K., Eisch, A. J. & Johnson, J. E. Ascl1 (Mash1) defines cells with long-term neurogenic potential in subgranular and subventricular zones in adult mouse brain. *PLoS One* **6**, e18472 (2011).
33. 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).
34. Shu, T., Butz, K. G., Plachez, C., Gronostajski, R. M. & Richards, L. J. Abnormal development of forebrain midline glia and commissural projections in Nfia knock-out mice. *J. Neurosci.* **23**, 203–212 (2003).
35. Bali, B., Erdélyi, F., Szabó, G. & Kovács, K. J. Visualization of stress-responsive inhibitory circuits in the GAD65-eGFP transgenic mice. *Neurosci. Lett.* **380**, 60–65 (2005).
36. Tamamaki, N. et al. Green fluorescent protein expression and colocalization with calretinin, parvalbumin, and somatostatin in the GAD67-GFP knock-in mouse. *J. Comp. Neurol.* **467**, 60–79 (2003).
37. Andrews, W. et al. Robo1 regulates the development of major axon tracts and interneuron migration in the forebrain. *Development* **133**, 2243–2252 (2006).
38. Plump, A. S. et al. Slit1 and Slit2 cooperate to prevent premature midline crossing of retinal axons in the mouse visual system. *Neuron* **33**, 219–232 (2002).
39. Yang, L. et al. Isl1Cre reveals a common Bmp pathway in heart and limb development. *Development* **133**, 1575–1585 (2006).
40. Yoshida, M. et al. Evidence that oxytocin exerts anxiolytic effects via oxytocin receptor expressed in serotonergic neurons in mice. *J. Neurosci.* **29**, 2259–2271 (2009).
41. Bäckman, C. M. et al. Characterization of a mouse strain expressing Cre recombinase from the 3′ untranslated region of the dopamine transporter locus. *Genesis* **44**, 383–390 (2006).
42. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
43. Lun, A. T. L. et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
44. Petukhov, V. et al. dropEst: pipeline for accurate estimation of molecular counts in droplet-based single-cell RNA-seq experiments. *Genome Biol.* **19**, 78 (2018).
45. van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
46. Hinman, M. N. & Lou, H. Diverse molecular functions of Hu proteins. *Cell. Mol. Life Sci.* **65**, 3168–3181 (2008).
47. Mili, S. & Steitz, J. A. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* **10**, 1692–1694 (2004).
48. Yoo, S. & Blackshaw, S. Regulation and function of neurogenesis in the adult mammalian hypothalamus. *Prog. Neurobiol.* **170**, 53–66 (2018).
49. Miranda-Angulo, A. L., Byerly, M. S., Mesa, J., Wang, H. & Blackshaw, S. Rax regulates hypothalamic tanycyte differentiation and barrier function in mice. *J. Comp. Neurol.* **522**, 876–899 (2014).
50. Shimogori, T. et al. A genomic atlas of mouse hypothalamic development. *Nat. Neurosci.* **13**, 767–775 (2010).
51. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
52. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
53. Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
54. Alpár, A., Benevento, M., Romanov, R. A., Hökfelt, T. & Harkany, T. Hypothalamic cell diversity: non-neuronal codes for long-distance volume transmission by neuropeptides. *Curr. Opin. Neurobiol.* **56**, 16–23 (2019).
55. Romanov, R. A., Alpár, A., Hökfelt, T. & Harkany, T. Unified classification of molecular, network, and endocrine features of hypothalamic neurons. *Annu. Rev. Neurosci.* **42**, 1–26 (2019).
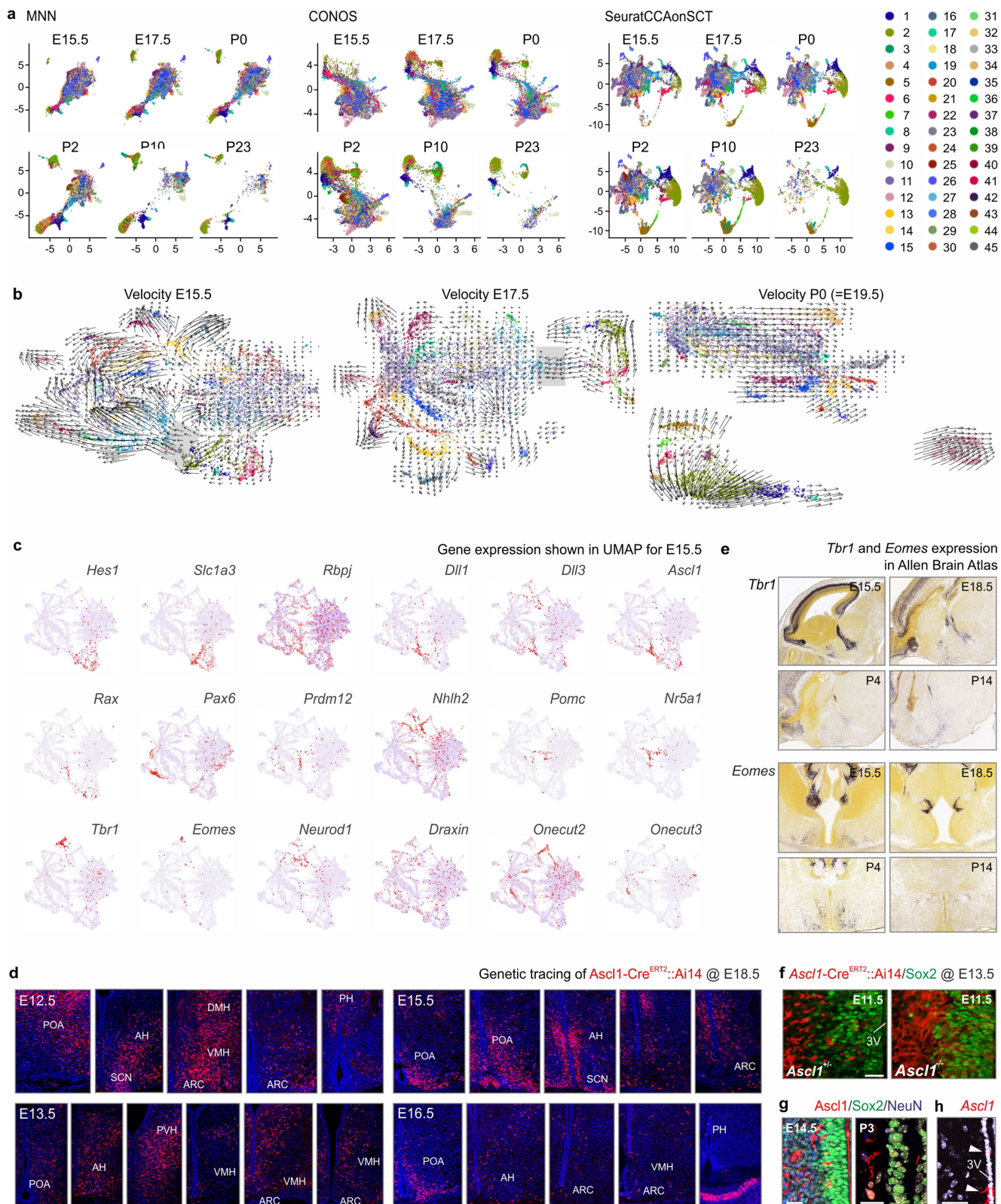
Article

56. Polański, K. et al. BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964–965 (2020).

57. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *JOSS* **3**, 861 (2018).

58. Johnson, J., Douze, M. & Jegou, H. Billion-scale similarity search with GPUs. *IEEE Trans. Big Data* https://doi.org/10.1109/TBDATA.2019.2921572 (2019).

59. Maneewongvatana, S. & Mount, D. in Data Structures, Near Neighbor Searches, and Methodology: Fifth and Sixth DIMACS Implementation Challenges Vol. 59 DIMACS Series in Discrete Mathematics and Theoretical Computer Science (eds. Goldwasser, M., Johnson, D. & McGeoch, C.) 105–123 (American Mathematical Society, 2002).

60. Hie, B., Bryson, B. & Berger, B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat. Biotechnol.* **37**, 685–691 (2019).

61. Welch, J. D. et al. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887.e1817 (2019).

62. Korsunsky, I. et al. Fast, sensitive, and accurate integration of single cell data with harmony. *Nature Methods* **16**, 1289–1296 (2019).

63. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).

64. Barkas, N. et al. Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nat. Methods* **16**, 695–698 (2019).

65. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).

66. Kowalczyk, M. S. et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* **25**, 1860–1872 (2015).

67. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).

68. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat. Commun.* **9**, 884 (2018).

69. Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-cell RNA-seq reveals hypothalamic cell diversity. *Cell Rep.* **18**, 3227–3241 (2017). https://doi.org/10.1016/j.celrep.2017.03.004.

70. Moffitt, J. R. et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018).

71. Vinh, N. X., Epps, J. & Bailey, J. Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. *J. Mach. Learn. Res.* **11**, 2837–2854 (2010).

72. Chang, F., Qiu, W., Zamar, R. H., Lazarus, R. & Wang, X. clues: an r package for nonparametric clustering based on local shrinking. *J. Stat. Softw.* **33**, (2010).

73. Pons, P. & Latapy, M. in *Intl Symp. Computer Information Sci.* 284–293 (Springer, 2005).

74. Kurrasch, D. M. et al. The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. *J. Neurosci.* **27**, 13624–13634 (2007).

75. Zhang, Y. et al. Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron* **89**, 37–53 (2016).

76. Kelley, K. W., Nakao-Inoue, H., Molofsky, A. V. & Oldham, M. C. Variation among intact tissue samples reveals the core transcriptional features of human CNS cell classes. *Nat. Neurosci.* **21**, 1171–1184 (2018).

77. Lin, Y. et al. Evaluating stably expressed genes in single cells. *Gigascience 8, giz106* (2019).

78. Extra Coordinate Systems. 'Geoms', Statistical Transformations, Scales and Fonts for 'ggplot2' [R package ggalt version 0.4.0] v. 0.4.0 https://CRAN.R-project.org/package=ggalt (2017).

79. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

80. Zhang, J. M., Kamath, G. M. & Tse, D. N. Valid post-clustering differential analysis for single-cell RNA-seq. *Cell Syst.* **9**, 383–392.e6 (2019).

81. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

82. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. Preprint at https://www.biorxiv.org/content/10.1101/820936v1 (2019).

83. Moerman, T. et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161 (2019).

84. Kanamori, M. et al. A genome-wide and nonredundant mouse transcription factor database. *Biochem. Biophys. Res. Commun.* **322**, 787–793 (2004).

85. Ravasi, T. et al. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* **140**, 744–752 (2010).

86. Forrest, A. R. et al. A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).

87. Ntranos, V., Yi, L., Melsted, P. & Pachter, L. Identification of transcriptional signatures for cell types from single-cell RNA-seq. *Nat. Methods* **16**, 163–166 (2019).

88. Bastian, M., Heymann, S. & Jacomy, M. *Gephi: An Open Source Software for Exploring and Manipulating Networks* https://gephi.org/publications/gephi-bastian-feb09.pdf (2009).

89. Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat. Genet.* **50**, 1593–1599 (2018).

90. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).

91. qvalue: Q-value estimation for false discovery rate control v. 2.18.0 (R package, 2019).

92. Galili, T., O'Callaghan, A., Sidi, J. & Sievert, C. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* **34**, 1600–1602 (2018).

93. Setty, M. et al. Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).

94. Levine, J. H. et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).

95. Espana, A. & Clotman, F. Onecut transcription factors are required for the second phase of development of the A13 dopaminergic nucleus in the mouse. *J. Comp. Neurol.* **520**, 1424–1441 (2012).

96. Miyazaki, T., Fukaya, M., Shimizu, H. & Watanabe, M. Subtype switching of vesicular glutamate transporters at parallel fibre-Purkinje cell synapses in developing mouse cerebellum. *Eur. J. Neurosci.* **17**, 2563–2572 (2003).

97. Bartesaghi, L. et al. PRDM12 is required for initiation of the nociceptive neuron lineage during neurogenesis. *Cell Rep.* **26**, 3484–3492.e4 (2019).

98. Wang, F. et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012).

99. Choi, H. M. T. et al. Mapping a multiplexed zoo of mRNA expression. *Development* **143**, 3632–3637 (2016).

100. Zhao, S. et al. Cell type–specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nat. Methods* **8**, 745–752 (2011).

Extended Data Fig. 1 | Marker genes to define molecular phenotypes.
a, Differential gene expression by glia (clusters 1–9) and neurons (clusters 10–45). Because of the integration of six stages, early-expressed TFs and spatially restricted genes amenable to cellular differentiation were identified. For neuronal clusters, fast neurotransmitter specificity is shown to the right. The relative diameter of the solid circles for each cluster is scaled to the fraction of cells that expresses a specific gene. Colour coding and numbering at the top correspond to those in Fig. 1a. **b**, Top, Differential TF expression in 45 ectoderm-derived cell groups in the hypothalamus. Bottom, subclass-specific TFs recapitulate the UMAP positions of neuronal (left) and glial (right) subtypes. **c**, Integrated molecular and anatomical annotation of hypothalamic clusters with their specific assignment to hypothalamic areas. ARC-Agrp, arcuate nucleus agouti-related peptide[+] neurons; ARC-Sst, arcuate nucleus somatostatin[+] neurons; ARC-TIDA, arcuate nucleus–tuberoinfundibular dopamine neurons; Gal, galanin; Ghrh/Vacht, growth hormone-releasing hormone/vesicular acetylcholine transporter[+] neurons; LH, lateral hypothalamus; LH-Lhx9, lateral hypothalamus LIM homeobox 9[+] cluster; Meis2, meis homeobox 2; MM, mammillary nucleus; MM-Lhx9, mammillary nucleus LIM homeobox 9[+] neurons; Pomc, proopiomelanocortin; PH, posterior hypothalamus; PMM, premamillary nucleus; Tbr1, T-box brain transcription factor 1.

Article

Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Molecular analysis of TFs involved in neurogenesis and neuronal differentiation. a, Comparative and time-resolved analysis of the cell bridge by MNN, CONOS and Seurat alignment. In UMAP space on separate developmental stages, MNN, CONOS and Seurat algorithms were compared for their ability to specifically resolve the transition of progenitors to immature cells (bridge). Colour codes correspond to those in Fig. 1a. **b**, RNA velocity at E15.5, E17.5 and P0. Colour codes are consistent with those in Fig. 1a. Note the presence of a bridge (grey background) between progenitor/glial and neuronal compartments at early developmental stages with its rupture being evident by birth. **c**, Gene expression in UMAP space at E15.5. Note a central role for Notch signalling in neurogenesis. **d**, Genetic tracing of $Ascl1^+$ cells produced in the developing hypothalamus during the E12.5–E16.5 period. **e**, In situ hybridization showing the distribution of $Tbr1$ and $Eomes$. Image credit: Allen Institute (https://www.brain-map.org). **f**, Genetic tracing of $Ascl1^+$ cells in the developing hypothalamus of $Ascl^{+/-}$ and $Ascl^{-/-}$ mice. Sox2 was localized by immunohistochemistry. **g**, $Sox2$, $Ascl1$ (Tomato) and $Rbfox3$ (NeuN) immunolocalization at successive developmental stages. **h**, Genetic tracing of $Ascl^+$ cells postnatally (as in Fig. 1f). Scale bars, 200 μm (**d**), 20 μm (**f**–**h**).
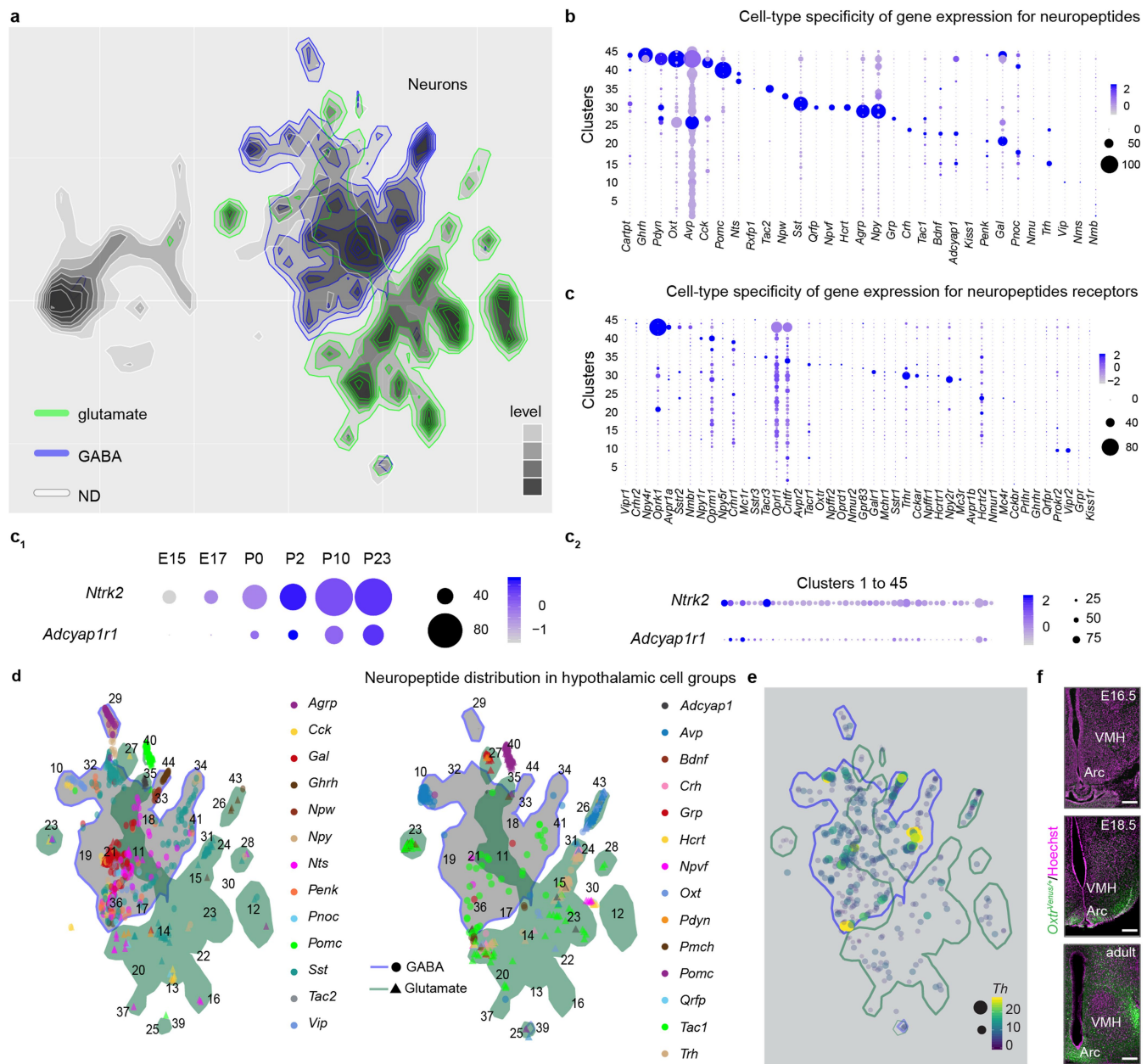
Extended Data Fig. 3 | Neurotransmitter and neuropeptide specificity and load in the developing hypothalamus. a–c, Coincident profiling of fast neurotransmitters (**a**), neuropeptides (**b**) and neuropeptide receptors (**c**, top) in 45 cell groups of ectodermal origin. **c**, Bottom left, given their abundance, *Ntrk2* and *Adcyap1r1* were plotted separately along the developmental timeline studied with appropriate scaling. Bottom right, likewise, the distribution of both receptors per cell cluster was mapped and scaled separately.

d, Coincident profiling of neuropeptides in neuronal clusters distinguished as GABA (blue) and glutamate (grey) phenotypes. **e,** Map of *Th* expression in GABA and glutamate neurons. Colour coding as in **d. f,** Developmental mapping of hypothalamic *Oxtr* expression in *Oxtr*[Venus/+] mice. Low-magnification image surveys are shown (see also Fig. 2e). Scale bars, 200 μm (**f**). Data shown as dot plots and scaled as previously described[6,51,65].
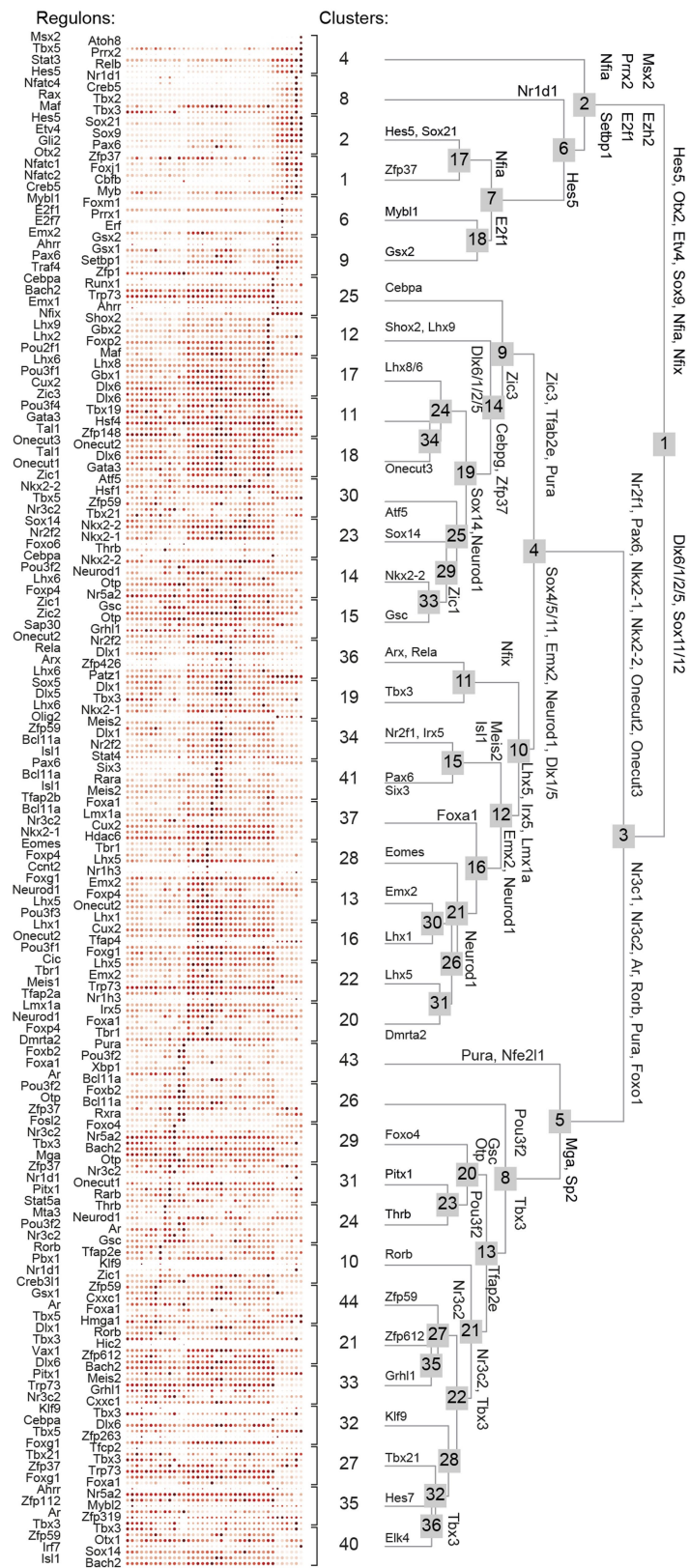
Extended Data Fig. 4 | Hierarchical relationship of GRNs (regulons). An AUC separability plot was used to assign regulons that determine cell cluster identities identified in SCENIC[23]. GRNs were reconstructed individually for each cell and then assigned as 'regulon representation' (Logreg test) to each cell group. TFs to the left are representative for each regulon. Marked dendrogram branchpoints were estimated by both the Wilcoxon and Logreg tests (see also https://doi.org/10.6084/m9.figshare.11867889).
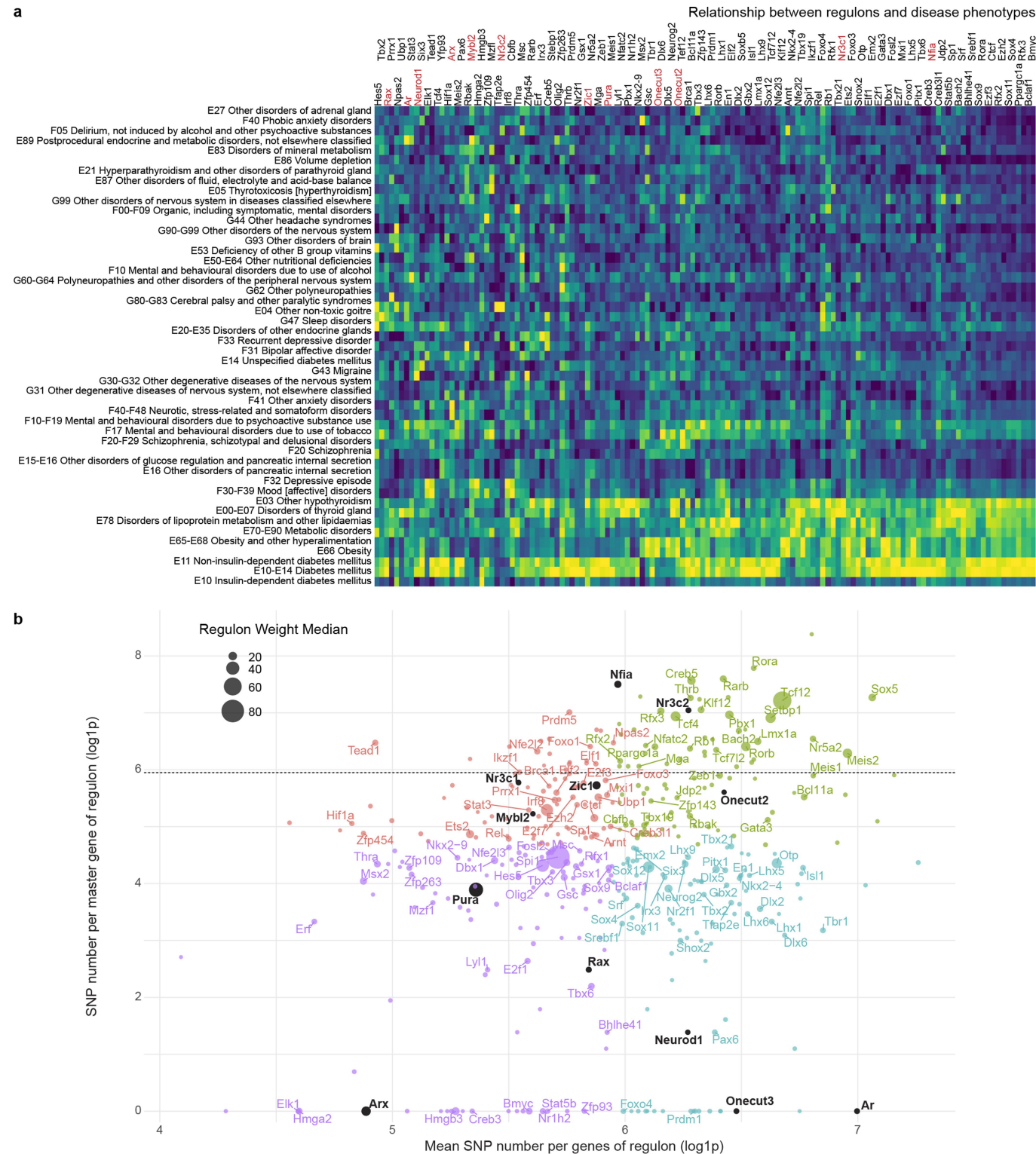
Extended Data Fig. 5 | Relationships between regulons and disease phenotypes in humans. a, Complete heat map of associations between regulon activity and clinical disease phenotype. Left: classifications of diseases as per phenotypic criteria of the UK biobank registry (https://www.ukbiobank.ac.uk). Top, master genes for each regulon. Genes presented in Fig. 3 are in red and highlighted in **b**. Colour coding from deep blue to bright yellow shows increasing correlation probability. **b**, Scatter plot reflecting the ratios of mutability in master genes versus all downstream target genes per regulon. Mutability and the constrains of TFs were expressed as the total number of mutations. Colours represent four quadrants that were separated on the basis of the total number of mutations per master gene (medians, y-axis) versus target genes (medians, x-axis). Horizontal line corresponds to the median of SNPs in all genes. Dot size reflects the median influence of a given regulon on its targets as per SCENIC output.
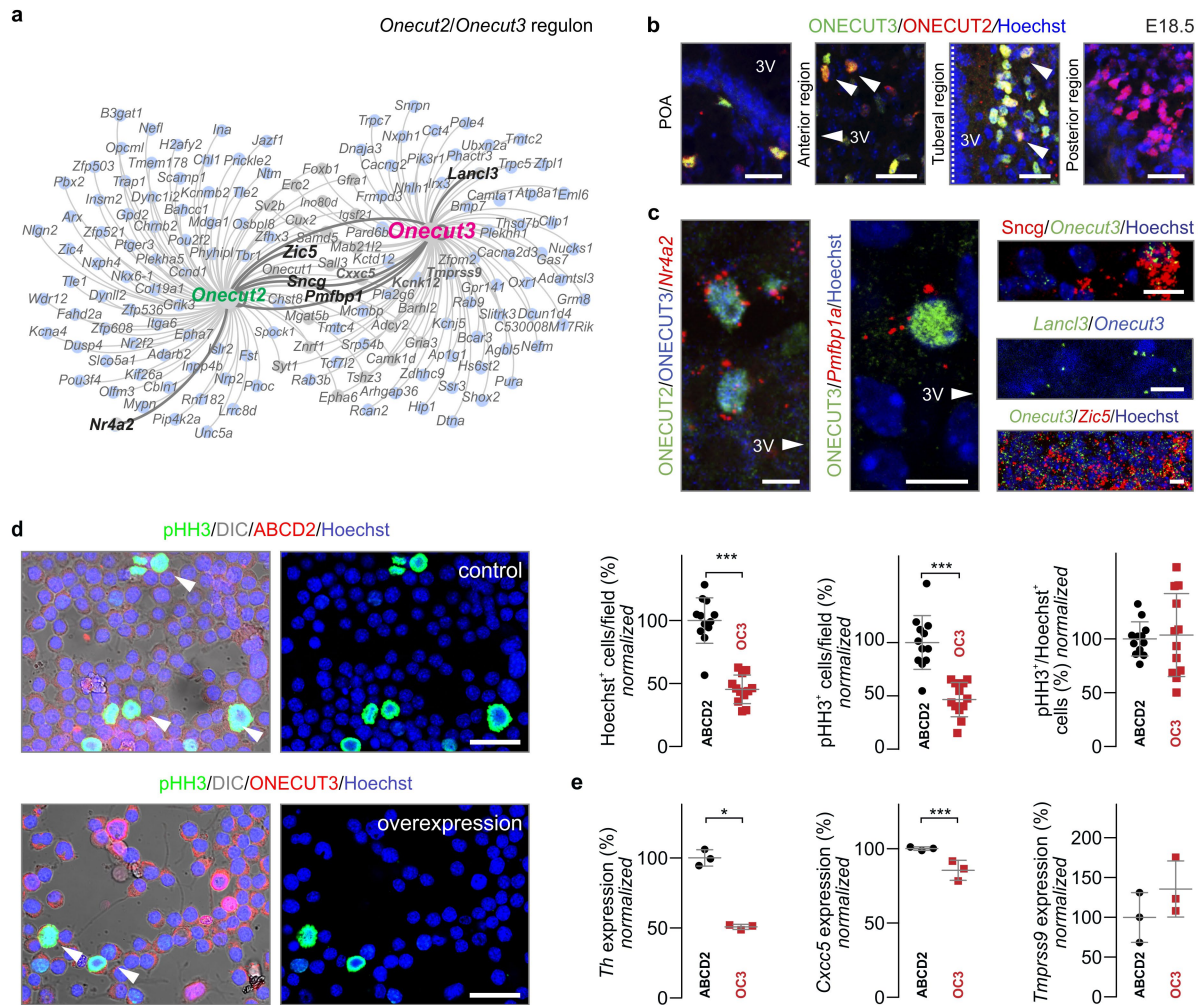
Extended Data Fig. 6 | Molecular complexity and function of the *Onecut3* regulon. a, Interlinked *Onecut2* and *Onecut3* regulons in hypothalamic neurons. Genes that were biologically validated (see below) are shown in black. **b**, Co-expression of *Onecut2* and *Onecut3* along the rostrocaudal axis of the hypothalamus. **c**, Co-localization of *Onecut3* and its target genes (from **a**). **d**, Overexpression of *Onecut3* (OC3) and ATP-binding cassette D2 (*Abcd2*, to control promoter activity) in Neuro2A cells. Left, representative images by multiple fluorescence labelling-differential interference microscopy. Right, quantification of Hoechst+ and phospho-histone H3 (pHH3)+ Neuro2A cells revealed significantly reduced proliferation upon *Onecut3* overexpression. No significant cell death was induced by either overexpressed plasmid or the transfection reagent alone. **e**, qPCR analysis of genes regulated by *Onecut3*: *Cxxc5*, *Tmprss9* and *Th*. All data were normalized to samples transfected with *Abcd2*, which were taken as technical controls. Scale bars, 50 μm (**d**), 20 μm (**b**, **f**), 10 μm (**g**).
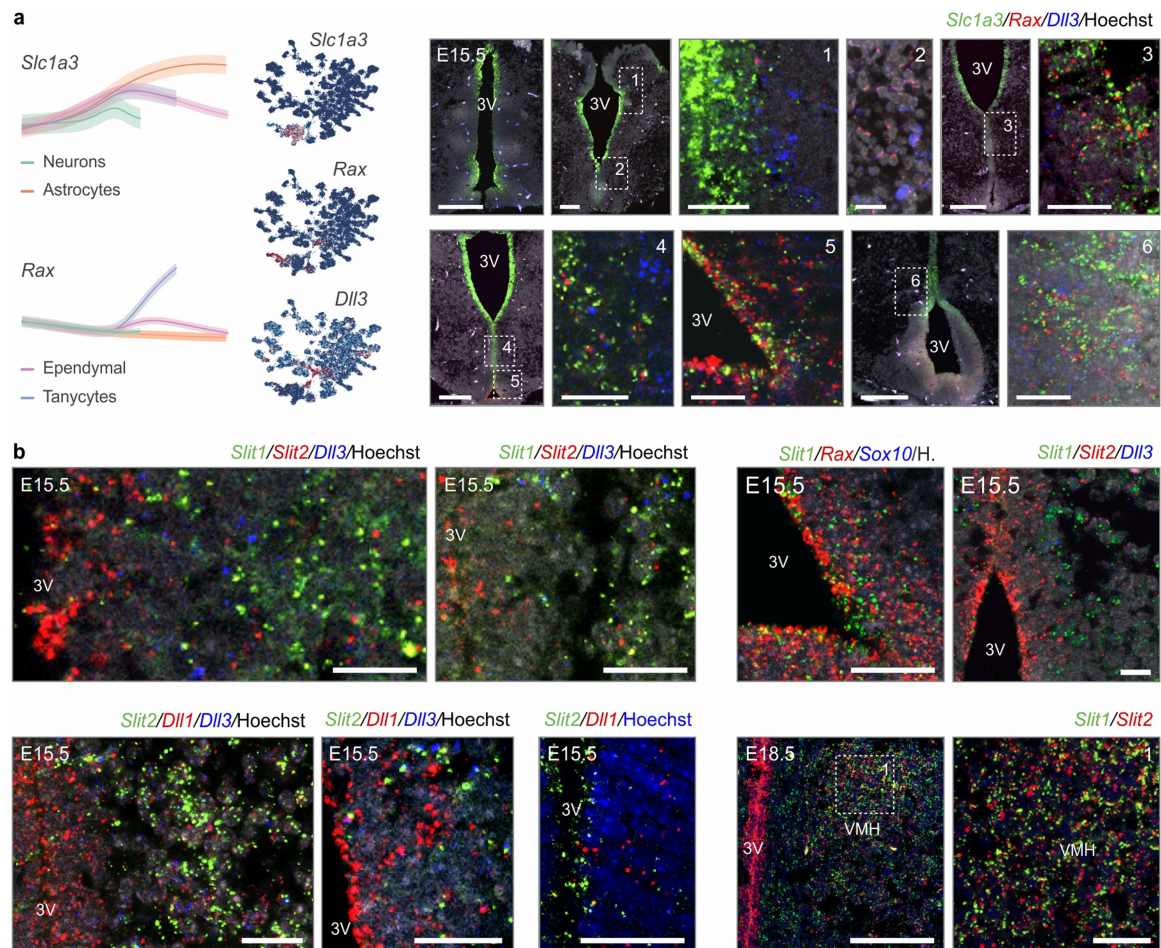
Extended Data Fig. 7 | Experimental validation of ventricle-restricted genes identified by scRNA-seq. a, Left, expressional dynamics of ventricle-associated marker genes: *Slc1a3*, *Rax* and *Dll3* on UMAP embedding (top) and trend lines (bottom). Right, validation by in situ hybridization. **b**, In situ hybridization for the co-existence of *Slit2* and *Rax* in ventricular progenitors and consequential medial-to-lateral *Slit1–Dll1–Dll3* patterns during neuronal differentiation and migration by E15.5 (left, top right). Left-to-right orientation corresponds to medial-to-lateral hypothalamic positions. Bottom right, localization of *Slit1* and *Slit2* mRNAs in the VMH at E18.5. Scale bars, 200 μm (**a**), 20 μm (**b**).
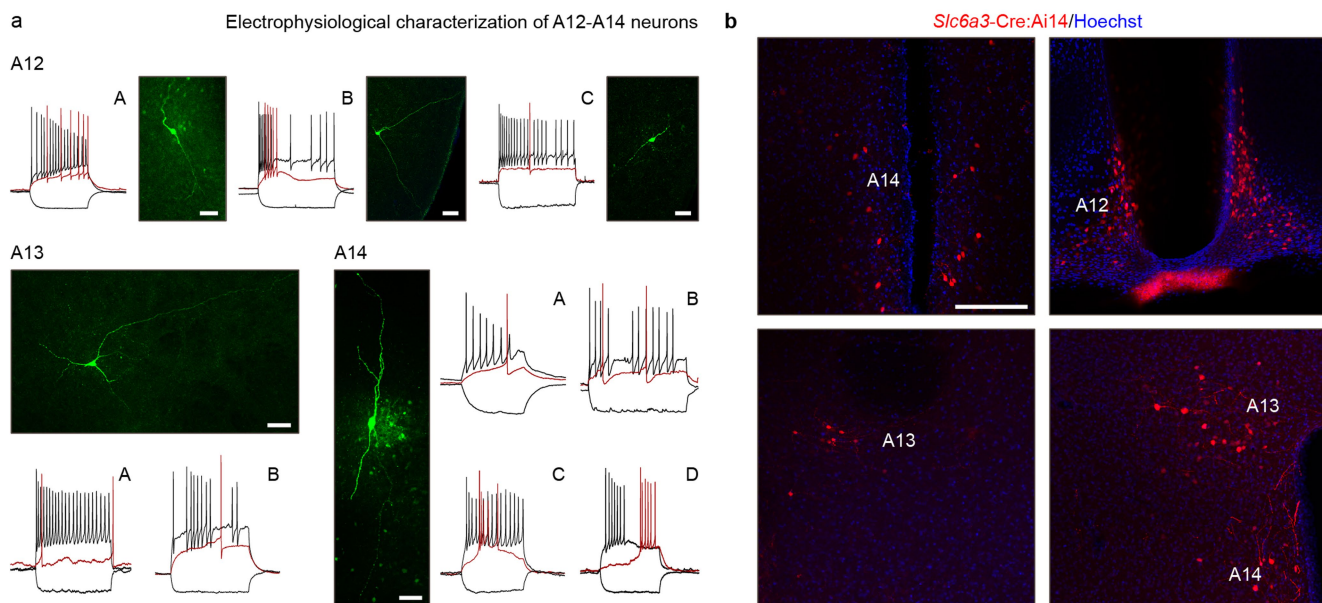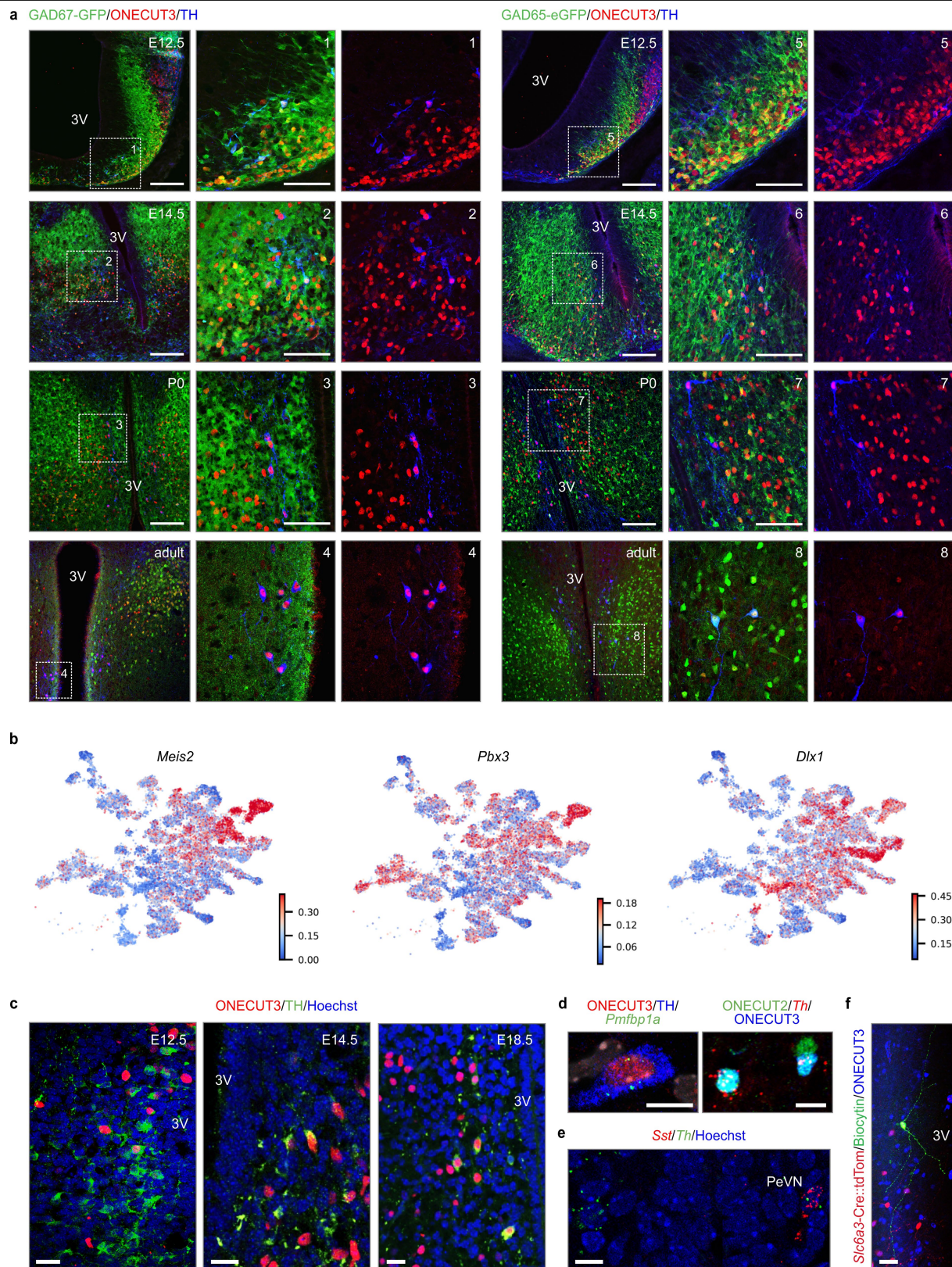
Extended Data Fig. 8 | Physiological and morphological subtypes of hypothalamic dopamine neurons. a, Action potential waveforms of dopamine neurons within the A12–A14 groups. Note the diversification of A14 dopamine cells into subgroups A–D with clearly different action potential signatures. Morphological reconstruction of biocytin-filled neurons is shown with each group. **b**, Distribution of *tdTomato*⁺ neurons in the hypothalamus of *Slc6a3*-Ires-*cre::Ai14* mice. Scale bars, 50 μm (**b**), 20 μm (**a**).
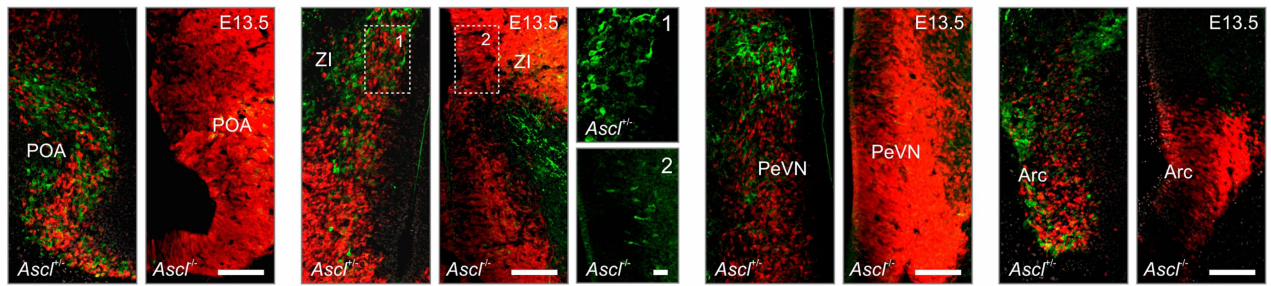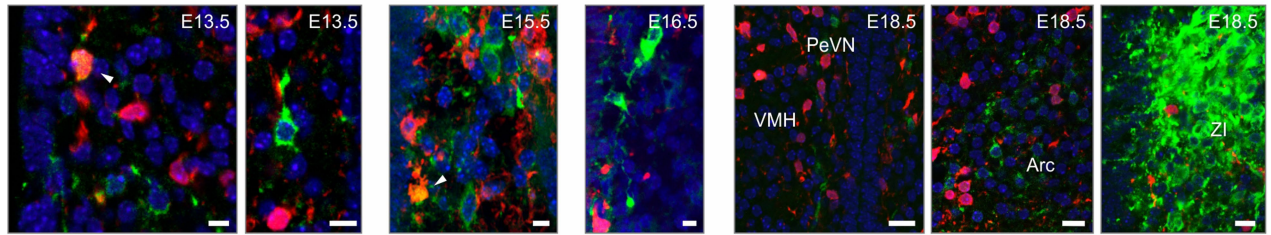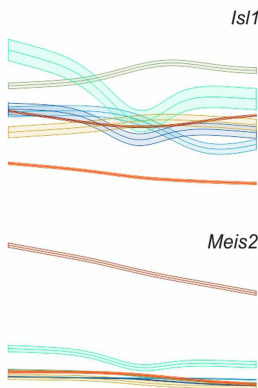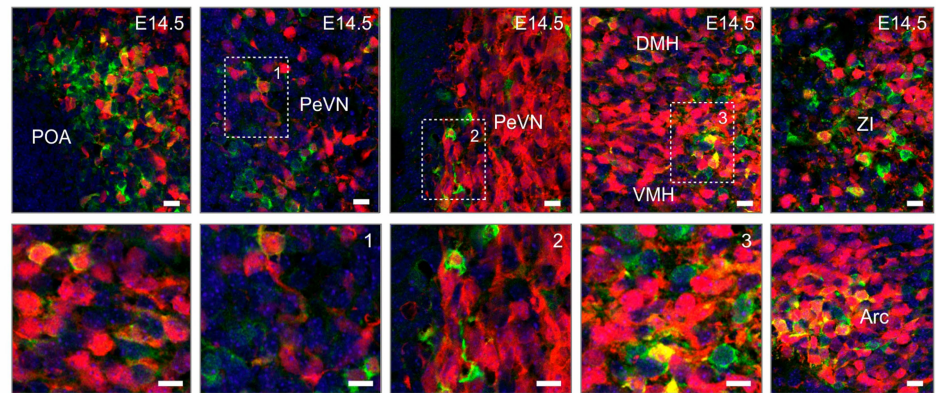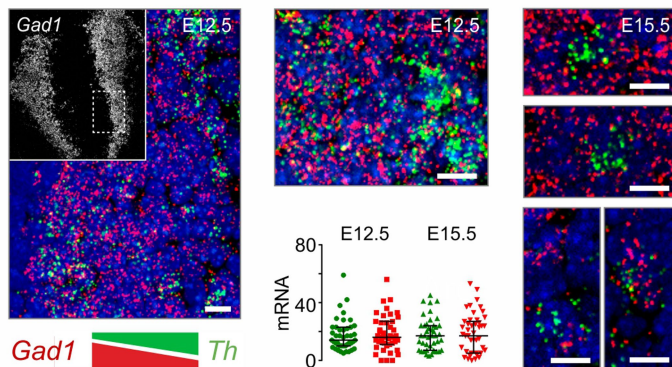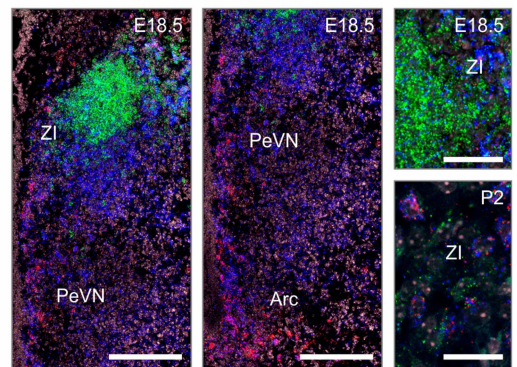
Extended Data Fig. 9 | Transcriptional and physiological features of dopamine neurons in the developing hypothalamus. a, *Ascl1-creER^T2/+::Ai14* (control) versus *Ascl1-cre^ERT2/ERT2::Ai14* mice (a knock-in mouse line with Cre disrupting the *Ascl1* gene, referred to as *Ascl1* ko), injected with tamoxifen at E11.5 and analysed at E13.5. Note the accumulation of tdTomato⁺ cells in the KO relative to controls. **b**, Genetic tracing in *Ascl1-creER^T2::Ai14* reporter mice identified *Ascl1⁺/Th⁺* neurons within the preoptic and periventricular nuclei. Meanwhile, *Ascl1⁻/Th⁺* neurons populated the Arc and zona incerta (ZI) by E18.5.

c, *Isl1* and *Meis2* transcriptional trends of differentiation for trajectories in *Th⁺* groups (clusters 1–9). Amplitudes are shown in log₁₀ scale. Line shading corresponds to mean ± s.e.m. **d**, Genetic lineage tracing using *Isl1-cre::Ai14* mice. **e**, In situ hybridization for *Gad1* and *Th* revealed anti-parallel expressional load for these genes as a factor of medial-to-lateral positioning. Scatter plots show the number of fluorescent puncta per cell (threshold >2). **f**, In situ hybridization for *Meis2*, *Th* and *Ddc* in the hypothalami of E18.5 and P2 mice. Scale bars, 120 μm (**a**, **f** (left)), 12 μm (**b**, **d**–**f** (right)).

a GAD67-GFP/ONECUT3/TH GAD65-eGFP/ONECUT3/TH

b *Meis2* *Pbx3* *Dlx1*

c ONECUT3/TH/Hoechst

d ONECUT3/TH/*Pmfbp1a* ONECUT2/*Th*/ONECUT3

e *Sst*/*Th*/Hoechst

f *Slc6a3*-Cre::tdTom/Biocytin/ONECUT3

Extended Data Fig. 10 | See next page for caption.

Article

Extended Data Fig. 10 | GABA origin of hypothalamic dopamine neurons.
a, Immunohistochemical analysis of TH and ONECUT3 protein expression in the hypothalamus of (BAC)GAD65–eGFP and GAD67–GFP mice at the developmental time-points indicated. Note a gradual GABA-to-dopamine transition as a factor of advancing age with ONECUT3 expression preceding that of TH. Dashed rectangles denote the positions of high-resolution insets. **b**, Expression patterns of regulon-forming TFs that directly drive *Th* transcription in the developing hypothalamus. *Meis2*, *Pbx3* and *Dlx1* were visualized on UMAP embedding for neuronal lineages. **c**, Histochemical localization of the migratory route of prospective PeVN dopamine neurons (cluster 9) through the coincident localization of TH and ONECUT3 during embryonic development. Dashed lines denote the ventricular surface. **d**, Localization of *Onecut2* and *Pmfbp1a* target genes within the *Onecut3* regulon to PeVN dopamine neurons by a combination of immunohistochemistry and in situ hybridization. **e**, *Sst* expression in PeVN dopamine neurons. **f**, Post hoc reconstruction of A14 *Onecut3*[+] dopamine neurons after patch-clamp recordings. Scale bars, 200 μm (**a**, overviews), 50 μm (**a**, insets), 20 μm (**f**), 12 μm (**c**–**e**).

nature research

Corresponding author(s):	HARKANY, Tibor
Last updated by author(s):	19.02.2020.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cells were sequenced using an Illumina HiSeq 3000/HiSeq 4000 System with corresponding Illumina commercial software (HCS v3.4.0, RRID:SCR_016386). For imaging, ZEN2010 (RRID:SCR_013672) was used as control software for a Zeiss LSM880 laser-scanning microscope. Patch-clamp recordings were carried out on an EPC-10 triple amplifier (HEKA, Germany) controlled by PatchMaster (v2.80, RRID:SCR_000034).
Data analysis	Cell Ranger (v2.2.0, RRID:SCR_017344), dropEst v0.8.3, Gephi (v0.9.2, RRID:SCR_004293), ZEN (Black, 2011, RRID:SCR_013672), ImageJ (1.52a, RRID:SCR_003070), Sigma13 (RRID:SCR_003210), GraphPad Prism (RRID:SCR_002798) were used, all in commercially-available configurations without the introduction of custom-made codes in our laboratory. We used existing R packages (RRID:SCR_001905) and Python modules (RRID:SCR_008394) as outlined in the on-line methods. CRAN packages (RRID:SCR_003005): Matrix (v1.2-17), MASS (v7.3-51.4), aricode (v0.1.2), clues (v0.6.1), fpc (v2.2-3), igraph (v1.2.4.1), Rtsne (v0.15, RRID:SCR_016342), irlba (v2.3.3), Seurat (v3.1.1, RRID:SCR_016341), ggalt (v0.4.0), heatmaply (v1.0.0), qvalue (v2.18.0, RRID:SCR_001073), reticulate (v1.13). Bioconductor (RRID:SCR_006442): MAST(v1.12.0, RRID:SCR_016340), org.Mm.eg.db (v3.7.0), GO.db (v3.7.0), topGO (v2.34.0), GOstats (v2.48.0), scater (v1.14.3, RRID:SCR_015954), MetaNeighbor (v1.6, RRID:SCR_016727), SC3 (v1.14.3, RRID:SCR_015953), schex (v1.0.0), scran (v1.14.3, RRID:SCR_016944). GitHub packages (RRID:SCR_002630): sctransform (v0.2.0, ChristophH/sctransform@8e48f49), dropestr (v0.7.9), pagoda2 (v0.1.0, hms-dbmi/pagoda2@dc550f0, RRID:SCR_017094), velocyto.R (v0.6 , velocyto-team/velocyto.R@666e1db), conos (v1.1.2, hms-dbmi/conos@b9ce17b), kBET (v0.99.6, theislab/kBET@4c9dafa), clustree (v0.4.1 Github (lazappi/clustree@39fd552), RRID:SCR_016293), liger (v0.4.2, MacoskoLab/liger@ae5f142), harmony (v1.0, immunogenomics/harmony@1a6d77a), Python modules (RRID:SCR_008394): scanpy (v1.4.5.dev180+g5d6f2769), scvelo (v0.1.25.dev10+gdf8ca57), pyscenic (v0.9.19, RRID:SCR_017247), arboreto (v0.1.5), leidenalg (v0.7.0), palantir (v0.2), umap-learn (v0.3.10), loompy (3.0.6, RRID:SCR_016666), anndata (v0.6.22.post2.dev105+g31659e6), scanorama (v1.5), bbknn (v1.3.6).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data generated in this study can be downloaded in raw and processed forms from the NCBI Gene Expression Omnibus, with accession number (GSE132730) provided in the text. Moreover, large data tables with tabulated entries are available from Figshare at DOI: 10.6084/m9.figshare.11867889.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A sample size of each measurement was determined by the practical limitations of the protocol utilised according to previously published criteria for single cell transcriptomics, in vitro and in vivo biological experimentation. While assuring reproducibility, re-sampling, parallel processing and repetitive iterations for animal experiments were limited as directed by available legal guidelines for animal ethics (see below). For qualitative experiments we utilized minimum n = 2/variable/time-point. For quantitative experiments we used at least n = 3 for each factor of comparison and time-point.
Data exclusions	The focus of this study was to investigate the development of ectoderm-derived cellular heterogeneity in a specific brain area. Therefore, when we identified non-ectoderm-derived populations of endothelial cells, pericytes, immune cells and mature oligodendrocytes, these were excluded from further analysis. Cell types unrelated to branches of neurogenesis (pars tuberalis, oligodendrocyte precursor cells) were also excluded from the final analysis, as described in detail the on-line methods because of the specific focus of this study. (Nevertheless, the full processed dataset with all cell types has been made available on GEO). In addition, exclusion of specific cells was made, in an unbiased manner, when data quality was low (see on-line methods and supplementary information for details). Otherwise, no data from bioinformatics, in vitro or in vivo experiments were excluded in a biased manner.
Replication	Multiple batches or timepoints served as replicates (showing consistency; for sequencing: data with/without pooling and resequencing (see below), from both sexes, in vitro: at least duplicate experiments, in vivo: multiple pregnancies for each experiment and processing of >2 fetuses/pregnancy for improved biological reliability (see below)). To minimize technical variability, we checked and combined data during library construction (n = 1 for P10, n = 2 for E15, E17, P0, P2, P10, P23, n = 3 for P2) including re-sequencing steps (n = 2 for E15, E17, P0, P2, P10, P23, n = 3 for E15). The hypothesis was tested independently on each sample using a published approach (velocito.R). Similarly, the assumption was tested on integrated data by two distinct approaches (directed PAGA using RNA velocity information and Palantir), which served as another type of replication to assure desire quality and integrity. For the analysis of embryonic hypothalami, the sample size was n = 7-8 embryos from 2 independent litters per time-point. Likewise, many adult samples were utilized (n = 3-4) without preference to sex to ensure reproducibility and to reduce the effect of individual variability.
Randomization	Randomization could not be performed due to a high risk of batch effect in case of sample separation.
Blinding	We did not use perturbation of conditions, instead we analysed different developmental time points which were processed in parallel. Thus, blinding was not possible due to obvious morphological differences between the samples. When processing ensuing datasets, we did not incorporate any information about the real age of the samples in any algorithm. Therefore, inferences on developmental trajectories were performed in an unsupervised manner as illustrated by Figure 1b.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary: rabbit anti-TH (1:500; Millipore Cat# AB152, RRID:AB_390204), #Lot 2593900, 3199177; sheep anti–TH (1:1000, Novus Cat# NB300-110, RRID:AB_10002491), #Lot ajo1217p; sheep anti-ONECUT2 (1:250; R and D Systems Cat# AF6294, RRID:AB_10640365), #Lot CDKS0116081; guinea pig anti-ONECUT3 (1:5,000); rabbit anti-VGLUT2 (1:800; a gift from M. Watanabe); goat anti-GFP (1:1,000; Abcam Cat# ab6662, RRID:AB_305635), #Lot GR311622-15, GR311622-7; chicken anti-GFP (1:500, Aves Labs Cat# GFP-1020, RRID:AB_10000240), #Lot GFP697986; rabbit anti-SOX2 (1:500, Abcam Cat# ab97959, RRID:AB_2341193), #Lot GR3244885-1; chicken anti-mCHERRY (1:1,000; EnCor Biotechnology Cat# CPCA-mCherry, RRID:AB_2572308), #Lot 7670-4; mouse anti-MASH1 (1:100, BD Biosciences Cat# 556604, RRID:AB_396479), #Clone: 24B72D11.1; guinea pig anti-GFAP (1:500, Synaptic Systems Cat# 173 004, RRID:AB_10641162), #Lot 2-15, 2-17; rabbit anti-phospho-Histone H3 (1:500; Cell Signaling Technology Cat# 3377S, RRID: AB_1549592), #Lot 7; chicken anti-NeuN (1:500, Millipore Cat# ABN91, RRID:AB_11205760), #Lot 3132967; mouse anti-FLAG-tag (1:1,000; Sigma-Aldrich Cat# F1804, RRID:AB_262044), #Lot SLBR7936V; mouse anti-HA-tag (1:600; Cell Signaling Technology Cat# 2367, RRID:AB_10691311) #Lot 1. Secondary: Cy3-AffiniPure Donkey Anti-Chicken (Jackson ImmunoResearch Labs Cat# 703-165-155, RRID:AB_2340363), Lot# 142225; Alexa Fluor 488-AffiniPure Donkey Anti-Guinea Pig (Jackson ImmunoResearch Labs Cat# 706-545-148, RRID:AB_2340472), Lot# 138058; Cy2-AffiniPure Donkey Anti-Rabbit (Jackson ImmunoResearch Labs Cat# 711-225-152, RRID:AB_2340612), Lot# 139999; Alexa Fluor 488-AffiniPure Donkey Anti-Goat (Jackson ImmunoResearch Labs Cat# 705-545-147, RRID:AB_2336933), Lot# 131669; Alexa Fluor 488 donkey anti-mouse (Jackson ImmunoResearch Labs Cat# 715-545-151, RRID:AB_2341099), Lot# 127820; Cy3-AffiniPure Donkey Anti-Mouse (Jackson ImmunoResearch Labs Cat# 715-165-150, RRID:AB_2340813), Lot# 116881; Cy3-AffiniPure Donkey Anti-Goat (Jackson ImmunoResearch Labs Cat# 705-165-147, RRID:AB_2307351), Lot# 134527; Cy3-AffiniPure Donkey Anti-Rabbit (Jackson ImmunoResearch Labs Cat# 711-165-152, RRID:AB_2307443), Lot# 141941; Cy3-AffiniPure Donkey Anti-Guinea Pig (Jackson ImmunoResearch Labs Cat# 706-165-148, RRID:AB_2340460), Lot# 134844; Alexa Fluor 647-AffiniPure Donkey Anti-Guinea Pig (Jackson ImmunoResearch Labs Cat# 706-605-148, RRID:AB_2340476), Lot# 135631; Alexa Fluor 647-AffiniPure Donkey Anti-Rabbit (Jackson ImmunoResearch Labs Cat# 711-605-152, RRID:AB_2492288), Lot# 127614.
Validation	For goat anti-GFP and chicken anti-mCherry: Species: Mouse; Applications: IHC; Kastriti, M. E. et al. Schwann cell precursors generate the majority of chromaffin cells in zuckerkandl organ and some sympathetic neurons in paraganglia. Front. Mol. Neurosci. 12, 6 (2019). For guinea pig anti-Onecut3: Species: Mouse; Applications: IHC; Espana, A. & Clotman, F. Onecut transcription factors are required for the second phase of development of the A13 dopaminergic nucleus in the mouse. J. Comp. Neurol. 520, 1424–1441 (2012). For sheep anti-Onecut2: Species: Mouse; Applications: IHC; Kabayiza, K. U. et al. The Onecut Transcription Factors Regulate Differentiation and Distribution of Dorsal Interneurons during Spinal Cord Development. Front. Mol. Neurosci. 10, 157 (2017). For rabbit anti-VGLUT2: Species: Mouse; Applications: IHC; Miyazaki, T., Fukaya, M., Shimizu, H. & Watanabe, M. Subtype switching of vesicular glutamate transporters at parallel fibre-Purkinje cell synapses in developing mouse cerebellum. Eur. J. Neurosci. 17, 2563–2572 (2003). For rabbit anti-TH: Species: Mouse; Applications: IHC; Romanov, R. A. et al. Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. Nat. Neurosci. 20, 176–188 (2017).

For sheep anti-TH:
Species: Mouse, Human; Applications: IHC;
La Manno, G. et al. Molecular diversity of midbrain development in mouse, human, and stem cells. Cell 167, 566-580.e19 (2016).

For chicken anti-GFP:
Species: Mouse; Applications: IHC;
Hoye, M. L. et al. MicroRNA profiling reveals marker of motor neuron disease in ALS models. J. Neurosci. 37, 5574–5586 (2017).

For rabbit anti-SOX2, mouse anti-MASH1:
Species: Mouse; Applications: IHC;
Hu, X.-L. et al. Persistent expression of VCAM1 in radial glial cells is required for the embryonic origin of postnatal neural stem cells. Neuron 95, 309-325.e6 (2017).

For guinea pig anti-GFAP:
Species: Mouse; Applications: IHC;
Hofmann, K. et al. Tanycytes and a differential fatty acid metabolism in the hypothalamus. Glia 65, 231–249 (2017).

For rabbit anti-phospho-Histone H3:
Species: Mouse; Applications: IHC;
Kim, S.-Y. et al. Cell autonomous phosphoinositide 3-kinase activation in oocytes disrupts normal ovarian function through promoting survival and overgrowth of ovarian follicles. Endocrinology 156, 1464–1476 (2015).
For chicken anti-NeuN:
Species: Mouse; Applications: IHC;
Davies, A. J. et al. Natural Killer Cells Degenerate Intact Sensory Afferents following Nerve Injury. Cell 176, 716-728.e18 (2019).

For mouse anti-FLAG-tag:
Species: Mouse; Applications: IHC;
Annunziata, I. et al. MYC competes with MiT/TFE in regulating lysosomal biogenesis and autophagy through an epigenetic rheostat. Nat. Commun. 10, 3623 (2019).

For mouse anti HA-tag:
Species: Mouse; Applications: IHC;
Yang, C. et al. Rewiring neuronal glycerolipid metabolism determines the extent of axon regeneration. Neuron 105, 276-292.e5 (2020).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Neuro2A (CLS Cat# 400394/p451_Neuro-2A, RRID:CVCL_0470)
Authentication	No specific authentication of cell line was performed.
Mycoplasma contamination	Microscopy analysis did not reveal any suspicion of mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	Lines from ICALC register were not used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mus musculus C57BL/6J (IMSR Cat# JAX:000664, RRID:IMSR_JAX:000664), Ai14 (IMSR Cat# JAX:007914, RRID:IMSR_JAX:007914), Ascl1-CreERT2 (IMSR Cat# JAX:012882, RRID:IMSR_JAX:012882), Th-GFP (IMSR Cat# RBRC03162, RRID:IMSR_RBRC03162), GAD65-GFP (MMRRC Cat# 011849-UCD, RRID:MMRRC_011849-UCD), GAD67-GFP (IMSR Cat# RBRC03674, RRID:IMSR_RBRC03674), Pomc-GFP (IMSR Cat# JAX:009593, RRID:IMSR_JAX:009593), Slc6a3-Ires-Cre (IMSR Cat# JAX:006660, RRID:IMSR_JAX:006660), Nfia-/- (MMRRC Cat# 010318-UNC, RRID:MMRRC_010318-UNC), Robo1-/- (IMSR Cat# APB:5320, RRID:IMSR_APB:5320), Slit1-/- (MMRRC Cat# 030404-MU, RRID:MMRRC_030404-MU), Slit2-/- (MMRRC Cat# 030405-MU, RRID:MMRRC_030405-MU), Isl1-cre (IMSR Cat# JAX:024242, RRID:IMSR_JAX:024242; IMSR Cat# JAX:007914, RRID:IMSR_JAX:007914), Oxtr-Venus (MGI:3838764) males and females. Ages used: E15.5, E17.5, E18.5, P0(E19.5), P2, P10, P23-P40.
Wild animals	The study did not involve the use of wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	Experiments on live animals conformed to the 2010/63/EU European Communities Council Directive and were approved by the Austrian Ministry of Science and Research (66.009/0145-WF/II/3b/2014, and 66.009/0277-WF/V/3b/2017). Particular effort was directed towards minimizing the number of animals used and their suffering during experiments.

Note that full information on the approval of the study protocol must also be provided in the manuscript.