Self-organization of human dorsal-ventral forebrain structures by light induced SHH

Riccardo De Santis, Fred Etoc, Edwin A. Rosado-Olivieri & Ali H. Brivanlou

Organizing centers secrete morphogens that specify the emergence of germ layers and the establishment of the body's axes during embryogenesis. While traditional experimental embryology tools have been instrumental in dissecting the molecular aspects of organizers in model systems, they are impractical in human in-vitro model systems to dissect the relationships between signaling and fate along embryonic coordinates. To systematically study human embryonic organizer centers, we devised a collection of optogenetic ePiggyBac vectors to express a photoactivatable Cre-loxP recombinase, that allows the systematic induction of organizer structures by shining blue-light on human embryonic stem cells (hESCs). We used a light stimulus to geometrically confine SHH expression in neuralizing hESCs. This led to the self-organization of mediolateral neural patterns. scRNA-seq analysis established that these structures represent the dorsal-ventral forebrain, at the end of the first month of development. Here, we show that morphogen light-stimulation is a scalable tool that induces self-organizing centers.
During embryogenesis, the central nervous system (CNS) is derived from an epithelial sheet of cells, the embryonic neural plate, that is induced in a polarized non-cell autonomous manner by a small group of cells called the Spemann organizer. Neural induction activity of the organizer occurs by a default mechanism that is exerted through the secretion of soluble inhibitors that block both branches (SMAD1/5/8 and SMAD2/3) of the TGFb signaling pathway. Dual SMADs inhibition directly converts pluripotent embryonic stem cells to the anterior neural tissue of the dorso-anterior forebrain. Downstream of these primary neural inducing signals, highly localized and dynamic organizing centers provide multiple morphogen sources that pattern the CNS. The embryonic neural tissue undergoes antero-posterior (A-P) and medio-lateral (M-L) patterning during neural plate stages. As the neural tube closes, cells in the lateral part of the plate become dorsal (roof plate) and those in the midline become ventral (floor plate). Thus, M-L patterns are converted to dorso-ventral polarity (D-V).

The interplay between two signaling pathways, BMP4 and SHH, both acting as morphogens guide the establishment of the M-L and D-V polarity of the embryonic neural tissue. BMP4, is expressed in the lateral edge of the neural plate, and subsequently in the dorsal neural tube, while SHH, is expressed in the ventral midline of the neural plate, and subsequently in the floor plate of the neural tube. Classical experimental embryology approaches such as ectopic presentation of SHH ligand by grafting coated beads, embryonic explants, or morphogen-secreting cells in mouse, chick or human stem cells demonstrated that SHH activity is sufficient to induce ventral neural fates via its transcriptional effector Gli3.

While these approaches have been instrumental in shaping our current understanding, they also suffer from technical shortcomings that have hindered a precise mapping of fate acquisition as a function of signaling dynamics in the context of early human development. For example, grafting experiments provide little control over the extent of the inductive field’s spatial limits, control of throughput, and non-specific effects due to wound healing. The development of tools that systematically control these parameters will lead to better dissection of morphogens patterning and will constitute a major step forward in experimental embryology. Optogenetic tools have been recently used to control gene expression with spatiotemporal resolution, taking advantage of different strategies. This carries the potential of creating exogeneous embryonic organizer centers in model tissues for quantitative studying of embryonic induction and for creating in vitro self-organized structures that present the axial organization, that is a key landmark of embryonic development. Light modulation of signaling pathways provides flexibility and high spatial resolution over the morphogenetic stimulus.

Here, we have devised a collection of optogenetic ePiggyBac vectors to conditionally express a photosactivatable Cre-loxP recombinase for creating spatially restricted organizing centers that break symmetry in self-organizing hESCs. This collection is an hESCs optimized blue-light inducible split-CRE system based on the Magnets-split-CRE. This system allows precise spatiotemporal control over the expression of a morphogen under a blue-light input. This experimental setup provides a highly quantitative and simplified method based on blue-light stimulation that can be used to induce organizing centers in vitro cultures of hESCs. To establish proof of feasibility, we tested our tool for its ability to break symmetry in hESC-derived neural tissue to light-induce M-L polarity by inducing stripes of SHH expression as observed in the midline of the neural plate in embryo. Light induced and polarized expression of SHH during neural induction in absence of exogenous WNT modulation, led to the self-organization of a 2D in vitro human dorsal-ventral forebrain structure, that include a ventral telencephalic-hypothalamic primordia. Polarization of morphogens using light provides a non-invasive approach to decipher the earliest events that underly symmetry breaking in the embryonic nervous system in stages of human development otherwise inaccessible for scrutiny.

### Results

#### Engineering a collection of optogenetic ePiggyBac vectors.

In order to provide light modulation of gene expression to human developmental studies, we re-engineered our original transposon ePiggyBac vector to conditionally express a light-inducible Cre-recombinase enzyme that takes advantage of the Magnets dimerization system (Magnet-CRE) (Fig. 1A). This allows for a stable integration of a blue-light dependent CRE enzyme in the genome. To avoid culturing cells in the dark, minimize leakage and to gain better control of light-sensitivity, we controlled the light-CRE enzyme using a Dox-inducible promoter and a second T2A peptide to improve the separation of its components (Fig. 1A, left panel). We paired this vector with a receiver ePiggyBac that carries two sequential ORFs (Red and Green modules) to be regulated by loxP recombination in a mutually exclusive manner (Fig. 1A, right panel). Both vectors were stably transfected into our female XX hESC line, RUES2 (NIHhESC-09-0013) and stimulated with DOX and blue-light through a light-blocking photomask (Fig. 1B, Supplementary Fig. 1A–C). Dox- and Blue-light induced hESCs showed robust expression of the Green module in patterns imposed by the shape of the photomask (Fig. 1B, C, Supplementary Fig. 1B, C). The Green module expressing cells showed a highly reproducible pattern of expression that tightly correlate with the photomask that slightly increase over time, probably due to cell proliferation (Supplementary Fig. 1D). The efficiency of light conversion by single-cell fluorescence measurement shows that 24 h of blue-light stimulation converts 78.3% of the total cells, while controls kept in the dark, or in absence of Dox, show less than 1% of Green positive cells (Fig. 1D). The activation of the green module depends on the duration of blue-light stimulation. Pulsed blue-light for 600 cycles (24 h) is the most efficient treatment without inducing cell death, as measured by Casp3 activation (Fig. 1D, Supplementary Fig. 1E). The blue-light-dependent induction of the Green module was also validated by measuring RNA levels by qRT-PCR in whole-illuminated samples (Fig. 1E). LoxP genomic recombination upon Dox and light treatment is shown by ampiclon Sanger sequencing of the selected genomic region (Fig. 1F). The light-induced pattern of gene expression is consistently and stably maintained over six days in the absence of continuous light stimulation (Supplementary Fig. 1F). Light induced gene expression modulation was not confined to a single hESC line as another of our hESC line (RUES1, genetic background, male XY, NIHhESC-09-0012) responded in the same manner (Supplementary Fig. 1G). Collectively, these experiments demonstrate that pairing a light inducible CRE enzyme with a stable and drug inducible ePiggyBac vector allows for a rapid and efficient spatiotemporal control of exogenous gene expression in hESCs.

**Light induced dorso-ventral hESC-derived neural tissue.** In order to study the M-L and D-V aspects of neural patterning in human models, we tested the ability of light stimulation to generate a localized SHH organizing center. We devised a LoxP inducible Green module to be co-expressed with SHH at the mRNA level, while producing two separate proteins using a T2A peptide. This setup can be regulated by blue-light stimuli through LoxP recombination (Fig. 2A). hESCs (RUES2) were differentiated using dual SMADs inhibition (SB431542 and
LDN193189) to induce an anterior forebrain fate. The application of DOX at day 0 confers light sensitivity for the first two days of differentiation (Fig. 2A). Neural induction using dual SMADs inhibition has been shown to generate progenitors representative of the embryonic neuroepithelium32–34. At this stage, a set of defined markers can be used to decipher patterning at various distances from the SHH source: PAX6 marks the most dorsal population, FOXA2 the floor plate and NKX2.1 the ventral neural progenitors in the basal telencephalon and hypothalamic primordia at E12.5 (Supplementary Fig. 2A)16,32.

Blue light was shone on day 1 during neural induction through a 1 mm rectangular mask for 24 h to produce SHH and a green fluorescent protein (NG-T2A-SHH) or a control fluorescent protein (NG-CNTRL) in a spatially restricted domain. Light
stimulation induced the expression of the Green module in both NG-CNTRL and NG-T2A-SHH lines (Fig. 2B). At day 2, FOXA2 positive cells were specifically induced by the NG-T2A-SHH but not in the NG-CNTRL line (Fig. 2B, C, Supplementary Fig. 2B). FOXA2 positive cells were detected in cells secreting SHH as well as in cells just next to the SHH secreting domain, providing functional evidence of autocrine as well as paracrine SHH activity (Fig. 2B, C, Supplementary Fig. 2B, C). Examination of the NG-T2A-SHH light-induced cells after 7 and 14 days display FOXA2, NKX2-1 and PAX6 cells that are organized in discrete domains while NG-CNTRL cells acquire PAX6 default neural fate, in absence of ventral cell types (Fig. 2D–F; Supplementary Fig. 3A–C). The expression NG-T2A-SHH is stably maintained in absence of light over the course of the differentiation (Supplementary Fig. 3D, E). SHH light induction unveiled M-L self-organization of the neural populations under the influence of an organizing center. Interestingly, at day 7, a population of cells co-expressing FOXA-2 and NKX2-1 was detected within and near by the light-induced organizer (Fig. 2D–F, Supplementary Fig. 3A). These FOXA2+/NKX2-1+ double-positive cells have been suggested to be human specific, as have not been detected in the mouse, while they are present in the ventral forebrain in human fetal samples at PCW5.516. PAX6, NKX2-1 and FOXA2 domains gradually segregate over time inside and outside the SHH induced domain, with PAX6 cells localized the farthest from the SHH source (Fig. 2D–F). Single-cell quantification shows that at day 7, the organizer induces population of cells double positive for NKX2-1+/FOX2A2+, both cell autonomously and non-cell autonomously (Fig. 2E, upper panel). A fraction of light converted cells, express high levels of FOXA2 but not NKX2-1 (Fig. 2D, E, upper panel). At day 14, the ventral cellular populations induced by SHH differentiate into a NKX2-1+/FOX2A2+ population, that is located both laterally and inside the light induced SHH organizer (Fig. 2D, E, lower panel). Also, there is a population of cells NG-T2A-SHH+/FOX2A2+/NKX2-1+ (Fig. 2E, F, Supplementary Fig. 4A). NKX2-1 domain juxtaposed to the SHH organizer is induced independently from the size of the SHH domain (Supplementary Fig. 4B). The RNA expression of the NG, the exogenous SHH and its downstream target GLI1 correlate with the expression of the NG-T2A-SHH module, validating our co-expression strategy (Supplementary Fig. 4C). Therefore, our analysis revealed a proximal distal pattern of ventral cell fates from the SHH organizer during neural induction. Spatiotemporal control of SHH induces ventral neural fates that are organized in a 2D space in vitro, resembling M-L and D-V neural populations (Supplementary Fig. 2D). Moreover, it validates the functionality of our optogenetic tool for its ability to induce and self-organize discrete fates in hESCs with a simple blue light stimulation.

### Molecular characterization of light-induced neural fates.

In order to precisely and unbiasedly identify the cell types present in our light-induced, self-organizing neural tissue, we characterized their transcriptome using scRNA-seq from two independent biological replica (12097 and 6207 cells) (Fig. 3A, B, Supplementary Fig. 5A–C). Cells were differentiated and stimulated with blue light as previously described in Fig. 2A and harvested at day 14 for single-cell RNA-sequencing analysis. Differentially expressed genes based on leiden clustering, RNA-velocity trajectories and cell-cycle predictions were used to classify 14 distinct cell identities: (i) FOXA2+/ARX Floor Plate; (ii) NKX-2.1+/RAX+/SIX6+ Ventral tuberal hypothalamic progenitors; (iii) NKX-2.1+/FOXG1+ Ventral telencephalic progenitors, (iv) NKX-2.1+/NHLH2+/OPT+ Ventral hypothalamic neurons; (v) TFAP2A+/KRT19+ Superficial ectoderm; (vi) SOX10+/PLP1+ Neural Crest, (vii) PAX6+/EMX2+/OTX2+ Dorsal forebrain progenitors; (viii) IRX3+/OLIG3+ Dorsal thalamic progenitors; (ix) TBR1+/LHX1+ Dorsal Neurons 1; (x) HES6+/DLL3+ Dorsal Neurons 2; dorsal and ventral proliferating progenitors, (xi) Dorsal, (xii) Dorsal thalamic, (xiii) Ventral and (xiv) an unidentified population (UnId) (Fig. 3B, Supplementary Fig. 5D, Source Data file). This analysis revealed the presence of multiple cell types that demarcate different domains along the embryonic D-V and A-P axes in agreement with what was previously shown by asymmetric SHH stimulation in 3D organoids.18 No endoderm, mesoderm or extra-embryonic markers were detected, locating cells in the ectodermal compartment.

To correlate the timing of our self-organizing in vitro tissues with in-vivo events, we integrated our dataset with a scRNA-seq collection of mouse brain samples at different time points, E8.5, E10, E12, E12.5, E13, and E1533, scRNA-seq transcriptomics of the human light-induced cells, grouped as neural precursors (NPCs), floor plate, superficial ectoderm, neurons and UnId cells, integrate with the in-vivo mouse brain developmental atlas (Supplementary Fig. 6A). The human NPCs display high correlation with the Radial glia population at E10, while in vitro derived neurons display high correlation with the mouse neuronal category at E12.5-E15 (Supplementary Fig. 6B, C), suggesting a temporal match with human development at PCW4-5 (https://embryology.med.unsw.edu.au/embryology/index.php/Carnegie_Stage_Comparison).

The expression of GLI3, GAS1 and PTCH was used to identify SHH receiving cells. In agreement with literature, we identified a...
GLI3-GAS1\textsubscript{low}/PTCH1\textsubscript{high} population as SHH stimulated while GLI3-GAS1\textsubscript{high}/PTCH1\textsubscript{low} cells as unstimulated (Fig. 3C, Supplementary Fig. 7A). We also validated the specific induction of SHH signaling in the NG-T2A-SHH line compared to a NG-CNTRL line by testing GLI1 and GLI3 expression at day 14 using qRT-PCR (Fig. 3D). Among the SHH induced populations, scRNA-seq analysis confirmed the presence of a FOXA2\textsuperscript{+}/ARX\textsuperscript{+}, floor plate population and revealed the identity of four distinct NKX2-1\textsuperscript{+} populations (Fig. 3B, C, E). The first is representative of the tuberal hypothalamic neural progenitors positive for NKX2-1, SIX6, SIX3, RAX (Fig. 3C, Fig. 4A–C, Supplementary Fig. 7C), the second is the NKX2-1\textsuperscript{+} FOXG1\textsuperscript{+} population representative of the ventral telencephalic population (Fig. 3A, B), the third is a ventral population of proliferating...
progenitors and the fourth is a small population of ventral neurons that we classified as hypothalamic neurons positive for NXX2-1, OTP, NNLH2 (Fig. 3B, C, Source Data file). Among the SHH unstimulated cells, we identified dorsal populations that consist of forebrain progenitors (PAX6+/EMX2+/OTX2+), thalamic progenitors (IRX3+/OLIG3+), two neuronal populations (TBR1+/LHX1+ and HES6+/DLI3+), and non-neural ectoderm derivatives such as superficial ectoderm and neural crest (Fig. 3B, Supplementary Fig. 5C, Source Data file). The non-neural ectoderm population derive mostly from the plate edge independently from the light organizer (Supplementary Fig. 7B).

Immunostaining for specific markers, SIX6, RAX, NXX2-2 and FOXG1, revealed the spatial segregation of telencephalic and hypothalamic territory (Fig. 4B, C, Supplementary Fig. 7C, D, E). We further showed that light-modulation of SHH not only self-organizes telencephalic and hypothalamic progenitors, but also neurons, since hypothalamic OTP+ neurons are preferentially located proximal to the light-induced organizer (Fig. 4D). While the differentiation of hypothalamic OTP cells has been previously observed in traditional cell culture or 3D organoids, confinement of a SHH source in 2D instructs the self-organization of a ventral telencephalic-hypothalamic structures that are spatially organized in monolayer.

Finally, the hypothalamic marker genes used in this study were in-vivo validated for their specific expression in the human fetal hypothalamus at PCW10 (Supplementary Fig. 8A) in order to capture the gene expression modules that are shared between our in vitro dataset and the fetal hypothalamus, we computed the gene regulatory networks (regulons) in each dataset using pySCENIC. Among the 427 active regulons identified in the human fetal hypothalamus dataset, the 72.5% (310 regulons) are shared with our in vitro dataset (Supplementary Fig. 8B, Source Data file). Based on RNA velocity analysis, we identified in the light-induced scRNA-seq dataset a ventral differentiation trajectory that starts from the ventral proliferating progenitors and ends at the ventral hypothalamic neurons (Fig. 3B). Performing gene ontology analysis of genes that are differentially expressed along this trajectory, we identified waves of gene expression linked to cell-cycle regulation, neural progenitor expansion and neuronal maturation (Supplementary Fig. 8C). We explored whether some important gene expression patterns recently described in the context of human fetal hypothalamic development were recapitulated in our model, TTYH1, HMGA2 and MYBL2 show the same progenitor-neuronal trend observed in the human fetal hypothalamus at PCW10 (Supplementary Fig. 8D).

Collectively, our experiments demonstrate the ability to generate organizing centers by a simple blue light-induction of a morphogen, which specify a morphogenetic source that patterns discrete cell types along a proximal distal axis in space, and lineage trajectory in time.

**Discussion**

We devised a collection of optogenetic ePiggyBac vectors to conditionally express a photoactivatable Cre-loxP recombinase, that allows the control of morphogen expression using a simple blue light stimulation in hESCs. In this study, we used this tool to induce a localized domain of SHH expression to generate a ventral organizer within a neutralizing tissue. This establishes proof of feasibility for a technological concept, which we predict to apply to a variety of gene expression modules, including signaling pathways, transcription factors and cell-cycle players. This experimental embryology tool has advantages over the classical techniques as it provides higher throughput and resolution. This system allowed us to control the asymmetric induction of the morphogen SHH during hESCs differentiation, which generated multiple locally organized cell types. It represents the first generation of light-induced self-organization of cell fates and patterns in hESCs. Future optimization will improve several aspects. For example, it is foreseeable to add a pulsatile temporal regulation to this setup by imposing another layer of regulation rather than a step induction as it is in this study. Traditional methods for inducing local perturbation of signaling pathways, such as beads, provide little control over the extent of the inductive fields and are incompatible with high-throughput setups. Our in vitro model system displays high reproducibility and spatial control among many independent wells of differentiating human pluripotent stem cells. Since our light-inducible experimental setup is unidirectionally activated, the forebrain D-V structures obtained in this study, are likely generated by the combination of direct SHH signaling and proliferation/expansion of the induced progenitors. We envision that manipulating multiple signals using different light wavelength provides the promise of more sophisticated inductive interaction studies. For example, pairing ligands and inhibitors to precisely decipher the cellular and molecular aspects of Turing based reaction diffusion events, which have been shown to have an instructing role in the establishment of embryonic patterns.

Light control of gene expression allows geometric confinement of the stimulus and its response in inducing the ventral organizer SHH. We have previously shown that physical geometrical confinement of hESCs is sufficient to induce self-organization in response to BMP4 in the context germ layers (gastruloids). Self-organization has also been demonstrated to occur in the context of a single embryonic germ layer, ectoderm and Dual SMADs...
inhibition leads to self-organization of telencephalic neural progenitor around a lumen generating rosette (cerebroids)\textsuperscript{34,47}. When Dual SMADs inhibition is followed by BMP4 presentation, confined hESCs colonies self-organize to generate the four derivatives of the ectodermal layers: neural, neural crest, sensory placode and epidermis that organize in radially symmetrical patterns\textsuperscript{34}. In the context of the ectoderm, this type of self-organization reflects the events that occur in the dorsal anterior part of the developing CNS. As for spatial geometrical confinement that provided the key element for gastruloid, cerebroid and ...
neuruloids self-organization, here we confined the geometry of a morphogenetic organizer by imposing a chemical edge with light stimulation. Complementing these results, our tool allowed the self-organization of the ventral aspect of the developing nervous system. It is tempting to speculate that self-organization imposed by geometrical confinement when combined with spatial manipulation of a morphogen will lead to more sophisticated and complex aspects of self-organization. Confinement of geometry and organizer structures promise to generate highly quantitative models that will more faithfully represent early human development in vitro.

We and others have shown that it is possible to deconvolve several aspects of embryonic self-organization by modeling developmental events in vitro. Our study uses a light inducible tool to generate 2D models of D-V forebrain development, composed of neural telencephalic and hypothalamic populations.

Interestingly, the hypothalamus is among the most conserved structures of the brain in vertebrates. In the adult, it is composed of several spatially distinct nuclei that control a wide range of functions, from body homeostasis to behavior and it is connected with the endocrine system through the pituitary gland. To the best of our knowledge, a comprehensive and systematic induction of discrete hypothalamic nuclei in the human context has not been achieved yet. Spatiotemporal induction of specific signaling pathways in the context of hypothalamic self-organization promise to recapitulate, in vitro, several aspects of the highly spatially organized hypothalamic development. Self-organization follows self-organization, beginning from an entire embryo (gastruloid) to a single embryonic germ layer such as the ectoderm (cerebroids and neuruloids), and ultimately the formation of discrete organs and cell types (dorsal and ventral brain). By this logic we expect that the development of discrete hypothalamic nuclei will also

Fig. 4 Spatial self-organization of telencephalic and hypothalamic fates upon light-induced SHH. A UMAP plot showing a selected population of cell NKX2-1+. Expression of markers that distinguish hypothalamic and telencephalic populations (NKX2-1, SIX6 and FOXG1). B Immunostaining shows the spatial segregation of ventral population arising from a light-induced SHH source at day 14. (Green NG, Magenta NKX2.1, Cyan SIX6, Red FOXG1, Gray DAPI). Scale bar = 100 μm. C Cumulative fluorescence intensity analysis (line profile) over the x-axis. x-axis displays the linear distance in μm. y-axis shows the cumulative fluorescent intensity profile in arbitrary units for each channel. Line profile shows the average (line) and SD (area) for each channel. The line profile is color-coded as the immunofluorescent channels, NG-Green, NKX2-1-Magenta, SIX6-Cyan, FOXG1-Red at day 14 quantification (n = 3 biologically independent samples). D Immunostaining shows OTP positive cells induced in proximity of the NG-T2A-SHH organizer but not in the NG-CNTRL. Light induced SHH drive the self-organization of both neural progenitors and neurons (Green NG, Magenta OTP, Gray DAPI), Scale bar = 100 μm.
follow the rules of self-organization and symmetry breaking. We speculate that the generation and patterning of hypothalamic territories are possible as long as the appropriate physical (confinement) and chemical (morphogen) cues are provided. Self-organizing models of the human hypothalamus in vitro will shed light on the basic principles that govern the development of a fundamental brain controlling center.

Methods

Human pluripotent stem cell culture. hESCs lines used in this study are part of the NIH Human Embryonic Stem Cell Registry (RUES1-NIHhESC-09-0012) and (RUES2-NIHhESC-09-0013). Human pluripotent cells were maintained in feeder-freeculture conditions on Geltran-coated dishes. Cells were fed with MEF conditioned medium (CM) supplemented with bFGF 20 ng/ml. Daily, cells were routinely passaged every 3–4 days using Gentle Cell Dissociation Reagent (STEMCELL-Technologies).

Neural induction for pluripotent stem cell. hESCs were passaged as single cells using Accutase (STEMCELL-Technologies) and seeded in CM media supplemented with bFGF and Y27632 at 100,000 cells/cm2 density. The differentiation started 1–2 days later when cells reached 100% confluency. Neural induction media is composed of: 50% Neurobasal and 50% DMEM-F12, supplemented with 2 N2 0.5%, B27 0.5%, Glutamax 0.5X and Essential amino acid 0.5X and Insulin (2.5 μg/ml) all from LifeTechnologies. Small molecule SB431542 (10 μM) and LDN193189 (100 nM) were supplemented to warm media.

DOX (1 μg/ml) induction was used at day “0” concomitantly with dual-SMAD inhibition, and then washed out at day “2”. Neural induction media supplied with dual-SMAD inhibitors is maintained throughout differentiation until day “14”.

piggyBac vectors generation. One of the two piggyBac DOX inducible and Puromycin selectable vectors was digested with HindIII-NotI and ligated with the Magnets-CRE expression cassette (pcDNA-MagnetsCRE plasmid is a kind gift from Dr. Yazzawa). The other that is Blasticidin selectable, and carries a constitutive promoter (CAG), was modified to express a Loxp regulated switch of the two ORFs. A first “Red module” that express a dsRed protein, is removed and upon Loxp recombination allows the expression of a second “Green module”, which express a Neon-Green protein fused with a T2A, allowing co-expression with other proteins (Fig. 1B). The Loxp-dsRed-Loxp was cloned using BamHI-BglII restriction site vector from the plV-CMV-Loxp-DsRed-Loxp-CAG, a gift from Jaco van Rheenen (Addgene plasmid # 65762). PCR-amplified SV40 polyadenylation site was introduced using NheI restriction enzyme after the dsRed ORF. The Green module consisting of the nuclear localized Neon-Green protein was synthesized using IDT genebuckeba and inserted using the restriction enzymes AgeI-NotI. The SHH coding sequence was PCR amplified and cloned using BsmBI-NotI restriction enzymes from pOEM1 pmcV3Shh-pP-PVsgvsg, a gift from Elly Tanaka (Addgene plasmid # 111156). Plasmids used in this study have been deposited on ADDEGENE.

Stable integration of piggyBac vectors in the genome of hESCs. hESCs were nucleofected using an Amaxa Nucleofector II (Lonza) according to manufacturer’s instructions for hESC nucleofection. A mixture (1 μg + 0.5 μg) of: pEBPURO-TP-PA-CRE, pEB-RBD-CAG-RLOXP-N2A or pEB-RBD-CAG-RLOXP-N2A-SHH, and the piggyBac transposase were nucleofected in 10 hESCs and plated in Geltrix (Life technologies) coated plates using CM media supplemented FGF2 (20 ng/ml) and ROCK-inhibitor (Y-27632, 10 μM). Cells were then selected using CM media supplemented with blasticidin 5 μg/ml and puromycin 1 μg/ml (all from Life Technologies). FGF2 (20 ng/mL) and ROCK-inhibitor (Y-27632, 10 μM) were added to hESCs and plated in Geltrex (Life technologies) coated plates using CM media supplemented FGF2 (20 ng/mL) and ROCK-inhibitor (Y-27632, 10 μM) and ROCK-inhibitor (Y-27632, 10 μM). hESCs were passaged as single cells using Accutase (STEMCELL-Technologies) and seeded in CM media supplemented with bFGF and Y27632 at 100,000 cells/cm2 density. The differentiation started 1–2 days later when cells reached 100% confluency. Neural induction media is composed of: 50% Neurobasal and 50% DMEM-F12, supplemented with 2 N2 0.5%, B27 0.5%, Glutamax 0.5X and Essential amino acid 0.5X and Insulin (2.5 μg/ml) all from LifeTechnologies. Small molecule SB431542 (10 μM) and LDN193189 (100 nM) were supplemented to warm media.

DOX (1 μg/ml) induction was used at day “0” concomitantly with dual-SMAD inhibition, and then washed out at day “2”. Neural induction media supplied with dual-SMAD inhibitors is maintained throughout differentiation until day “14”.

RNA velocity and pseudotime analysis. Kalilto Bustools1 was used to align and quantify layered spliced and unspliced gene counts. The two scRNA-seq replicates were aligned independently, using kb-python=0.24.4 to capture spliced and unspliced transcripts, applying the -lanmann function with the GRCm38 human genome. The Kalilto unfilted count matrix was imported in Scany and further processed using scvelo (v0.22.3)22, using UMAP coordinates and clusters annotation from previously annotated Scany object. Count matrices carrying the spliced/unspliced layer annotation were filtered and normalized for minimum shared counts=30 and the top 2000 high variable genes. The first 30 principal components and 30 neighbors were used to calculate ‘Mi’ and ‘Mi’ moments of spliced/unspliced abundance. We used the steady-state model3 to determine velocities using the “scvelo” velocity function with default parameters.

We computed pseudotime analysis on a subset of cell as classified as NKX2-1 hypothalamic cells, consisting of: NKX2-1+/RAX+/SIX6 Ventral tibular hypothalamic progenitors, NKX2-1+/NHLH2+/OTP+ Ventral hypothalamic neurons and NKX2-1+/TOP2A+ ventral proliferating progenitors. Using the python wrapper of Slingshot, ”scprep.run.Slingshot” (scprep 1.0.0, https://pypi.org/project/scprep/), we calculated the pseudotime order of cells using full covariance matrix and default parameters. The normalized and log transformed count matrix for gene trends investigation has been denosed using MAGIC (3.0.0)24 (default parameters, t=3). Trajectory scatterplots are displayed as a function of the pseudotime, x-axis, and the gene expression values, y-axis.

Gene patterns have been identified using the scaled high variable genes and the Agglomerative Clustering algorithm (from scikit-learn=0.22), https://scikit- learn.org/leiden (cluster Numb/stable) along the reconstructed pseudotime. Selected gene pattern clusters have been tested for enrichment of specific categories using GSEA (10.0.5)-enrich (GO biological process 2018) (https://pypi.org/project/gsea/)25.

Analysis and integration of publicly available scRNA-seq datasets. scRNA-seq samples from the mouse brain developmental atlas25 were downloaded as a Loom file and according to the cell type annotation, we used E8.5, E12, E15, E18.5. Extracted count matrices were further normalized, log transformed and filtered for mitochondria content less than 15% and genes to be expressed in at least 5 cells. Gene annotations were converted to their human orthologous using ”scaply- queries:biomart_ annotations”. UMAP coordinates were calculated using (n.neighbors = 50, PCA components= 30). Human and mouse datasets were integrated using Ingest “scit.ingest” function, using the common expressed high variable genes between the two datasets, using default settings. Correlation analysis of the integrated dataset was performed using the “Class” annotation from the mouse annotation and the human light-induced cells, grouped as neural precursors (NPCs), floor plate, superficial ectoderm, neurons and Udl cells. Correlation analysis is calculated using the scannpy correlation matrix function with default pearson parameters. Dendrogram is calculated using default settings and 30 PCA.

cRNA-seq count matrix of the human fetal hypothalamus42 was processed using the following settings: genes expressed in more than 5 cells and cells with less than 20% of mitochondrial genes were used for further analysis. Count matrices have been normalized and log transformed. UMAP coordinates and clusters were obtained using (n.neighbors = 10, PCA components= 30) and leiden clustering (resolution=1).

Regulations identification. pyScenic (0.11.2) is used to identify gene regulatory networks in scRNA-seq datasets, using default parameters44. In brief, count matrices were preprocessed using the following parameters: minimum number of genes = 1, minimum cell fraction = 0.2 and number of nearest neighbors = 20 and at least 0.2 correlation in the correlation matrix is used to infer co-expression modules using the human transcription factors ranked database (https://resources.aertslab.org/cistarget/databases/). The
Arboreto algorithm (grnboost2) is used to calculate adjacency matrices. Regulons are computed using the adjacencies of each dataset independently, followed by the pruning module for targets with cis regulatory footprints (ReciTarget). Commonly detected regulons between the fetal hypothalamic dataset and our scRNA-seq dataset have been tested for enrichment of gene ontology, using the “Tissues protein expression from the human proteome map” from GSEA/Py-enrich.

**Imaging.** Confocal images were acquired on a Zeiss Inverted LSM 780 laser scanning confocal microscope with a x20 dry objective.

**Imaging analysis.** Images were preprocessed to display the Maximum Intensity Projection (MIP) of at least 4 z-stacks using ZEN black software. Fiji, V2 (Version 2.1.0/1.53g) was used for image display and formatting. MIP images were subsequently imported into Python using the Ciffile package (2019.7.2). Line profile analysis consists of a single channel exported and the intensity values for each pixel over the y-axis were summed and displayed as a curve. Relative internal normalization is performed when indicated. Line profile in a 0–1 range was obtained normalizing values for each channel according to their maximum and minimum value. Single-cell intensity quantification was performed by identification of individual cells using an Otu2 binarized DAPI image followed by distance transform and Watershed algorithm (scikit-image, 0.16.2) to separate overlapping nuclei. The pixel intensity of each object is plotted.

Cell death is quantified by calculating the area of CASP3 positive staining normalized on the total area. CASP3 positive signal is obtained upon Otu2 thresholding of CASP3 staining using the python package (scikit-image 0.16.2, threshold_otsu). Plotting was performed after data processing using nupmy (1.19.5) and pandas (1.1.5) libraries, using matplolib (3.2.2) and seaborn (0.11.2) plotting libraries.

**qRT-PCR.** RNA from individual wells was extracted using the Qiagen RNeasy Plus Mini Kit. 1ug of RNA was retrotranscribed using Transcriptor First Strand cDNA Synthesis Kit (Roche 0489703001) and Real-time quantification was performed using SYBR Green Master Mix (Roche 0489703001). HPRT or ATP5O endogenous control is used for internal normalization and results are expressed as fold change over a reference sample.

**Sanger sequencing.** PCR-amplified amplicons (Hotstart Q5 polymerase, NEB) were sanger sequenced using Genewiz sanger sequence service under standard developments.brain-map.org)56.

**Brain atlas showing ISH staining. Image credit: Allen Institute (https://genouscontrol is used for internal normalization and results are expressed as fold using SYBR Green Master Mix (Roche 04887352001). HIPRT or ATP5O endogenous control is used for internal normalization and results are expressed as fold change over a reference sample.

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

**Data availability.** The scRNA-seq data generated in this study have been deposited in the GEO database under accession code GSE163505. The mouse brain developmental atlas used in this study from La manno et al., was downloaded from http://mousebrain.org/downloads.html. The human fetal hypothalamic dataset used in this study from Zhou et al. is available on GEO under the GSE118487. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding author upon reasonable request. Source data are provided with this paper.

**Materials availability.** Plasmids are available on ADDGENE and cell lines upon request to the corresponding author.

**References.**

1. Spemann, H. & Mangold, H. Induction of embryonic primordia by implantation of organizers from a different species. 1923. Int. J. Dev. Biol. 45, 13–38 (2001).

2. Harland, R. & Gerhart, J. Formation and function of spemann’s organizer. Cell Dev. Biol. 33, 611–667 (1997).

3. Sasai, Y. et al. Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. Cell 79, 779–790 (1994).

4. Hemmati-Brivanlou, A. & Melton, D. A. Inhibition of activin receptor signaling promotes neuralization in Xenopus. Cell 77, 273–281 (1994).

5. Ozair, M. Z., Kintner, C. & Brivanlou, A. H. Neural induction and early patterning in vertebrates. Wiley Interdiscip. Dev. Biol. 2, 479–498 (2013).

6. Santis, R. D. & Brivanlou, A. H. From neural induction in frogs to human brains on chips. eLS 1, 508–515 (2020).

7. Chambers, S. M. et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat. Biotechnol. 27, 275–280 (2009).

8. Arnold, S. J. & Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat. Rev. Mol. Cell Biol. 10, 91–103 (2009).

9. Bordiis, U. & Pierani, A. Patterning the cerebral cortex: traveling with morphogens. Curr. Opin. Genet. Dev. 20, 408–415 (2010).

10. Kiecker, C. & Lumsden, A. The role of organizers in patterning the nervous system. Neuroscience 35, 347–367 (2012).

11. Cox, W. G. & Hemmati-Brivanlou, A. Caudalization of neural fate by tissue recombination and bFGF. Dev. Camb. Biol. Engl. 121, 4349–4358 (1995).

12. Nikolopoulou, E., Galea, G. L., Rolo, A., Greene, N. D. E. & Copp, A. J. Neural tube closure: cellular, molecular and biomechanical mechanisms. Development 144, 552–566 (2017).

13. Lupo, G., Harris, W. A. & Lewis, K. E. Mechanisms of ventral patterning in the vertebrate nervous system. Nat. Rev. Neurosci. 7, 103–114 (2006).

14. Kuschel, S., Rüther, U. & Theil, T. A disrupted balance between Bmp/Wnt and Fgf signaling underlies the ventralization of the Gli3 mutant telencephalon. Dev. Biol. 260, 484–495 (2003).

15. Marklund, M. et al. Retinoic acid signalling specifies intermediate character in the developing telencephalon. Development 131, 4323–4332 (2004).

16. Marooof, A. M. et al. Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell 12, 559–572 (2013).

17. Placzek, M. & Briscoe, J. Sonic hedgehog in vertebrate neural tube development. Int J. Dev. Biol. 62, 225–234 (2018).

18. Cedernaquist, G. Y. et al. Specification of positional identity in forebrain organoids. Nat. Biotechnol. 37, 436–444 (2019).

19. Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nat. Methods 7, 973–975 (2010).

20. Kawano, F., Okazaki, R., Yazawa, M. & Sato, M. A photoactivatable Cre recombinase system for optogenetic genome engineering. Nat. Chem. Biol. 12, 1059–1064 (2016).

21. Nihongaki, Y. et al. CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. Nat. Methods 14, 963–966 (2017).

22. Quejada, J. R. et al. Optimized light-inducible transcription in mammalian cells using Flavin Kelch-repeat F-box/GIANTEA and CRY2/ClB1. Nucleic Acids Res. 45, gkx804 (2017).

23. Shao, J. et al. Synthetic far-red light-mediated CRISPR-Cas9 device for inducing functional neuronal differentiation. Proc. Natl Acad. Sci. 115, 201802448 (2018).

24. Morikawa, K. et al. Photocatalytic Cre recombinase 3.0 for in vivo mouse applications. Nat. Commun. 11, 2141 (2020).

25. Rogers, K. W. & Müller, P. Optogenetic approaches to investigate spatiotemporal signaling during development. Curr. Top. Dev. Biol. 137, 37–77 (2019).

26. Kim, N. et al. Spatiotemporal control of fibroblast growth factor receptor signals by blue light. Chem. Biol. 21, 903–912 (2014).

27. Sako, K. et al. Optogenetic control of nodal signaling reveals a temporal pattern of nodal signaling regulating cell fate specification during gastrulation. Cell Rep. 16, 866–877 (2014).

28. Johnson, H. E. et al. The spatiotemporal limits of developmental Erk signaling. Dev. Cell 40, 185–192 (2017).

29. Repina, N. A. et al. Engineered illumination devices for optogenetic control of cellular signaling dynamics. Cell Rep. 31, 107737 (2020).
Acknowledgements
The authors would like to thank Zeeshan M. Ozair for discussion, criticisms, advise, and comments on the manuscript. Manon Valet for computational suggestions and Sandra Moreu for supporting plasmids generation. We also would like to thank Jean-Marx Santel and Peter Ingrassia for revising the manuscript. We are grateful to Brivanlou and Siggia lab members for inputs and criticism. We also would like to thank the Genomics Resource Center and the Bio-Imaging Resource Center for technical support. We wish to thank Kavli-foundation for supporting this research with the Kavli-NSH-Pilot grant to R.D.S. and A.H.B. R.D.S. was supported by EMBO-LTF:254:2019.

Author contributions
R.D.S. conceived the project, performed experiments and analyzed the data. F.E. contributed to experimental discussions, imaging analysis and to establish the LED-light setup. E.A.R. contributed to scRNA-seq analysis. A.H.B. supervised the project and secured funding. R.D.S. and A.H.B. wrote the manuscript with input from all authors.

Competing interests
A.H.B. is the co-founder of RUMI Scientific, RUMI Viro and OvaNova. A.H.B., F.E., and E.A.R. are shareholders of RUMI Scientific and RUMI Viro. R.D.S. has no competing interests to disclose.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26881-w.

Correspondence and requests for materials should be addressed to Ali H. Brivanlou.

Peer review information Nature Communications thanks Corinne Houart, Moritoshi Sato and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

30. Lacoste, A., Berenshteyn, F. & Brivanlou, A. H. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. Cell Stem Cell 5, 332–342 (2009).
31. Rosa, A., Papaioannou, M. D., Kryspiak, J. E. & Brivanlou, A. H. miR-373 is regulated by TGFβ signaling and promotes mesendoderm differentiation in human Embryonic Stem Cells. Dev. Biol. 391, 81–88 (2014).
32. Fasano, C. A., Chambers, S. M., Lee, G., Tomishima, M. J. & Studer, L. Engineering derivatives of the floor plate from human embryonic stem cells. Cell Stem Cell 6, 336–347 (2010).
33. Ozair, M. Z. et al. hPSC modeling reveals that fate selection of cortical deep projection neurons occurs in the subplate. Cell Stem Cell 23, 60–73.e6 (2018).
34. Haremaki, T. et al. Self-organizing neuroloids model developmental aspects of Huntington’s disease in the ectodermal compartment. Nat. Biotechnol. 37, 1219–1230 (2019).
35. Manno, G. L. et al. Molecular architecture of the developing mouse brain. Nature 596, 92–96 (2021).
36. Shimogori, T. et al. A genomic atlas of mouse hypothalamic development. Nat. Neurosci. 13, 767–775 (2010).
37. Martinez-Ferre, A. & Martinez, S. Molecular regionalization of the diencephalon. Front. Neurosci.-switz. 6, 73 (2012).
38. Morales-Delgado, N. et al. Regionalized differentiation of CRH, TRH, and GHRH peptidergic neurons in the mouse hypothalamus. Brain Struct. Funct. 219, 1083–1111 (2014).
39. Merkle, F. T. et al. Generation of neuropeptidergic hypothalamic neurons from human pluripotent stem cells. Development 142, 633–643 (2015).
40. Qian, X. et al. Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. Cell 165, 1238–1254 (2016).
41. Kasai, T. et al. Hypothalamic contribution to pituitary functions is recapitated in vitro using 3D-cultured human iPS cells. Cell Rep. 30, 18–24.e5 (2020).
42. Zhou, X. et al. Cellular and molecular properties of neural progenitors in the developing mammalian hypothalamus. Nat. Commun. 11, 4063 (2020).
43. Sande, R., Vde A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. 15, 2247–2276 (2020).
44. Yu, Y. et al. Engineering a far-red light–activated split-Cas9 system for remote-controlled genome editing of internal organs and tumors. Sci. Adv. 6, eabb1777 (2020).
45. Yen, S. et al. CreLite: an optogenetically controlled CreloxP system using red light. Dev. Dynam. 249, 1394–1403 (2020).
46. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitate early embryonic spatial patterning in human embryonic stem cells. Nat. Methods 11, 847–854 (2014).
47. Knight, G. T. et al. Engineering induction of singular neural rosette emergence within hPSC-derived tissues. Elife 7, e37549 (2018).
48. James, D., Noggle, S. A., Swigut, T. & Brivanlou, A. H. Contribution of human embryonic stem cells to mouse blastocysts. Dev. Biol. 295, 90–102 (2006).
49. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19, 15 (2018).
50. 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).
51. Melsted, P. et al. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat Biotechnol. 1–6. https://doi.org/10.1038/s41587-021-00870-2 (2021).
52. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. 38, 1408–1414 (2020).
53. Manno, G. L. et al. RNA velocity of single cells. Nature 560, 494–498 (2018).
54. Dijk, D. van et al. Recovering gene interactions from single-cell data using data diffusion. Cell 174, 716–729.e27 (2018).
55. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. Nucleic Acids Res. 14, 128 (2013).
56. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. Nature 445, 168–176 (2007).

The authors would like to thank Zeeshan M. Ozair for discussion, criticisms, advise, and comments on the manuscript. Manon Valet for computational suggestions and Sandra Moreu for supporting plasmids generation. We also would like to thank Jean-Marx Santel and Peter Ingrassia for revising the manuscript. We are grateful to Brivanlou and Siggia lab members for inputs and criticism. We also would like to thank the Genomics Resource Center and the Bio-Imaging Resource Center for technical support. We wish to thank Kavli-foundation for supporting this research with the Kavli-NSH-Pilot grant to R.D.S. and A.H.B. R.D.S. was supported by EMBO-LTF:254:2019.