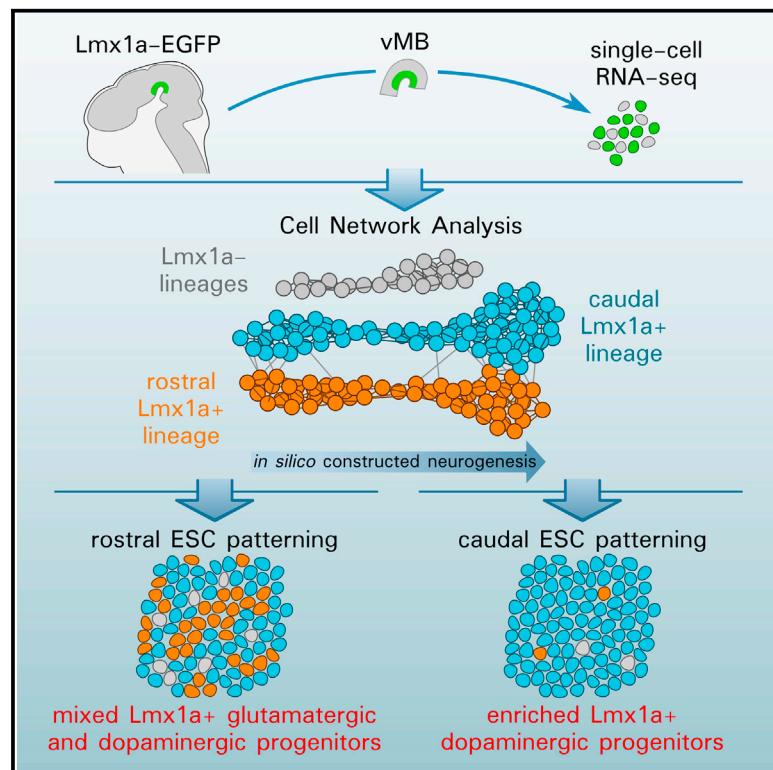


# Single-Cell Analysis Reveals a Close Relationship between Differentiating Dopamine and Subthalamic Nucleus Neuronal Lineages

## Graphical Abstract



## Authors

Nigel Kee, Nikolaos Volakakis,  
Agnete Kirkeby, ..., Rickard Sandberg,  
Malin Parmar, Thomas Perlmann

## Correspondence

nigel.kee@licr.ki.se (N.K.),  
thomas.perlmann@ki.se (T.P.)

## In Brief

Kee et al. use single-cell RNA-seq to reconstruct Lmx1a+ differentiation in silico, revealing an unexpectedly close relationship between mesDA and STN neuronal lineages during differentiation. Application of markers that distinguish the two can help optimize mesDA differentiation of hESCs with a view to improving therapeutic translation.

## Highlights

- Single-cell RNA-seq can reconstruct Lmx1a+ neuronal development
- The Lmx1a+ population contains distinct mesDA and STN lineages
- The STN and mesDA lineages have related, yet distinct, transcription factor profiles
- Application of distinct markers can help optimize mesDA differentiation of hESCs

## Accession Numbers

GSE87069

# Single-Cell Analysis Reveals a Close Relationship between Differentiating Dopamine and Subthalamic Nucleus Neuronal Lineages

Nigel Kee,<sup>1,2,\*</sup> Nikolaos Volakakis,<sup>1</sup> Agneta Kirkeby,<sup>3</sup> Lina Dahl,<sup>1</sup> Helena Storvall,<sup>1</sup> Sara Nolbrant,<sup>3</sup> Laura Lahti,<sup>1</sup> Åsa K. Björklund,<sup>4</sup> Linda Gillberg,<sup>1</sup> Eliza Joodmardi,<sup>1</sup> Rickard Sandberg,<sup>1,2</sup> Malin Parmar,<sup>3</sup> and Thomas Perlmann<sup>1,2,5,\*</sup>

<sup>1</sup>Ludwig Institute for Cancer Research, Box 240, 171 77 Stockholm, Sweden

<sup>2</sup>Department of Cell and Molecular Biology, Karolinska Institutet, 171 77 Stockholm, Sweden

<sup>3</sup>Department of Experimental Medical Science, Lund Stem Cell Centre, Lund University, 221 84 Lund, Sweden

<sup>4</sup>Department for Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, 751 24 Uppsala, Sweden

<sup>5</sup>Lead Contact

\*Correspondence: [nigel.kee@licr.ki.se](mailto:nigel.kee@licr.ki.se) (N.K.), [thomas.perlmann@ki.se](mailto:thomas.perlmann@ki.se) (T.P.)

<http://dx.doi.org/10.1016/j.stem.2016.10.003>

## SUMMARY

Stem cell engineering and grafting of mesencephalic dopamine (mesDA) neurons is a promising strategy for brain repair in Parkinson's disease (PD). Refinement of differentiation protocols to optimize this approach will require deeper understanding of mesDA neuron development. Here, we studied this process using transcriptome-wide single-cell RNA sequencing of mouse neural progenitors expressing the mesDA neuron determinant Lmx1a. This approach resolved the differentiation of mesDA and neighboring neuronal lineages and revealed a remarkably close relationship between developing mesDA and subthalamic nucleus (STN) neurons, while also highlighting a distinct transcription factor set that can distinguish between them. While previous hESC mesDA differentiation protocols have relied on markers that are shared between the two lineages, we found that application of these highlighted markers can help to refine current stem cell engineering protocols, increasing the proportion of appropriately patterned mesDA progenitors. Our results, therefore, have important implications for cell replacement therapy in PD.

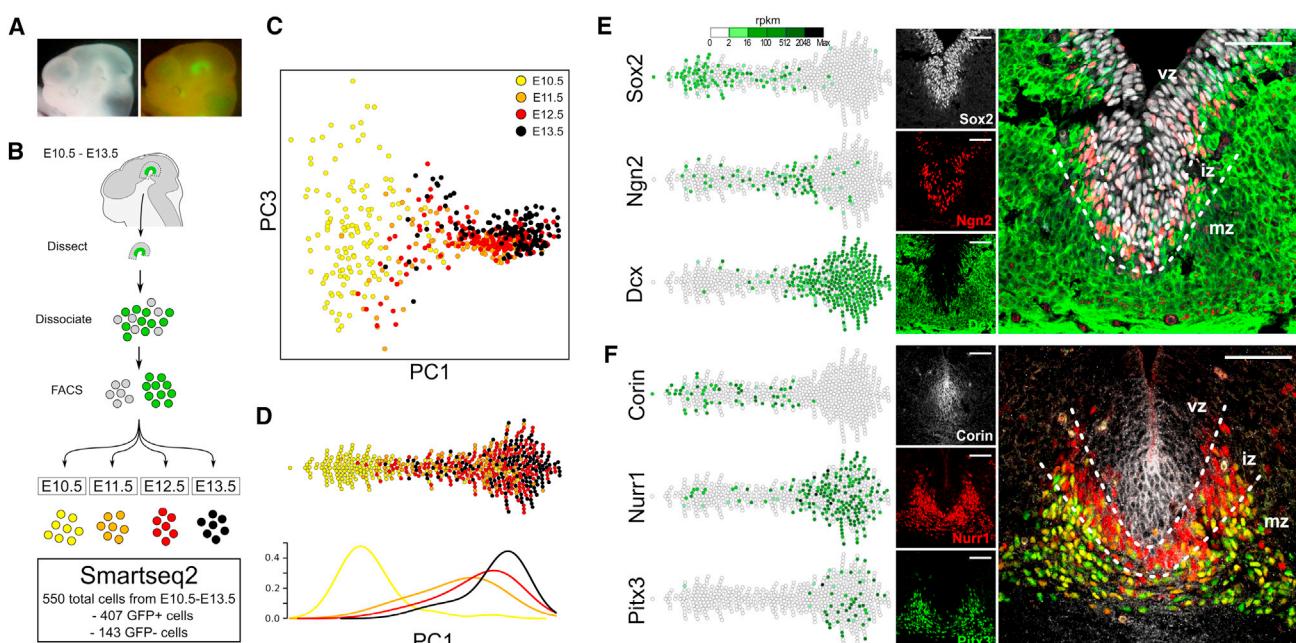
## INTRODUCTION

A promising treatment for neurodegenerative disorders is cell-replacement of clinically relevant neuron types engineered from stem cells (Steinbeck and Studer, 2015). Parkinson's disease (PD) is particularly attractive from this perspective, as demonstrated by clinical studies using fetal grafts containing mesencephalic dopamine (mesDA) neurons that degenerate in disease (Barker et al., 2015). Intense efforts have focused on engineering human pluripotent embryonic stem cells (hESCs) or induced pluripotent stem cells into transplantable and functional mesDA neurons (Doi et al., 2014; Jaeger et al., 2011; Kirkeby

et al., 2012; Kriks et al., 2011). For the development of such protocols, understanding the mechanisms underlying the normal development of mesDA neurons has been critical, since the principle strategy has been to culture stem cells under conditions that mimic the signaling environment of the developing ventral midbrain and caudal diencephalon, where mesDA neurons are normally generated. However, major challenges remain including cellular heterogeneity, batch-to-batch variability in differentiation outcome, and a need to identify improved predictive markers that can indicate successful graft outcome prior to transplantation. Critical for moving forward toward robust protocols is therefore to achieve a more comprehensive understanding of the neuronal lineages developing in the vicinity of mesDA neurons. Such knowledge would elucidate the likely contaminating neuron types, thereby providing gene profiles that can inform the development of efficacious, clinically viable protocols (Kirkeby et al., 2016, in this issue of *Cell Stem Cell*).

mesDA neurons are derived from neural progenitor cells located in the ventral mesencephalon and caudal diencephalon (Arenas et al., 2015; Marín et al., 2005). At early developmental stages, spatially organized signaling events induce the patterned expression of transcription factors including Lmx1a. Lmx1a has a restricted expression pattern in progenitor cells in the ventral mesencephalic and diencephalic (VMD) region and was for this reason identified as a selective mesDA neuron-specifying determinant (Andersson et al., 2006b). Together with the highly related Lmx1b, Lmx1a is indeed essential in the specification of mesDA neural progenitors as demonstrated by gene ablation studies in mice (Deng et al., 2011; Yan et al., 2011). Following mesDA neuronal lineage specification, progenitors exit the cell cycle and initiate both pan-neurogenic and mesDA neuron-specific gene expression programs that are dependent on additional transcription factors including Nurr1 and Pitx3 (Smidt et al., 2004; Zetterström et al., 1997). Thus, a mesDA neuron-specific transcription factor signature has been defined and shown to mediate essential functions in specification of neural stem cells and in the acquisition of mesDA neuron-specific traits.

Despite these advances, the precise orchestration of mesDA neurogenesis and how it relates to other types of neurons developing within the VMD region remains poorly understood and has



**Figure 1. Strong Axis of Differentiation Is Illustrated in Single-Cell Data**

- (A) Whole-mount micrographs of heterozygous E11.5 *Lmx1a*-EGFP embryos. EGFP fluorescence is shown to the right.
- (B) Schematic depicting isolation strategy for EGFP-positive and -negative cells from the VMD region at different embryonic stages. The green domain and cells indicate EGFP-positive cells. The cells in the lower image are colored according to embryonic age: E10.5 yellow; E11.5 orange; E12.5 red; and E13.5 black.
- (C) PCA plot of all single cells based on 3,006 differentially expressed genes.
- (D) All cells plotted according to their position along only PC1, arranged so as to be adjacent, but not overlapping. The cells are colored according to embryonic age as in (A). The lower image depicts the frequency distribution along PC1 for each embryonic age. The lines are colored according to embryonic age.
- (E) Expression levels of pan-neuronal markers *Sox2*, *Ngn2*, and *Dcx* plotted along PC1. The right hand images show triple co-immunostainings for *Sox2*, *Ngn2*, and *Dcx* in VMD tissue in a coronal section at E12.5.
- (F) Expression levels of mesDA neuronal lineage markers *Corin*, *Nurr1*, and *Pitx3* plotted along PC1. The right hand images show triple co-immunostainings for *Corin*, *Nurr1*, and *Pitx3* in the VMD region in a coronal section at E12.5 (the scale bar represents 100  $\mu$ m).

not been resolved at the global gene expression level. As a consequence, contaminating neuron types in stem cell engineering remain insufficiently characterized. Newly developed methods for analyzing mRNA expression in single cells (RNA sequencing [RNA-seq]) are therefore of major interest since they offer new opportunities to define cell relationships within highly complex and dynamic tissues including the adult or fetal brain (Darmanis et al., 2015; Jaitin et al., 2014; Klein et al., 2015; Macosko et al., 2015; Pollen et al., 2014; Shalek et al., 2014; Treutlein et al., 2014; Zeisel et al., 2015).

To study mesDA and neighboring neuronal lineages, we performed single-cell RNA-seq of cells from the developing mouse VMD region that either express or do not express the transcription factor *Lmx1a*. Computational analysis of these data uncovered an unexpected strong relationship between mesDA and STN neural lineage development. Further histological, genetic, and in vitro stem cell analyses confirmed this relationship in mouse and human embryos. Importantly, our analysis uncovered new markers that can distinguish between early mesDA and STN development, while classical dopaminergic markers previously used in stem cell engineering are shared between the two lineages. These data therefore have important implications for strategies aimed at robust generation of highly enriched mesDA neurons for transplantation in PD (Kirkeby et al., 2016).

## RESULTS

### Single-Cell RNA-Seq Reveals a Strong Axis of Differentiation

To resolve mesDA neuron development at the genome-wide gene expression level, we took advantage of a mouse line expressing EGFP from the *Lmx1a* locus (*Lmx1a*<sup>EGFP</sup> mice) (Deng et al., 2011). Proliferating mesDA neural progenitors exit the cell cycle roughly between embryonic (E) day 10 to E13.5 in the mouse. Fluorescent-activated cell sorting (FACS) was used to isolate both EGFP-positive and negative single cells (in total 550 cells) from dissected VMD tissue from *Lmx1a*<sup>EGFP</sup> heterozygous mice at 4 embryonic days from E10.5 to E13.5 (Figures 1A, 1B, and S1; *Supplemental Experimental Procedures*). Libraries for RNA-seq were generated using the Smart-seq2 method (Picelli et al., 2013), and sequenced to a depth of at least 400,000 uniquely mapped reads, displaying a mean Spearman correlation of 0.45 when considering all genes expressed above a reads per kilobase transcript and million mappable reads (RPKM) cut-off of one (Figure S1A). Additionally, 100-cell pool bulk RNA samples were prepared in parallel, with six replicates in four groups (EGFP negative and positive cells from E10.5 and E13.5) that clustered in agreement with *in silico* bulk samples constructed from corresponding single cells (Figure S1B). The 100-cell pools were used to generate an instructive gene set consisting of 3,006

differentially expressed genes that were used in further analysis (**Figure S1C**).

Considering this gene set, principal-component analysis (PCA) illustrated a strong axis of age along principal component one (PC1; **Figure 1C**). Thus, the density of all cells' ages along PC1 followed an expected timing, with E10.5 cells occupying negative ranges, and E13.5 cells occupying positive ranges (**Figure 1D**). EGFP positive and negative cells were dispersed along this component (**Figure S1D**).

Expression of established neural differentiation markers also conformed to PC1's age axis. Higher expression of progenitor markers including *Sox2* was seen toward the left, while expression of the post-mitotic neuronal markers including doublecortin (*Dcx*) was higher toward the right (**Figures 1E** and **S1E**). Pro-neural genes, including *Ngn2*, *Ascl1*, and *Neurod1*, and regulators of the Notch signaling pathway are crucial for the transition of proliferating progenitors into post-mitotic neurons and were expressed toward the center of PC1 (**Figures 1E** and **S1F**). This transition was further illustrated by the expression of additional pro-neural genes that, as predicted from previous studies, describe neurogenic cascades along PC1 (**Figure S1G**).

Known mesDA neuronal lineage markers also showed a predictable distribution along PC1 (**Figures 1F** and **S1H**). mesDA neuron progenitor markers including *Corin* were high toward the left, while *Nurr1* (*Nr4a2*) and *Pitx3*, two critical post-mitotic mesDA transcription factors, as well as mature mesDA neuron markers *Th*, *Slc6a3*, and *Slc18a2*, were high toward the right of PC1 (**Figures 1F** and **S1H**). Notably, subtle timing differences were also recapitulated along PC1. For example, *Ascl1* expression precedes the highest density of *Ngn2* expression (**Figure 1G**) (Andersson et al., 2006a; Kele et al., 2006), while *Nr4a2* is expressed slightly earlier than *Pitx3*, an order also seen in vivo (**Figure 1F**). Additionally, non-dopaminergic markers such as the *Gli* transcription factors, *Nkx6-1*, *Isl1*, and *Lhx5* were detected (**Figure S1I**), illustrating that cells were derived from both EGFP-positive and EGFP-negative lineages. Together, these results show that single-cell RNA-seq has a remarkable potential to reconstruct the VMD lineage development at the genome-wide expression level.

### Identification of Parallel Lmx1a-Expressing Neuronal Lineages

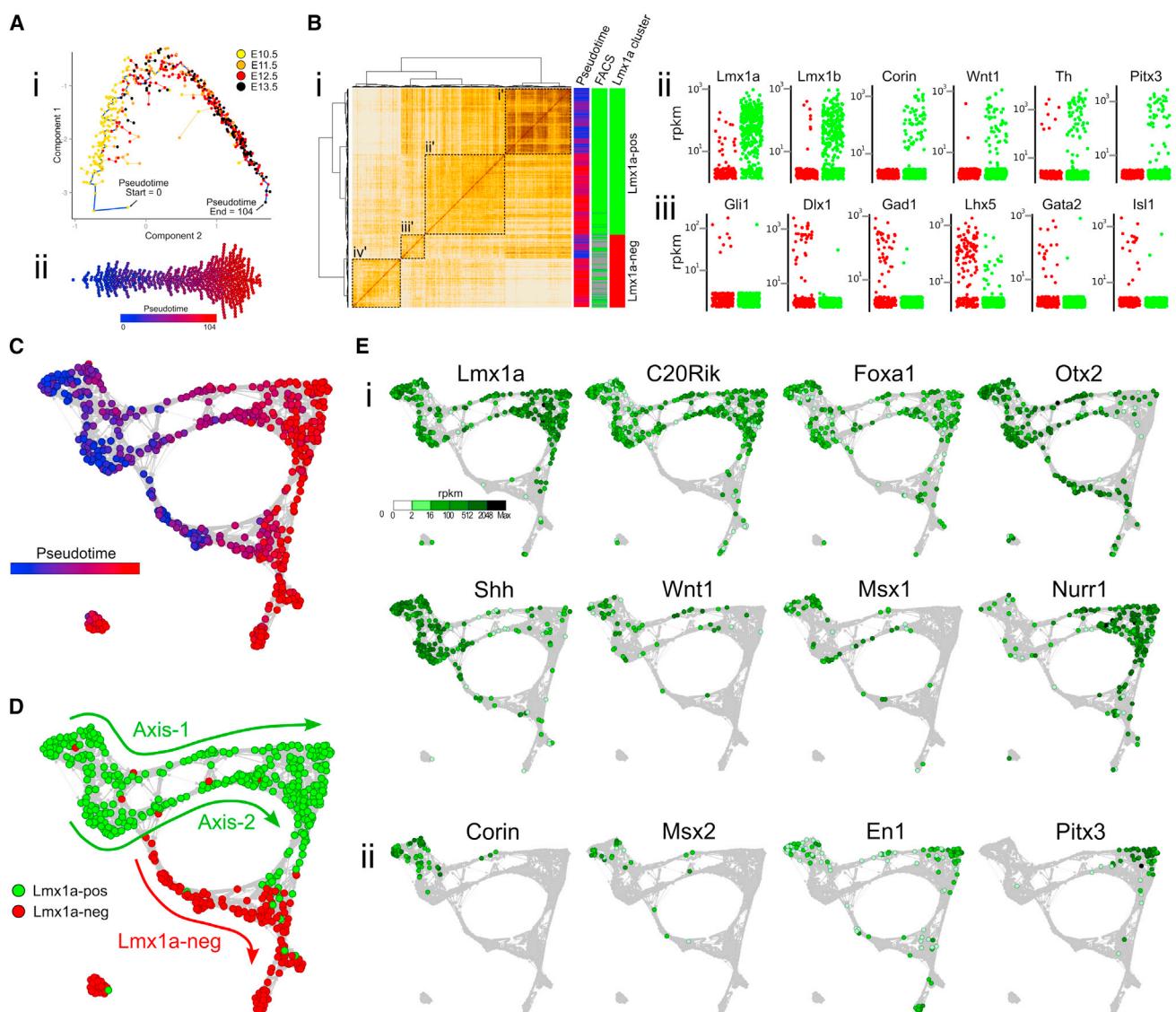
We next employed Monocle (Trapnell et al., 2014) to determine a temporal order in our single-cell data (**Experimental Procedures**). The Monocle-derived pseudotime differentiation axis matched very closely with PC1 (**Figure 2A**), showing the same distribution of cell age. Additionally, Monocle's unsupervised clustering generated multiple clusters of co-varying genes (**Figure S2A**), each enriched for gene ontology (GO) terms dominating successive phases of differentiation.

After establishing an axis of neural differentiation, explaining by far the largest source of variation in the data, we next sought to delineate more subtle differences between neuronal subtypes. In principle, Lmx1a-EGFP positive and negative lineages should be distinct and represented by cells found across all stages of pseudotime. A trend for such separation is recognized along PC3 (**Figure S1D ii**). However, Spearman correlation considering the original 3,006 genes, all expressed genes, or

the top 1,000 variable genes, could not recapitulate these expected FACS-based Lmx1a-EGFP positive and negative groups (**Figures S2B–S2D**). Thus, we curated an instructive lineage-defining gene set (termed Lineage Intersect gene set), depleted of genes involved in pan-neuronal differentiation, from commonly differentially expressed genes between EGFP positive and negative cell pools derived from both E10.5 and E13.5 (see **Figure S2E**). Spearman correlation clustering based on this gene set strongly divided into four main branches: two progenitor and two post-mitotic branches, as indicated by the cells' pseudotime values (**Figure 2B i**). Additionally, one progenitor and one post-mitotic branch each were classified into Lmx1a-EGFP-positive and -negative clusters that matched closely with the cells' original FACS-based classification. Only five cells considered negative by FACS clustered in the Lmx1a-positive group, while 45 cells considered FACS positive clustered in the Lmx1a-negative group. This matched our expectations, as a relaxed fluorescence gating was intentionally used while sorting to avoid missing cells expressing low levels of EGFP (**Figure S2F**). Scatterplots depicting gene expression across these two groups confirmed our clustering, with mesDA neuron-specific genes including *Lmx1a*, *Lmx1b*, *Corin*, *Wnt1*, *Th*, and *Pitx3* expressed at high levels in the Lmx1a-positive group (**Figure 2B ii**). Conversely, *Gli1*, *Dlx1*, *Gad1*, *Lhx5*, *Gata2*, and *Isl1* were expressed at high levels in the Lmx1a-negative group indicating, as expected, that non-dopaminergic neurons such as oculomotor neurons and GABAergic interneurons are included in this cluster (**Figure 2B iii**).

An important challenge is to uncover previously unknown subdivisions in Lmx1a-expressing cells. We considered two bioinformatic tools; t-distributed neighbor embedding (t-SNE) and graph-based clustering, recently proposed to effectively resolve high-dimensionally rich non-parametric data sets (Levine et al., 2015). We combined these tools in a novel workflow, using Euclidean distance after t-SNE projection to construct a force-directed graph of the nearest-neighbor network across all cells (**Figures 2C–2E** and **S2G–S2K**) (van der Maaten, 2014). Additionally, we employed a weighted PCA as input into the t-SNE projection (**Figures S3A–S3D**; **Experimental Procedures**). This resulted in a network map that strongly segregated early and late cells based on pseudotime, while also confirming our earlier classification of Lmx1a-positive and Lmx1a-negative cells. Reads were also mapped to an eGFP reference sequence and, as expected, expression was clearly restricted to Lmx1a-positive cells (**Figure S2G**). Interestingly, the network map also indicated the existence of two segregating Lmx1a-positive sublineages displaying parallel, unbroken pseudotime axes (**Figure 2D**), a pattern we were unable to replicate using pre-existing bioinformatic approaches (**Figures S3E–S3H**). These axes were also clear when using the alternative expression normalization approach (transcripts per million; **Figure S3I**). We refer to the three axes as Axis-1, Axis-2, and Lmx1a-neg.

The emergence of two Lmx1a-positive axes (Axis-1 and Axis-2) suggested the existence of at least two Lmx1a-expressing sublineages. To investigate this, expression values were plotted for known dopaminergic and non-dopaminergic markers. As expected, non-dopaminergic markers were highly expressed in Lmx1a-neg axis (**Figure S2H**). Several known dopaminergic markers such as *Lmx1a*, *Lmx1b*, *C130021I20Rik*,



**Figure 2. Separation of Two Distinct Lmx1a-Positive and One Lmx1a-Negative Pseudotime Axes**

(A) (i) Cells are colored as in Figure 1B. The monocle's pseudotime trajectory is depicted by a thick black line traversing through a sprawling minimum spanning tree. (ii) The cell density is plotted along PC1 and colored by pseudotime. Blue: early pseudotime and red: late pseudotime.

(B) (i-iii) Euclidean distance heatmap of cells based on Spearman correlation when considering the Lineage Intersect gene list. The clustering tree (left and top) divides into four main branches (i': early Lmx1a-positive cluster; ii': late Lmx1a-positive cluster; iii': early Lmx1a-negative cluster; and iv': late Lmx1a-negative cluster). The right hand bars indicate cells' pseudotime age (blue: early and red: late), FACS classification (green: GFP-positive and gray: GFP-negative), and assignment into either Lmx1a-positive (green) or Lmx1a-negative (red) lineages. (ii) The expression of mesDA markers in Lmx1a-negative (red) and Lmx1a-positive (green) lineages is shown. (iii) The expression of non-dopaminergic VMD markers in Lmx1a-negative (red) and Lmx1a-positive (green) lineages is shown.

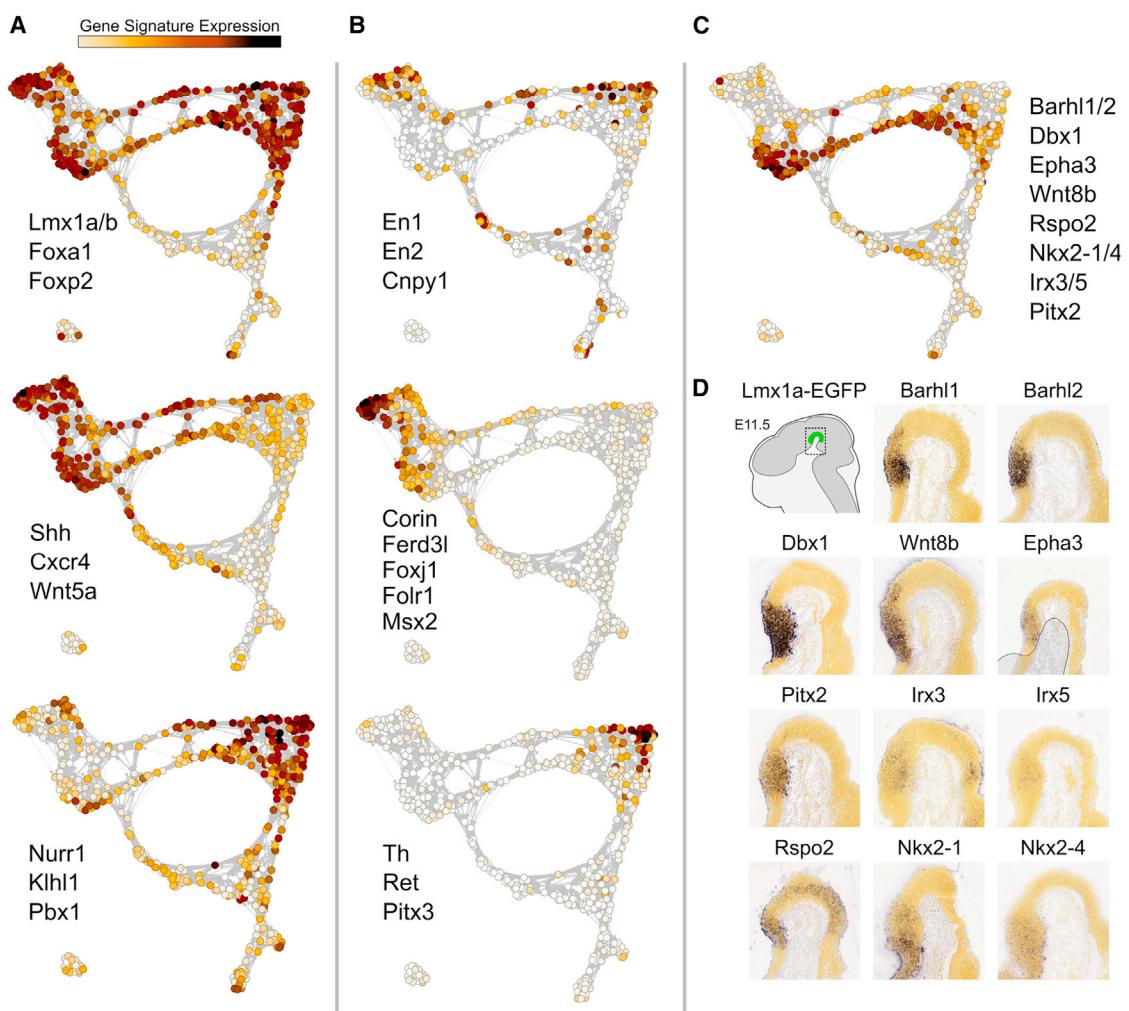
(C) t-SNE network plot shows a left-to-right distribution of cells roughly conforming to their established pseudotime order. Blue: early pseudotime and red: late pseudotime.

(D) Cells plotted as in (C). The assignment to either Lmx1a-positive (green) or Lmx1a-negative (red) lineages from Figure 2B i is indicated. The t-SNE plot separates into two parallel Lmx1a-positive axes (Axis-1 and Axis-2) and one Lmx1a-negative axis (Lmx1a-neg).

(E) t-SNE plots displaying expression of known mesDA markers.

*Foxa1*, *Foxa2*, *Otx2*, *Msx1*, and *Nurr1* (*Nr4a2*), in addition to the developmental signaling molecules *Shh*, *Wnt1*, *Wnt5a*, and *Cxcr4* were expressed at their expected early or late pseudotime points and were equally expressed in both Axis-1 and Axis-2 (Figures 2E i and S2I). Notably, many of these genes are essential markers commonly used to assess differentiation outcome when generating mesDA neurons from human pluripotent stem cells

(Doi et al., 2014; Jaeger et al., 2011; Kirkeby et al., 2012; Kriks et al., 2011). Surprisingly, however, multiple classic dopaminergic markers including *Corin*, *Msx2*, *En1*, and *Pitx3* were confined to only Axis-1 (Figure 2E ii). Additionally, at late pseudotime, Axis-2 contained lower levels of *Th*, *Slc6a3*, and *Slc18a2*, compared to the late pseudotime region of Axis-1 (Figure S2J). Thus, Axis-2 appeared to be non-dopaminergic.



**Figure 3. Gene Signatures Identified by WGCNA Distinguish Axis-1 from Axis-2**

(A) Relative expression levels (dark brown: high expression; yellow: low expression; and white: no expression) of gene signatures identified to be expressed in both Axis-1 and Axis-2. The representative genes are indicated for each signature. The top signature is strongly expressed from early through to late pseudotime; the middle signature is mostly active at early pseudotime, while the bottom signature is active only at late pseudotime.

(B) Relative expression levels of gene signatures identified to be expressed selectively in Axis-1. The representative genes are indicated for each signature and include many known mesDA markers. The top, middle, and bottom signatures are strongly expressed throughout, at early or at late pseudotime, respectively.

(C) Relative expression levels of the identified gene signature active in Axis-2.

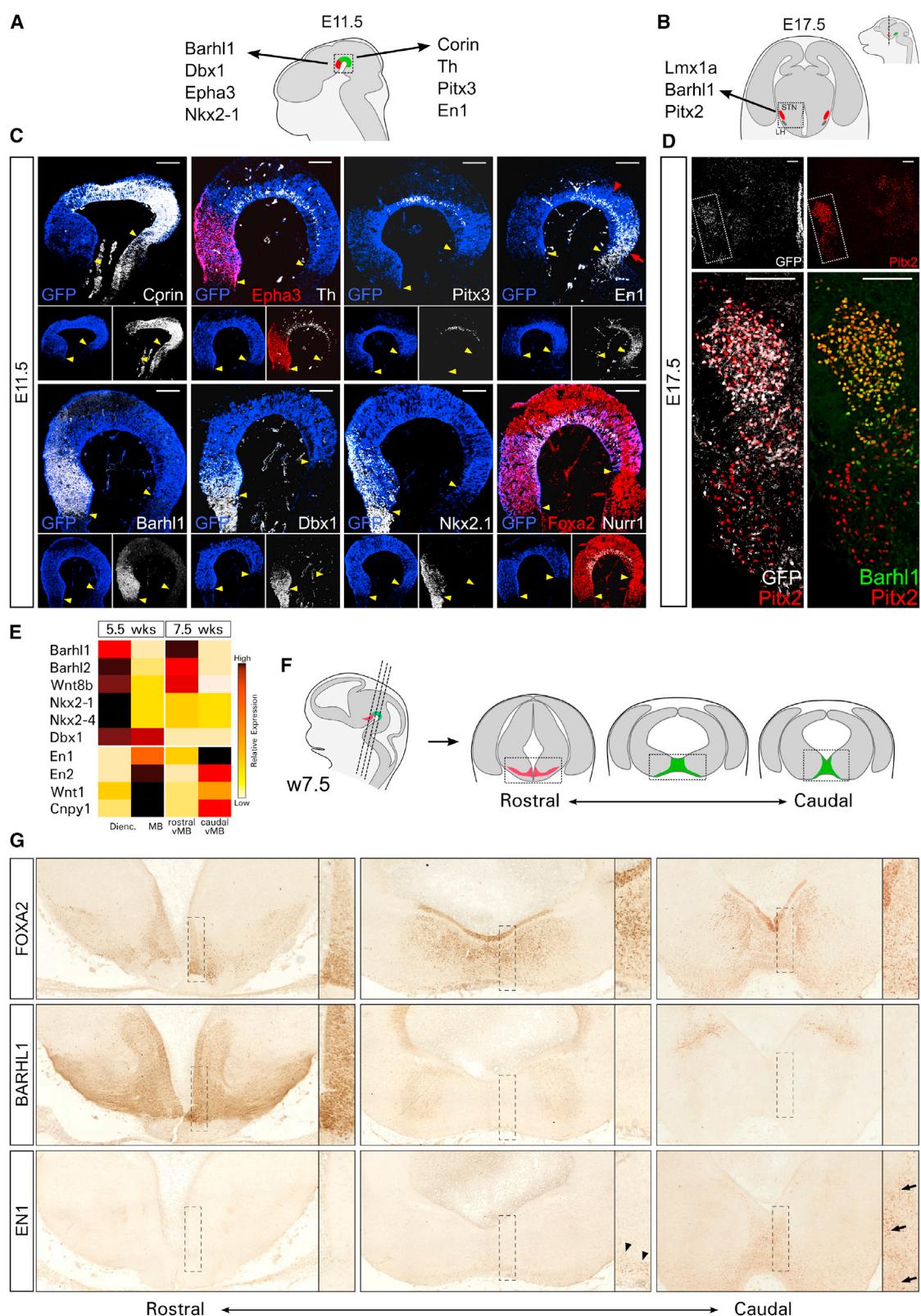
(D) In situ hybridization expression patterns from sagittal sections of E11.5 mouse embryos from the Allen Developing Mouse Brain Atlas of genes present in (C).

The expression was consistently confined to a domain in the developing diencephalon that appears rostral to developing mesDA neurons.

### Neuronal Identity of Axis-2

To investigate what genes may describe Axis-2, neighboring cells within the graph were assigned to one of 27 unsupervised clusters using Infomap graph-based community detection (Rosvall and Bergstrom, 2008) (Figure S4A i). To specifically isolate only lineage related genes by an unbiased approach, single-cell differential expression (SCDE) (Kharchenko et al., 2014) was performed between cluster pairs only of similar pseudotime age. Significant genes were then analyzed using weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008), a clustering approach that groups similarly behaving genes into gene signatures based on the topological overlap of their expression across all cells. Cells were then correlated against each gene signature and their score plotted on the

t-SNE network (Figure 3). Considering the Lmx1a-neg axis, multiple gene signatures described GABAergic and glutamatergic lineages known to develop within the VMD region and caudal hypothalamus (Figure S4B). When considering the two Lmx1a-positive axes, most gene signatures fell into two categories, those present in both Axis-1 and Axis-2 (Figure 3A) or those confined to only Axis-1 cells (Figure 3B). Strikingly, one gene signature including *Barhl1*, *Barhl2*, *Dbx1*, *EphA3*, *Wnt8b*, *Rspo2*, *Nkx2-1*, *Nkx2-4*, *Irx3*, *Irx5*, and *Pitx2* was confined to Axis-2, distinguishing this gene signature from the dopaminergic lineage (Figure 3C). Almost all genes in this signature were strongly expressed from the beginning of pseudotime, while many genes, including *Barhl1/2* and *Pitx2*, remained strongly expressed into late pseudotime (Figure S4C). Interrogation of Allen



**Figure 4. Immunohistochemistry of Identified Axis-2 Markers**

(A) Sagittal schematic of an E11.5 embryo, boxed section depicts image boundary in (C). The caudal Lmx1a domain expressing mesDA genes is indicated in green, while the rostral Lmx1a domain expressing STN genes is indicated in red.

(legend continued on next page)

Developing Mouse Brain Atlas in E11.5 mouse embryos demonstrated specific expression in a rostral Lmx1a-expressing domain (**Figure 3D**) (Thompson et al., 2014) (Allen Institute for Brain Science available from <http://www.developingmouse.brain-map.org>). Importantly, the expression of Pitx2, one of eight transcription factors present in this gene signature, has previously been mapped to the rostral/ventral diencephalon where it has been implicated in the development of the subthalamic nucleus (STN) (Asbreuk et al., 2002; Martin et al., 2004; Smidt et al., 2000), raising the possibility that this glutamatergic lineage may arise from rostral Lmx1a-expressing progenitor cells.

To clarify this *in silico*-derived rostro-caudal division of Axis-1 and Axis-2, we performed immunohistochemistry on sagittal sections from E11.5 Lmx1a<sup>EGFP</sup> heterozygote embryos (**Figures 4C** and **S5A**). EGFP extended rostrally from the midbrain into the developing diencephalon (**Figures 4C** and **S5A**, arrowheads). Corin, TH, Pitx3, and En1 were confined to caudal regions of the EGFP domain, thus defining the boundaries for mesDA development and genes confined to Axis-1. In contrast, markers confined to Axis-2 (Epha3, Barhl1, Dbx1, Nkx2-1, and Pitx2) were expressed immediately rostral to the mesDA neuron markers, extending further into the diencephalon. Consistent with their expression in both Axis-1 and Axis-2, Nurr1 (Nr4a2) and FoxA2 spanned both rostral and caudal Lmx1a domains (**Figure 4C**, bottom right). Nkx2-1 was only weakly expressed within the rostral Lmx1a-expressing domain, however, much stronger Nkx2-1 staining extended past the rostral boundary of Lmx1a expression (**Figure S5A**, top, yellow bracket), overlapping a domain also expressing Foxb1, a marker for developing mammillary bodies and hypothalamic neural populations (**Figure S5A**, bottom) (Zhao et al., 2008). Together, the expression of these markers defines two Lmx1a-expressing domains: a caudal domain undergoing mesDA neurogenesis, and an adjacent rostral domain expressing many classical mesDA progenitor and early differentiation markers, but, additionally, the unique gene signature of Axis-2. Interestingly, Lmx1a-expressing caudal mesDA progenitors can be further distinguished based on En1 expression, with little or no En1 protein more anteriorly (**Figure 4C**, top right, red arrowhead) and strong expression closer to the mid/hindbrain boundary (**Figure 4C**, top right, red arrow).

Analysis of Lmx1a<sup>EGFP</sup> heterozygous embryos at later stages corroborates the identity of the Axis-2 lineage. At E13.5, at a level immediately rostral to mesDA neurons, EGFP co-localizes with Barhl1, Pitx2, and Epha3 in rostro-tangentially migrating neuroblasts (**Figures S5B** and **S5C**, top). At E17.5, EGFP, Barhl1, and Pitx2 are co-localized in the maturing STN, but not in ventrally

located Pitx2 expressing neurons of the lateral hypothalamus (**Figure 4D**). At this age, EGFP staining was also observed outside the STN. To clearly visualize other neuron types emerging from rostral Lmx1a-expressing progenitors, a mouse line with a tamoxifen-inducible Cre targeted to the *Lmx1a* locus was crossed with a reporter mouse line. Expression of this fluorescent reporter showed that, in addition to the STN, Lmx1a expressing progenitors contribute to other diencephalic-hypothalamic nuclei (for details, see **Figures S6A–S6T**). Interestingly, Lmx1a, Pitx2, and Barhl1 are co-expressed in virtually all cells of the STN and show permutations of co-expression in other diencephalic/hypothalamic glutamatergic neurons (**Figures S6M–S6P**). This characterization demonstrates that Axis-2 corresponds to Lmx1a-positive neural progenitors with an expression profile matching that of glutamatergic neurons of the STN.

Next, we investigated the segregation of Axis-1 and Axis-2 genes in the developing human brain. qPCR analysis of dissected 5.5 week human diencephalon and midbrain showed segregation of Axis-2 (*BARHL1/2*, *WNT8B*, *NKX2-1/4*, and *DBX1*) and Axis-1 (*EN1/2*, *WNT1*, and *CNPY1*) genes, a pattern that was repeated in 7.5-week-old rostral versus caudal VMD samples (**Figure 4E**). 7.5 week human expression patterns of key transcription factors recapitulated the mouse data, as shown by DAB immunostaining of FOXA2, BARHL1, and EN1 through either the developing diencephalon (**Figure 4G**, left) or more posterior rostral and caudal VMD region (**Figure 4G**, middle and right, respectively).

### Related Transcription Factor Signature in STN and mesDA Neurons

Our data revealed a strong transcriptional relationship between mesDA and STN neuronal lineages and show that several critical mesDA neuron transcription factors are co-expressed in STN neural progenitors including *Lmx1a/b*, *Foxa1/2*, *Otx2*, *Foxp1/2*, and *Msx1*. *Lmx1a*-EGFP knockout mice display a modest reduction in the number of mesDA neurons, while genetic ablation of both *Lmx1a* and *Lmx1b* results in a complete disruption of mesDA neurogenesis (Deng et al., 2011; Yan et al., 2011). Interestingly, *Lmx1a*-EGFP homozygous knockout embryos displayed a reduction in size of the STN (**Figure S5F**). To determine if, similar to mesDA neuron development, generation of the STN requires both *Lmx1a* and *Lmx1b*, strains of floxed *Lmx1a* and *Lmx1b* mice were crossed with CAGG-CreERT2 mice (Hayashi and McMahon, 2002) to generate animals in which both *Lmx1a* and *Lmx1b* could be conditionally ablated by tamoxifen treatment (see **Experimental Procedures**). Embryos tamoxifen treated at E10.5, a time point after the developmental window

(B) Sagittal schematic of an E17.5 embryo is shown in the top right corner indicating developing mesDA (green) and STN (red) neurons. STN, lateral hypothalamus: LH.

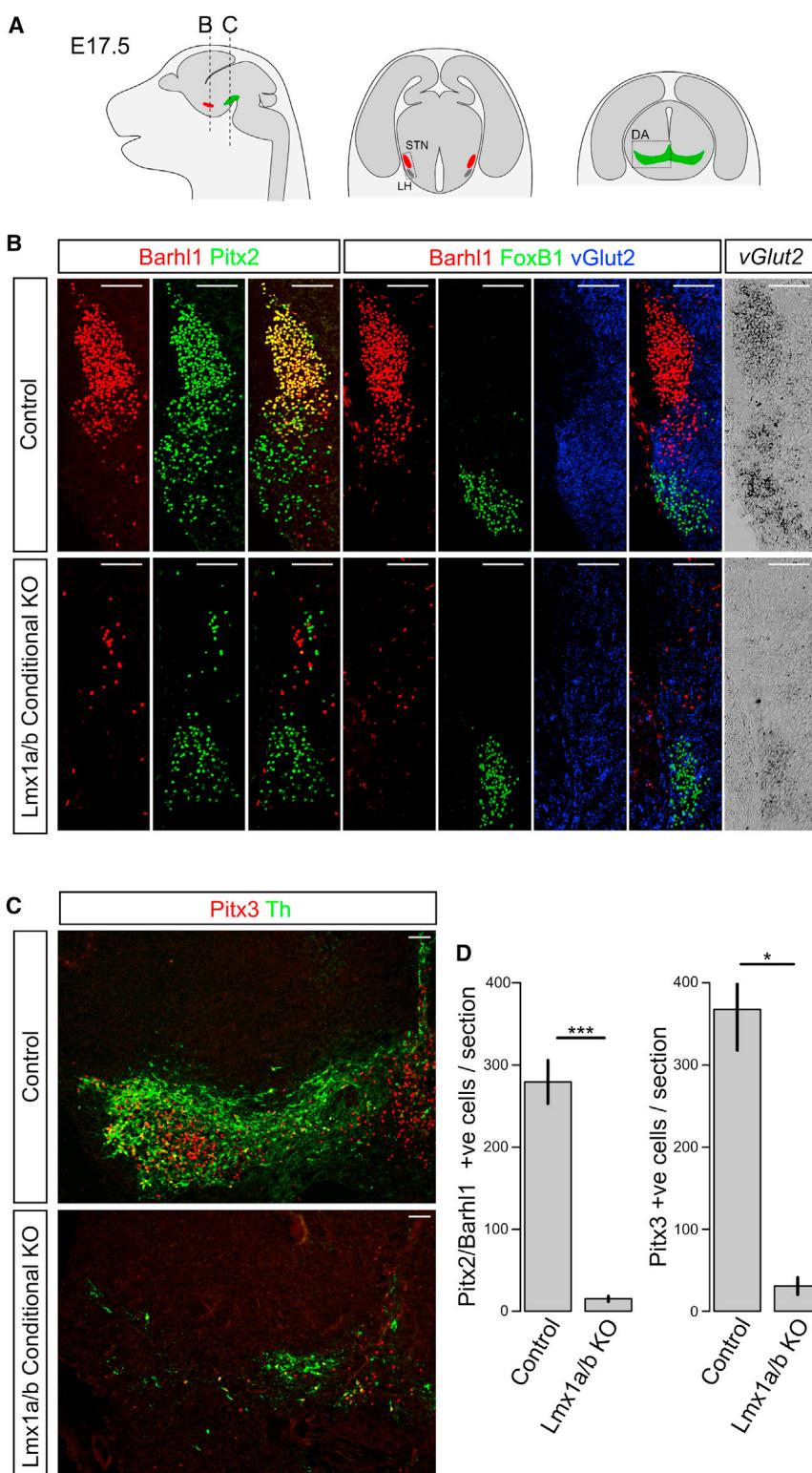
(C) Immunostaining of indicated markers in sagittal sections from E11.5 Lmx1a<sup>EGFP</sup> heterozygous embryos. All sections were stained for EGFP. The arrowheads indicate the rostro-caudal boundaries of the EGFP expression. The red arrowhead or arrow indicates weak or strong En1 expression, respectively (the scale bar represents 100 μm).

(D) Immunostaining of indicated markers in coronal sections from E17.5 Lmx1a<sup>EGFP</sup> heterozygous embryos. The top shows low-magnification images (boxed section in B). The white boxed regions are shown in higher magnification below, where Lmx1a, Pitx2, and Barhl1 are strongly co-expressed within the STN. The ventrally located lateral hypothalamic neurons expressing Pitx2 do not co-express Lmx1a or Barhl1.

(E) qPCR analysis of primary dissected human tissue displays the segregation of Axis-1 and Axis-2 genes in week 5.5 and 7.5 VMD tissue.

(F) Schematic of week 7.5 human embryo showing rostral-to-caudal levels of sections in (G).

(G) FOXA2, BARHL1, and EN1 expression at three rostro-caudal levels of week 7.5 human ventral diencephalon and midbrain. The boxed regions are shown in adjacent higher magnifications. As in the mouse, the strongest EN1 expression is seen at the most caudal level (black arrows).



**Figure 5. Lmx1a and Lmx1b Are Essential for STN Neurogenesis**

(A) Left: sagittal schematic of an E17.5 mouse brain showing rostro-caudal levels at which embryos were sectioned (B and C). The rostral level coronal schematic depicting STN neurons (red) and lateral hypothalamic (LH) neurons (gray) is shown (middle). The caudal level coronal schematic depicting mesDA neurons (green) is shown (right).

(B) Immunohistochemistry of Barhl1/Pitx2 double or Barhl1/FoxB1/vGlut2 triple stains in control (top row) and Lmx1a/Lmx1b double conditional knockout embryos (bottom row). The right most image shows *in situ* hybridization for *vGlut2* mRNA. The STN nucleus is absent in conditional knockout embryos, while ventrally located lateral hypothalamic neurons remain (the scale bar represents 100  $\mu$ m).

(C) Immunohistochemistry of mesDA neuron markers Pitx3 and TH in control (top) and Lmx1a/b double conditional knockout embryos (bottom). The mesDA neurons are almost completely lost in the conditional knockout.

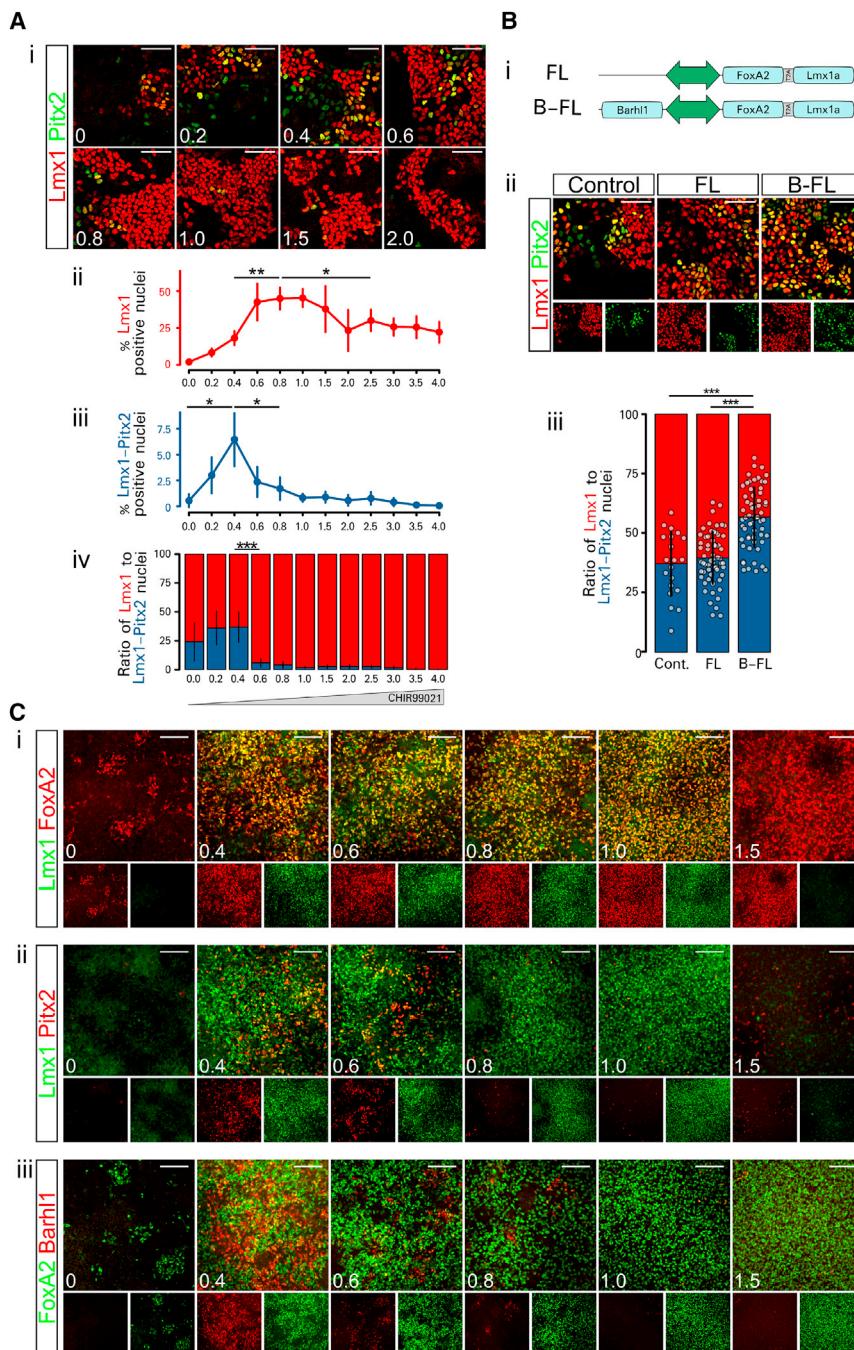
(D) Left: two-way Student's t test statistical quantification of (B) showing significant loss of Pitx2 and Barhl1 co-expressing cells ( $p < 0.001$ , mean  $\pm$  SD,  $n = 4$  control, and five knockout [KO] animals). Two-way Student's t test of (C) displays significant loss of Pitx3 expressing cells ( $p < 0.05$ , mean  $\pm$  SD,  $n = 3$  control, and three KO animals) (right).

position compared to control littermates (Figure S5C, lower), and at E17.5 the characteristic olive shaped STN displaying co-localized Barhl1, Pitx2, and *vGlut2* expression in control embryos was completely lost in mutants (Figures 5B and 5D, left). Ventrally located lateral hypothalamic neurons displaying expression of Pitx2, Foxb1, and *vGlut2* were not lost. As expected, mesDA neurogenesis was abolished in these embryos as shown by an almost complete absence of mesDA neuron markers Pitx3 and TH (Figures 5C and 5D, right). In conclusion, the close lineage relationship indicated by a co-expressed transcription factor code extends to a functional co-dependency on Lmx1a and Lmx1b for development of both mesDA and STN neurons.

#### Rostrocaudal Fine Patterning of ESCs Can Bias Differentiation of Axis-1 or Axis-2 Lineages

We wished to investigate if fine-tuned rostrocaudal patterning of mouse ESCs could bias differentiation toward either Axis-1 or Axis-2 lineage development. GSK-3 beta inhibitors have been used to direct rostrocaudal patterning of differentiating hESCs into mesDA neurons (Kirkeby et al., 2012; Kriks et al., 2011; Xi et al., 2012). VMD patterned mouse ESCs (mESCs) were cultured

where Lmx1b influences the mid/hindbrain border (Guo et al., 2007), appeared morphologically normal; however, STN neurogenesis was severely disrupted. Thus, at E13.5, developing STN neurons could not be seen in their appropriate rostral/lateral



**Figure 6. Rostrocaudal Patterning of ESC Differentiations Segregates Axis-1 and Axis-2 Genes**

(A) (i–iv) mESCs were differentiated across a gradient of CHIR99021 and stained for Lmx1a/b (Lmx1) and Pitx2 (the scale bar represents 50 μm). The quantification across the gradient of (ii) percent Lmx1a/b (Lmx1) single positive, (iii) Lmx1a/b-Pitx2 double positive cells, or (iv) the ratio of Lmx1a/b (Lmx1) versus Lmx1a/b-Pitx2 double positive cells (mean ± SD) is shown.

(B) (i–iii) Schematic illustrating lentiviral vectors, which drive the expression of either Foxa2/Lmx1a or Foxa2/Lmx1a/Barhl1. (ii) Cells cultured as in (A), using 0.4 μM CHIR99021, transduced with indicated viruses, and immunostained for Lmx1/Pitx2. (iii) Two-way Student's t test displays a significant increase in Lmx1a/b-Pitx2 double positive cells after forced expression of Barhl1 ( $p < 0.001$ , mean ± SD, non-transduced  $n = 20$ , FL  $n = 60$ , and B-FL  $n = 60$ ).

(C) (i–iii) hESCs were differentiated according to Kirkeby et al. (2012), along a gradient of CHIR99021 and immunostained for Lmx1/FoxA2 (i), Lmx1/Pitx2 (ii), or FoxA2/Barhl1 (iii) (the scale bar represents 100 μm).

along a concentration gradient of GSK-3 inhibitor CHIR99021. Consistent with their profile in both rostral and caudal domains, Lmx1/Foxa2 double positive and Lmx1 single positive cells were observed at both low and high levels of inhibitor (Figures S6U ii and iii and 6A i and ii). By contrast, rostral domain Lmx1/Pitx2 double positive cells were observed only at low concentrations of CHIR99021, with their proportion dropping sharply above 0.4 μM (Figures 6A iii and iv). Additional Axis-1 and Axis-2 genes reproduced this rostro-caudal segregation (Figure S6U i).

The persistent expression of Barhl1 from early progenitors into mature STN neurons indicates an important role for this

transcription factor in distinguishing rostral versus caudal Lmx1a-expressing lineages. We therefore tested if forced expression of Barhl1 also influenced the ratio of rostral Lmx1/Pitx2 double positive cells versus caudal Lmx1 single positive cells in mESC-derived progenitors patterned by 0.4 μM CHIR99021. Indeed, forced expression of Barhl1 (together with Lmx1a and Foxa2) significantly increased the proportion of Lmx1/Pitx2 double positive cells as compared to non-transduced cells or Lmx1a and Foxa2 transduced cells (Figure 6B). Considering that roughly 15% of all progenitors were transduced (Figure S6V), the results indicate a high efficiency induction of Lmx1/Pitx2-positive progenitors. Together, these results suggest that Barhl1 may function as a lineage determinant in ventral diencephalic Lmx1a progenitor cells.

The finding that classical mesDA neuron markers used to assess differentiation outcome are all co-expressed in the more rostral STN lineage is of considerable relevance for human therapeutic stem cell engineering and raises the concern that previous stem cell protocols may have been guided toward generation of both mesDA neurons and more rostral glutamatergic neurons. We therefore tested if fine-tuned rostrocaudal patterning using CHIR99021 has potential to further refine differentiation outcome of VM patterned hESCs (Kirkeby et al., 2012). Differentiation at CHIR99021 concentrations ranging from 0.4–1 μM resulted in the majority of neural progenitors expressing LMX1A,

FOXA2, and OTX2 (Figures 6C and S6W). Importantly, many cells were PITX2/LMX1 and BARHL1/FOXA2 double positive at 0.4–0.6  $\mu$ M, while these progenitors were mostly absent in more caudalized cultures differentiated at higher concentrations of CHIR00021. Thus, fine patterning of rostrocaudal level has potential to eliminate more rostral contaminating cell types under conditions that would otherwise appear appropriately patterned based on co-expression of LMX1A, FOXA2, and OTX2. Indeed, consistent with this conclusion, Kirkeby et al. (2016) demonstrate that caudalizing culture conditions eliminates STN neuron contamination and markedly improves hESC differentiation and grafting outcome.

## DISCUSSION

A motivation for understanding mesDA neuron development is the prospect of using stem cell derived mesDA neurons in cell-replacement therapy, a strategy requiring detailed understanding of mesDA neurogenesis for the production of clinically viable transplantable cells. Here, we demonstrate that single-cell RNA-seq can be used to robustly reconstruct mesDA neurogenesis with high temporal resolution. Importantly, we show that essentially all markers used to verify developing stem cell-derived human mesDA progenitor cells prior to grafting in PD animal models are shared between caudal dopaminergic and more rostral glutamatergic Lmx1a-expressing progenitor cells giving rise to the STN. These findings suggest that more robust generation of human mesDA neurons from stem cells may require precise rostro-caudal patterning that surveys the expression of STN markers, in particular when considering that GSK-3 inhibitor mediated manipulation of the rostrocaudal axis may vary between different hESC and patient specific iPSC lines. Moreover, grafted cells are transplanted at an immature stage before terminal dopaminergic markers are expressed, emphasizing the need to identify specific early markers that can better predict graft outcome. Importantly, Kirkeby et al. (2016) show that STN neurons are indeed a source of contamination in grafted cells and, in line with the single cell analysis, demonstrate that fine-tuned rostro-caudal patterning reduces STN marker expression and significantly improves graft outcome.

Additional mesDA neuron subtypes, recently described using a single cell qPCR approach (Poulin et al., 2014), were not seen in our data, presumably because mesDA neuron subtypes become distinguished at later stages of development than analyzed here. However, our single cell analysis has elucidated diversity of VMD cells that neighbor the developing Lmx1a lineage, including *Dlx/Pax6* or *Gata2-3/Tal1/Sox14* expressing GABAergic neurons, *Onecut1-2/Foxd2* expressing glutamatergic neurons, and *Isl1/Phox2a-b* expressing oculomotor neurons. Since mesDA stem cell protocols are designed to mimic the regional context of the developing VMD region, we expect these data will contribute an important resource in refining stem cell engineering protocols enriching for mesDA or neighboring lineages. To facilitate this, we provide a web-based resource whereby any individual gene can be projected onto our single cell data set ([http://rshiny.nbis.se/shiny-server-apps/shiny-apps-scrnaseq/Kee\\_2016/](http://rshiny.nbis.se/shiny-server-apps/shiny-apps-scrnaseq/Kee_2016/)).

Understanding how cell type diversity in the CNS is generated has been a long-standing research focus, and numerous studies

have unravelled principles underlying lineage specification and pan-neuronal differentiation. Of note, analysis of the single-cell data resolved multiple, but distinct, phases of progressive neuronal maturation. For example, the reconstituted differentiation axis illustrated poorly understood transitions in progenitor cell maturation. *In vivo*, the ventricular zone harbors proliferating progenitors at least until E13.5. However, *in silico* progenitor cells exhibit diverse attributes along PC1 that distinguish an early phase of rapid cell division from a later phenotype associated with slower proliferation and induction of factors promoting pan-neuronal differentiation (Figures S1E–S1G). The regulatory program guiding this temporal switch is not clearly understood, and additional analysis of single-cell data thus holds potential to resolve key mechanisms governing such transitions. From a clinical perspective, further understanding of temporal transitions can potentially help to control multiple parameters relevant to stem cell therapies, including efforts to isolate appropriately staged progenitors (Ganat et al., 2012) or derive adult-like neurons more rapidly during *in vitro* stem cell differentiation (Nicholas et al., 2013).

Computationally, we found that interrogation of our data was more successful when considering the Lineage Intersect gene list, as compared to top variable genes or all expressed genes. In data sets where a differentiation path contributes substantially more variation than lineage differences, applying lineage-centric gene lists may prove useful in resolving subtle features from whole transcriptome data. Further, while we employed differential expression to weight genes before tSNE projection, in principal, other rationales can be used, such as consideration for the biological relevance of a gene or gene set or the expression of reporter transgenes.

Three transcription factors in the identified STN gene signature, Lmx1a, Barhl1, and Pitx2, were maintained in STN cells at all ages. Permutations of these transcription factors were expressed in additional hypothalamic neurons, suggesting that the STN may head a family of related neural subtypes emerging from the diencephalic Lmx1a-expressing domain. The STN is a key relay station of the indirect output pathway of the basal ganglia, a circuitry that is critically modulated by mesDA neurotransmission (DeLong, 1990). As a consequence of mesDA neuron degeneration in PD, glutamatergic excitatory output from the STN is increased, providing the rationale for deep brain stimulation of the STN to relieve motor symptoms in PD patients (Benabid et al., 2009). Our analysis reveals a remarkably close relationship between mesDA and STN neuronal lineages, but additionally suggests candidate factors critical for specification of STN neurons. Given their selective expression in diencephalic Lmx1a progenitors, Barhl1/2, Pitx2, and Dbx1 in particular stand out. Further, our finding that forced expression of Barhl1 in mouse ESC-derived progenitors can robustly induce Pitx2/Lmx1 expressing cells is also consistent with a function in STN neuron determination. Interestingly, no transcription factors previously found to be essential for mesDA neurogenesis are specifically expressed in all mesencephalic Lmx1a progenitors. Taken together, these observations suggest that Barhl1 may function to distinguish between these two neuron types by suppressing the mesDA neural fate while promoting STN neuron commitment.

In summary, our single cell analysis uncovered mesDA neurogenesis at high resolution at the genome-wide expression level

and identified a close relationship with developing STN neurons. Many signaling molecules and transcription factors are shared between the two lineages, several of which have been essential to predict successful differentiation of human pluripotent stem cells into mesDA progenitors. Moving forward, it is likely that graft heterogeneity between different hESC lines, patient specific iPSC lines, and protocols (Kirkeby et al., 2016) will present a considerable barrier. Our results thus have important implications for faithful and highly reproducible production of clinical grade transplantable mesDA cells.

## EXPERIMENTAL PROCEDURES

### Library Preparation and Sequencing

VMD tissue was dissected from Lmx1a-EGFP ± embryos, dissociated, and sorted into positive or negative fractions by FACS. Single cells or 100-cell pools were manually picked and processed using the Smartseq2 protocol (Picelli et al., 2014). Reads were aligned to the mm10 genome, and expression values calculated using rpkmforgenes. See [Supplemental Experimental Procedures](#) for further details.

### Bioinformatics Analyses

For bioinformatic analyses, see [Supplemental Experimental Procedures](#).

### Immunohistochemistry and Knockout Mice

All mouse tissue was collected, frozen, and immunostained as described in [Supplemental Experimental Procedures](#). Lmx1a/b double flox mice were crossed with CAGG-CreERT2 mice, pregnant mothers were treated with tamoxifen by gavage at E10.5, embryos collected at E13.5 or E17.5, and finally processed for immunohistochemistry. All animal experiments were approved by the local animal ethics committee and conform to the relevant regulatory standards.

### Human Tissue

Human fetal tissue was obtained from legally terminated embryos with approval of the Swedish National Board of Health and Welfare in accordance with existing guidelines including informed consent from women seeking elective abortions. The gestational age of each embryo was determined by measuring the crown-to-rump length and neck-to-rump length. See [Supplemental Experimental Procedures](#) for further details.

### ESC Differentiations

For the monolayer protocol, mESCs were differentiated for 7 days, RNA was extracted using RNEasy micro kits (QIAGEN), cDNA was synthesized, and SYBR Green qPCR was performed. For the 5-stage protocol, mESCs were differentiated for 10 days and processed for immunohistochemistry. hESCs were differentiated as previously described (Kirkeby et al., 2012). Lentiviral particles were produced as previously described (Zufferey et al., 1997). See [Supplemental Experimental Procedures](#) for further details.

### ACCESSION NUMBERS

The accession number for the raw read sequence data, cell annotations, and RPKM and read count expression matrices for all cells reported in this paper is GEO: GSE87069.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.10.003>.

### AUTHOR CONTRIBUTIONS

Conceptualization, N.K. and T.P.; Methodology, N.K. and T.P.; Software, N.K.; Analyses, N.K., H.S., and Å.K.B.; Investigation, N.K., N.V., A.K., L.D., S.N.,

L.L., L.G., and E.J.; Writing – Original Draft, N.K. and T.P.; Writing – Review & Editing, N.K., N.V., and T.P.; Visualization, N.K.; Supervision, T.P., R.S., and M.P.; Project Administration, N.K. and T.P.; and Funding Acquisition, T.P. and M.P.

### ACKNOWLEDGMENTS

We wish to thank Daniel Edsgård for statistical advice, Daniel Ramsköld for bioinformatic assistance, and members at the Stockholm Branch of the Ludwig Institute for support. We give further thanks to Jenny Nelander Wahlestet and Bengt Mattson for valuable help with human expression data, Mattias Karlén for illustrations, and Jonas Muhr and Johan Holmberg for valuable manuscript suggestions. This work was supported by grants from the Knut and Alice Wallenberg's Foundation (grant 2013.0075 to T.P.), grants from the European Union's, 7th Framework Programme under grant agreements mdDANeurodev (grant 222999) and NeuroStemCell (grant 222943 to T.P.), the Swedish Strategic Research Foundation (SSF; to T.P.), the Sigrid Juselius Foundation (to L.L.), European Union's 7th Framework Programme (FP/2007-2013) / ERC grant 309712 (to M.P.), NeuroStemcellRepair grant 602278 (to M.P.), and the Swedish Research Council (VR, K2014-61X-20391-084; to M.P.).

Received: February 29, 2016

Revised: June 23, 2016

Accepted: October 1, 2016

Published: October 27, 2016

### REFERENCES

- Andersson, E., Jensen, J.B., Parmar, M., Guillemot, F., and Björklund, A. (2006a). Development of the mesencephalic dopaminergic neuron system is compromised in the absence of neurogenin 2. *Development* 133, 507–516.
- Andersson, E., Tryggvason, U., Deng, Q., Friling, S., Alekseenko, Z., Robert, B., Perlmann, T., and Ericson, J. (2006b). Identification of intrinsic determinants of midbrain dopamine neurons. *Cell* 124, 393–405.
- Arenas, E., Denham, M., and Villaescusa, J.C. (2015). How to make a midbrain dopaminergic neuron. *Development* 142, 1918–1936.
- Asbreuk, C.H.J., Vogelaar, C.F., Hellemans, A., Smidt, M.P., and Burbach, J.P.H. (2002). CNS expression pattern of Lmx1b and coexpression with ptx genes suggest functional cooperativity in the development of forebrain motor control systems. *Mol. Cell. Neurosci.* 21, 410–420.
- Barker, R.A., Drouin-Ouellet, J., and Parmar, M. (2015). Cell-based therapies for Parkinson disease—past insights and future potential. *Nat. Rev. Neurol.* 11, 492–503.
- Benabid, A.L., Chabardes, S., Mitrofanis, J., and Pollak, P. (2009). Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. *Lancet Neurol.* 8, 67–81.
- Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. USA* 112, 7285–7290.
- DeLong, M.R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* 13, 281–285.
- Deng, Q., Andersson, E., Hedlund, E., Alekseenko, Z., Coppola, E., Panman, L., Millonig, J.H., Brunet, J.-F., Ericson, J., and Perlmann, T. (2011). Specific and integrated roles of Lmx1a, Lmx1b and Phox2a in ventral midbrain development. *Development* 138, 3399–3408.
- Doi, D., Samata, B., Katsukawa, M., Kikuchi, T., Morizane, A., Ono, Y., Sekiguchi, K., Nakagawa, M., Parmar, M., and Takahashi, J. (2014). Isolation of human induced pluripotent stem cell-derived dopaminergic progenitors by cell sorting for successful transplantation. *Stem Cell Reports* 2, 337–350.
- Ganat, Y.M., Calder, E.L., Kriks, S., Nelander, J., Tu, E.Y., Jia, F., Battista, D., Harrison, N., Parmar, M., Tomishima, M.J., et al. (2012). Identification of embryonic stem cell-derived midbrain dopaminergic neurons for engraftment. *J. Clin. Invest.* 122, 2928–2939.

- Guo, C., Qiu, H.-Y., Huang, Y., Chen, H., Yang, R.Q., Chen, S.-D., Johnson, R.L., Chen, Z.F., and Ding, Y.-Q. (2007). Lmx1b is essential for Fgf8 and Wnt1 expression in the isthmic organizer during tectum and cerebellum development in mice. *Development* 134, 317–325.
- Hayashi, S., and McMahon, A.P. (2002). Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* 244, 305–318.
- Jaeger, I., Arber, C., Risner-Janiczek, J.R., Kuechler, J., Pritzsche, D., Chen, I.-C., Naveen, T., Ungless, M.A., and Li, M. (2011). Temporally controlled modulation of FGF/ERK signaling directs midbrain dopaminergic neural progenitor fate in mouse and human pluripotent stem cells. *Development* 138, 4363–4374.
- Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 343, 776–779.
- Kele, J., Simplicio, N., Ferri, A.L.M., Mira, H., Guillemot, F., Arenas, E., and Ang, S.-L. (2006). Neurogenin 2 is required for the development of ventral midbrain dopaminergic neurons. *Development* 133, 495–505.
- Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. *Nat. Methods* 11, 740–742.
- Kirkeby, A., Grealish, S., Wolf, D.A., Nelander, J., Wood, J., Lundblad, M., Lindvall, O., and Parmar, M. (2012). Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. *Cell Rep.* 1, 703–714.
- Kirkeby, A., Nolbrant, S., Tiklova, K., Heuer, A., Kee, N., Cardoso, T., Ottosson, D.R., Lelos, M.J., Rifes, P., Dunnett, S.P., et al. (2016). Predictive Markers Guide Differentiation to Improve Graft Outcome in Clinical Translation of hESC-Based Therapy for Parkinson's Disease. *Cell Stem Cell* 20. Published online October 27, 2016. <http://dx.doi.org/10.1016/j.stem.2016.09.004>.
- Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161, 1187–1201.
- Kriks, S., Shim, J.-W., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., et al. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* 480, 547–551.
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559.
- Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* 162, 184–197.
- Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214.
- Marín, F., Herrero, M.-T., Vyas, S., and Puelles, L. (2005). Ontogeny of tyrosine hydroxylase mRNA expression in mid- and forebrain: neuromeric pattern and novel positive regions. *Dev. Dyn.* 234, 709–717.
- Martin, D.M., Skidmore, J.M., Philips, S.T., Vieira, C., Gage, P.J., Condie, B.G., Raphael, Y., Martinez, S., and Camper, S.A. (2004). PITX2 is required for normal development of neurons in the mouse subthalamic nucleus and midbrain. *Dev. Biol.* 267, 93–108.
- Nicholas, C.R., Chen, J., Tang, Y., Southwell, D.G., Chalmers, N., Vogt, D., Arnold, C.M., Chen, Y.-J.J., Stanley, E.G., Elefant, A.G., et al. (2013). Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. *Cell Stem Cell* 12, 573–586.
- Picelli, S., Björklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* 10, 1096–1098.
- Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9, 171–181.
- Pollen, A.A., Nowakowski, T.J., Shuga, J., Wang, X., Leyrat, A.A., Lui, J.H., Li, N., Szpankowski, L., Fowler, B., Chen, P., et al. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* 32, 1053–1058.
- Poulin, J.-F., Zou, J., Drouin-Ouellet, J., Kim, K.-Y.A., Cicchetti, F., and Awatermani, R.B. (2014). Defining midbrain dopaminergic neuron diversity by single-cell gene expression profiling. *Cell Rep.* 9, 930–943.
- Rosvall, M., and Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* 105, 1118–1123.
- Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 510, 363–369.
- Smidt, M.P., Cox, J.J., van Schaick, H.S., Coolen, M., Schepers, J., van der Kleij, A.M., and Burbach, J.P. (2000). Analysis of three Ptx2 splice variants on transcriptional activity and differential expression pattern in the brain. *J. Neurochem.* 75, 1818–1825.
- Smidt, M.P., Smits, S.M., Bouwmeester, H., Hamers, F.P.T., van der Linden, A.J.A., Hellemons, A.J.C.G.M., Graw, J., and Burbach, J.P.H. (2004). Early developmental failure of substantia nigra dopamine neurons in mice lacking the homeodomain gene Pitx3. *Development* 131, 1145–1155.
- Steinbeck, J.A., and Studer, L. (2015). Moving stem cells to the clinic: potential and limitations for brain repair. *Neuron* 86, 187–206.
- Thompson, C.L., Ng, L., Menon, V., Martinez, S., Lee, C.-K., Glattfelder, K., Sunkin, S.M., Henry, A., Lau, C., Dang, C., et al. (2014). NeuroResource. *Neuron* 83, 309–323.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386.
- Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 509, 371–375.
- van der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* 15, 1–21.
- Xi, J., Liu, Y., Liu, H., Chen, H., Emborg, M.E., and Zhang, S.-C. (2012). Specification of midbrain dopamine neurons from primate pluripotent stem cells. *Stem Cells* 30, 1655–1663.
- Yan, C.H., Levesque, M., Claxton, S., Johnson, R.L., and Ang, S.-L. (2011). Lmx1a and Lmx1b function cooperatively to regulate proliferation, specification, and differentiation of midbrain dopaminergic progenitors. *J. Neurosci.* 31, 12413–12425.
- Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnérberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142.
- Zetterström, R.H., Solomin, L., Jansson, L., Hoffer, B.J., Olson, L., and Perlmann, T. (1997). Dopamine neuron agenesis in Nurr1-deficient mice. *Science* 276, 248–250.
- Zhao, T., Szabó, N., Ma, J., Luo, L., Zhou, X., and Alvarez-Bolado, G. (2008). Genetic mapping of Foxb1-cell lineage shows migration from caudal diencephalon to telencephalon and lateral hypothalamus. *Eur. J. Neurosci.* 28, 1941–1955.
- Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., and Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* 15, 871–875.