This Accepted Manuscript has not been copyedited and formatted. The final version may differ from this version. A link to any extended data will be provided when the final version is posted online.

eNeuro

Research Article: New Research | Sensory and Motor Systems

Molecular profiling defines evolutionarily conserved transcription factor signatures of major vestibulospinal neuron groups

Anders Lunde[1], Benjamin W. Okaty[2], Susan M. Dymecki[2] and Joel C. Glover[1]

[1]*Dept. of Molecular Medicine, University of Oslo, 0372, Oslo, Norway*
[2]*Dept. of Genetics, Harvard Medical School, Boston, MA, 02115*

https://doi.org/10.1523/ENEURO.0475-18.2019

Received: 6 December 2018

Revised: 4 January 2019

Accepted: 25 January 2019

Published: 1 February 2019

AL, BWO, SMD and JCG Designed Research; AL, BWO and JCG Performed Research; AL, BWO and JCG Analyzed data, AL, BWO, SMD, and JCG Wrote the paper.

Funding: South-East Norway Regional Health Authority
2013022

Conflict of Interest: Authors report no conflict of interest.

South#East Norway Regional Health Authority, Project number 2013022.

Correspondence should be addressed to Joel C. Glover, joel.glover@medisin.uio.no

Cite as: eNeuro 2019; 10.1523/ENEURO.0475-18.2019

Alerts: Sign up at www.eneuro.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Accepted manuscripts are peer-reviewed but have not been through the copyediting, formatting, or proofreading process.

Copyright © 2019 Lunde et al.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

eNeuro Accepted Manuscript

eNeuro Accepted Manuscript

Manuscript title

Molecular profiling defines evolutionarily conserved transcription factor signatures of major vestibulospinal neuron groups

Abbreviated title

Vestibulospinal transcription factor signatures

Author Names and Affiliations

Anders Lunde[1], Benjamin W. Okaty[2], Susan M. Dymecki[2], Joel C. Glover[1]

[1]Dept. of Molecular Medicine, University of Oslo, 0372 Oslo, Norway; [2]Dept. of Genetics, Harvard Medical School, Boston, MA 02115

Author contributions

AL, BWO, SMD and JCG Designed Research; AL, BWO and JCG Performed Research; AL, BWO and JCG Analyzed data, AL, BWO, SMD, and JCG Wrote the paper.

Corresponding author

Correspondence should be addressed to Joel C. Glover.

Address: Department of Molecular Medicine, University of Oslo, 0372 Oslo, Norway.

Email: joel.glover@medisin.uio.no

Number of figures: 10

Number of tables: 4

eNeuro Accepted Manuscript

Number of multimedia: 0

Number of words for Abstract: 149

Number of words for Significance Statement: 81

Number of words for Introduction: 750

Number of words for Discussion: 2674

Acknowledgements

This project was funded by South-East Norway Regional Health Authority, project number 2013022. We thank Jean-François Brunet for anti-chicken Phox2b antibody, Frederic Clotman for anti-Onecut antibodies, Martyn Goulding for anti-Evx1 antibody, Thomas Müller for anti-Lbx1 antibody, Ryuichi Shirasaki for anti-Evx2 antibody, Sasha Ilovar, Margaret Russell and Anne Helene Fosby for assistance with some of the immunohistology, and Marian Berg Andersen and Kobra Sultani for additional technical assistance. We thank Bernd Fritzsch and Christo Goridis for critical comments on an earlier draft of the manuscript. We thank the Norwegian Sequencing Centre for guidance and sequencing services.

Conflict of interest

Authors report no conflict of interest

Funding sources

South-East Norway Regional Health Authority, Project number 2013022.

eNeuro Accepted Manuscript

All animal procedures were were approved by the Norwegian Animal Research Authority (Forsøksdyrutvalget, FDU ID. 8473) and performed in accordance with the University of Oslo animal care committee's regulations and followed the Federation of European Laboratory Animal Science Associations (FELASA) guidelines.

eNeuro Accepted Manuscript

47 **Molecular profiling defines evolutionarily conserved transcription factor**

48 **signatures of major vestibulospinal neuron groups**

49

eNeuro Accepted Manuscript

ABSTRACT:

Vestibulospinal neurons are organized into discrete groups projecting from brainstem to spinal cord, enabling vertebrates to maintain proper balance and posture. The two largest groups are the lateral vestibulospinal tract (LVST) group and the contralateral medial vestibulospinal tract (cMVST) group, with different projection lateralities and functional roles. In search of a molecular basis for these differences, we performed RNA sequencing on LVST and cMVST neurons from mouse and chicken embryos followed by immunohistofluorescence validation. Focusing on transcription factor (TF)-encoding genes, we identified TF signatures that uniquely distinguish the LVST from the cMVST group and further parse different rhombomere-derived portions comprising the cMVST. Immunohistofluorescence assessment of the CNS from spinal cord to cortex demonstrated that these TF signatures are restricted to the respective vestibulospinal groups and some neurons in their immediate vicinity. Collectively, these results link the combinatorial expression of TFs to developmental and functional subdivisions within the vestibulospinal system.

eNeuro Accepted Manuscript

Significance statement:

The molecular underpinnings of hodological and functional subdivisions within brainstem-to-spinal cord projection neurons are poorly understood. Transcriptomic profiling is an important step towards obtaining a molecular characterization of individual projection neuron groups and identifying candidate genes potentially involved in their specification. Here we use this approach to identify transcription factor signatures conserved in mouse and chicken that distinguish the two major vestibulospinal projection neuron groups and define coherent subpopulations within them, using whole transcriptome sequencing and immunohistofluorescence combined with retrograde tracing.

eNeuro Accepted Manuscript

Introduction:

Projection neurons in the brainstem represent the major source of descending inputs to spinal circuits and comprise a large set of neuron groups with diverse functions. Their original anatomical classification led to the definition of projection systems such as the reticulospinal, vestibulospinal, tectospinal and rubrospinal. Many of these are markedly heterogeneous, composed of distinct subpopulations with differing projection pathways, spinal targets and functional roles. How this diverse collection of projection neuron groups arises and collectively regulates spinal function is poorly understood. A key element in understanding this lies in the molecular programs that specify neuron groups within each projection system. A first step towards elucidating such programs is to identify molecular signatures that distinguish the different neuron subpopulations.

The vestibulospinal (VS) projection is one of the better characterized brainstem projection systems, and plays a central role in the control of posture and movement. VS neurons receive direct and indirect input from the peripheral vestibular organs, and innervate alpha motoneurons in the spinal cord monosynaptically and polysynaptically, permitting rapid regulation of musculature in response to movements of the head and body (Grillner et al., 1970; Akaike, 1983; Uchino and Kushiro, 2011; Di Bonito et al., 2015; Kasumacic et al., 2015; Murray et al., 2018). VS neurons also receive inputs from the cerebellum and from other brainstem regions through which their activity can be integrated more broadly into ongoing motor control.

The classical description of the VS system has included the unilaterally projecting lateral vestibulospinal tract (LVST) and the bilaterally-projecting medial vestibulospinal tract (MVST). We showed earlier that the LVST derives from the ipsilaterally-projecting LVST neuron group and the MVST derives from the distinct ipsilaterally-projecting iMVST and contralaterally-projecting cMVST neuron groups (J. C. Glover and Petursdottir, 1988, 1991; Diaz et al., 1998; Diaz et al., 2003; Pasqualetti et al., 2007; Kasumacic et al., 2010; Di Bonito et al., 2015). These groups differ not only in axonal trajectory but also in developmental origins, locations and spinal termination patterns. The LVST neuron group derives from hindbrain rhombomere (r)

eNeuro Accepted Manuscript

4, lies primarily within the lateral vestibular nucleus, and projects along an initially lateral trajectory within the hindbrain and then down the length of the spinal cord. The iMVST neuron group derives from r6 and lies primarily within the descending vestibular nucleus, whereas the cMVST neuron group derives from r4 and 5 and overlaps the lateral, medial and descending vestibular nuclei. Both project in the medial longitudinal fascicle (MLF; respectively ipsilaterally and contralaterally) to cervical and upper thoracic spinal segments. Axons from all three groups synapse on motoneurons and premotor interneurons, with clear examples of differential targeting (Shinoda et al., 2006; Kasumacic et al., 2010; Kasumacic et al., 2015). A fourth VS group, originating from the more caudal portion of the posterior vestibular nucleus, is less well characterized (Peterson and Coulter, 1977; Peterson et al., 1978; Donevan et al., 1992).

The LVST and cMVST neuron groups are the largest of the VS groups and the most clearly conserved across the vertebrate radiation, having been described in mammals, birds, amphibians and fish (mouse: (Pasqualetti et al., 2007); chicken: (J. C. Glover and Petursdottir, 1988); frog: (Straka et al., 2001); fish: (Glover and Fraser (unpublished observations); Suwa et al. (1996))). Their different anteroposterior and dorsoventral origins, axon trajectories and termination patterns suggest that they are genetically programmed by specific profiles of transcription factor (TF) expression (Auclair et al., 1999; J. C. Glover, 2000b; Cepeda-Nieto et al., 2005). An initial effort at profiling the expression of TFs in VS neurons was made by Chen et al. (2012), who demonstrated that the TFs Lbx1 and Phox2a/b are expressed by a population of VS neurons located in the lateral vestibular nucleus (LVN) of the mouse during embryonic development. The LVN contains the LVST and also portions of the cMVST and is thus hodologically heterogeneous (Diaz et al., 2003). To better define how TF expression relates to these VS groups, more comprehensive expression analysis is necessary.

Here, we use whole transcriptome sequencing of LVST and cMVST neurons isolated from mouse and chicken embryos, followed by immunohistochemical validation, to define specific TF profiles distinguishing the LVST and cMVST groups. Immunohistochemical assessment from spinal cord to cortex shows that these TF signatures are restricted the LVST and cMVST groups plus a few vicinal neurons. Additional TFs

eNeuro Accepted Manuscript

define spatially coherent subpopulations within these groups. These results expand our understanding of the molecular identity of VS neuron groups, provide a first step towards unraveling transcriptional heterogeneity within them, and suggest testable hypotheses about potential fate-specifying programs.

eNeuro Accepted Manuscript

Materials and Methods:

Animal handling and dissection

All animal procedures were approved by the Norwegian Animal Research Authority (Forsøksdyrutvalget, FDU ID. 8473) and performed in accordance with the [Author University] animal care committee's regulations and followed the Federation of European Laboratory Animal Science Associations (FELASA) guidelines. In compliance with these regulations, all efforts were made to minimize the number of mice used and their suffering. Unless otherwise noted, experiments were performed on mice of either sex from the Crl:CD1(ICR) line (Research Resource Identifier (RRID):IMSR_CRL:22). For r4 lineage tracing, the *b1r4-Cre* transgenic line (Di Bonito et al., 2013), which expresses c*re recombinase* exclusively in r4 under the control of the *Hoxb1* r4 enhancer (Studer et al., 1994), was used in combination with the *Ai14* Cre reporter line harboring a loxP-flanked STOP cassette preventing transcription of a CAG promoter-driven red fluorescent protein variant (tdTomato) (strain 7914 - Jackson Laboratories, Bar Harbor, ME, USA), both of either sex. The morning of vaginal plug observation was defined as embryonic day (E) 0.5. Pregnant dams were anesthetized with isoflurane before cervical dislocation. Dissected embryos were kept in ice cold (4°C), oxygenated (95% O2-5% CO2), artificial cerebrospinal fluid (ACSF; containing in mM: 128 NaCl, 3 KCl, 11 D-glucose, 2.5 CaCl2, 1 MgSO4, 1.2 NaH2PO4, 5 HEPES, and 25 NaHCO3). Brainstems with cervical spinal cord were dissected out in cold ACSF under a dissection microscope.

Fertilized Ross II chicken eggs of either sex acquired from Nortura, Norway, were stored at 14°C, and incubated at 37.5°C in a humidified forced draft incubator, counting start of incubation as incubation day 0 (d0). On the day of dissection, eggs were cracked open and embryos transferred to ice cold (4°C), oxygenated (100% O2), chicken ringer solution (containing in mM: 137 NaCl, 5 KCl, 11 D-glucose, 2 CaCl2, 1 MgSO4, 1 NaPO4 buffer pH=7.4, 5 HEPES). Embryo stage was determined according to HamburgerHamilton (1992), with d7.5 and d9 corresponding to stages HH30 and HH35, respectively. Brainstems with cervical spinal cord attached were dissected out under a dissection microscope.

eNeuro Accepted Manuscript

Retrograde labeling with conjugated dextran amines

Isolation of LVST and cMVST neurons for RNA sequencing requires that they first be selectively retrogradely labeled via their axons to the spinal cord. To determine the earliest stages at which reliable retrograde labeling of LVST and cMVST neurons could be obtained, we performed such labeling at different stages and anteroposterior levels in mouse and chicken embryos. We found that the LVST group could be well labeled from cervical level (C)1 at E12.5 in the mouse and at d6 (HH stage 28) in the chicken. By contrast, we found that the cMVST group could not be well labeled from C1 until much later, but that we could label cMVST neurons as early as E13.5 in the mouse and d7.5 in the chicken, if tracer was applied in the medial longitudinal fascicle (MLF) midway between cranial nerve nVIII and C1 (mid-medulla oblongata). To minimize sample variation within each species, we chose E13.5 in mice and d7.5 in chicken to retrogradely label both VS groups for manual cell isolation and RNA sequencing, applying tracer at C1 for the LVST, and the mid-medulla oblongata for the cMVST. For immunohistofluorescence, we always retrogradely labeled the LVST group from C1 regardless of stage, and we labeled the cMVST group from mid-medulla oblongata at E13.5 but from C1 at all later stages.

Vestibulospinal neurons were retrogradely labeled for manual cell isolation with tetramethylrhodamine-conjugated dextran amine (RDA; 3 kDa; Invitrogen, Carlsbad, CA), or for immunohistofluorescence with a 1:1 mixture of fluorescein dextran amine (FDA; 3 kDa; Invitrogen, Carlsbad, CA) and biotin dextran amine (BDA; 3 kDa; Invitrogen, Carlsbad, CA) or pure BDA (J. Glover, 1995; Auclair et al., 1999). For labeling LVST neurons, a hemi-transection of the ventral half of the spinal nerve was made at C1, and several small crystals of conjugated dextran amines were applied successively for at least 4 minutes to the transection. Preparations were then incubated in room temperature oxygenated ACSF or chicken ringer for at least 7 hours for immunohistofluorescence, or at least 3 hours for manual cell isolation. For labeling cMVST neurons, conjugated dextran amine crystals were applied to a hemi-transection of the brainstem, extending from the midline to about 400 um laterally, at the level midway between cranial nerve nVIII and C1. Tracer application at C1 is preferable, because it limits labeling to bona fide vestibulospinal axons, as opposed to axons that might terminate within the medulla. Thus, to control that this more rostral

eNeuro Accepted Manuscript

application did not label other neuron populations in the vicinity of the cMVST, we applied RDA unilaterally at C1, waited 6 hours, and then applied FDA/BDA to the MLF on the same side at the mid-medullary level, in d11 chicken embryos (n=2) and postnatal day 1 (P1) mice (n=3). At these late stages most cMVST neurons have extended their axons to C1, and neurons labeled exclusively from mid-medulla could be quantified (FDA/BDA positive, RDA negative). In the chicken, 22% and 26%, and in the mouse, 14%, 15%, and 22% additional cells were labeled in the area of the cMVST from the mid-medullary compared to the C1 application. The additional cells for the most part were interspersed among the cMVST neurons labeled from C1 and appeared to be part of the same coherent neuron group (Extended data 1-1), confirming that mid-medullary labeling does not lead to contamination of the cMVST group by non-cMVST neurons in that region.

RNA sample acquisition

Retrograde labeling was performed with RDA as described above, with the addition of control lesions to minimize labeling of unwanted axonal pathways. The control lesions were performed within 10 minutes after RDA application, and only after complete removal of RDA from the preparation by focal superfusion followed by washing the preparation at least 3 times with ACSF or chicken ringer. After 3 hours of incubation, tissue chunks restricted as closely as possible to the labeled LVST or cMVST groups were carefully dissected out, and placed in 1 mg/mL Pronase (Sigma-Aldrich) in ACSF or chicken ringer for 10 min at RT. Tissue was washed for 10 minutes in ACSF or chicken ringer, triturated to complete dissociation and dissociated RDA-labeled neurons manually sorted as described in Hempel et al. (2007). The isolated neurons were transferred to RNA extraction buffer (PicoPure RNA Isolation Kit; Applied Biosystems, Foster City, California), heated for 30 min at 42°C, and stored at -80°C.

We chose control tissue to include genes that would likely be commonly expressed at the same levels as the VS groups along the anteroposterior (AP) or the dorsoventral (DV) axis (note that in the hindbrain, the DV axis is anatomically displaced to the lateromedial axis, due to the dorsal opening of the fourth ventricle). Caudal control samples consisted of manually sorted non-fluorescent cells from a region (spanning 3-4

eNeuro Accepted Manuscript

rhombomeres) immediately caudal to the LVST neuron group, and thus at a similar DV level. They were collected from the same preparations as the LVST neuron samples and contained only a few hundred cells each. Medial control samples were collected as bulk tissue from separate preparations, after bilateral RDA labeling from C1, with the sample extent delimited rostro-caudally and laterally by the locations of the RDA-labeled LVST and cMVST neuron groups, and thus derived from the same AP level (Figure 2A). Because they were bulk tissue samples, they contained many thousands of cells. The medial control samples were lysed directly after dissection, without trituration and dissociation of cells. In the mouse, we collected 6 LVST, 3 cMVST, 4 caudal control, and 4 medial control biological replicates. In the chicken, we collected 4 LVST, 4 cMVST, and 4 medial control biological replicates, but no caudal control samples. Each biological replicate was processed and sequenced independently, with approximately 110 cells on average collected for each LVST and cMVST replicates. cMVST replicates each contained neurons pooled from at least 3 retrogradely labeled preparations, whereas each LVST replicate was obtained from a single preparation.

RNA sequencing

mRNA was converted to cDNA and amplified with the Ovation RNA-seq System v2 kit (Nugen, San Carlos, CA). Following amplification, cDNA was fragmented (~250 bp) using a Covaris S2 sonicator, and ~50 ng of fragmented cDNA was introduced into the Ovation Ultralow DR Multiplex System (Nugen, San Carlos, CA) to generate bar-coded libraries. Quantification and quality control were assessed using Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) and qPCR. 50 bp, single-end reads were generated on an Illumina HiSeq2500 platform for 2 of the mouse LVST samples and of the 2 caudal control samples, whereas 75 bp single-end reads were generated on an Illumina NextSeq500 for all other samples. Sequencing depth was minimum 19 million reads per sample. RNAseq data files have been uploaded to the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/).

RNAseq data analysis: mapping and alignment, dendrograms, clustergrams, MDS plots

eNeuro Accepted Manuscript

Sequencing data in FASTQ format were aligned to the mouse genome (mm9) using the RUM pipeline (version 1.11 (Grant et al., 2011)) or to the chicken genome (Galgal5) using the STAR aligner (Dobin et al., 2013). Genome feature quantification for mouse data was performed with RUM, using RefSeq annotations, and for chicken data with STAR using the Gallus_gallus.Gallus_gallus-5.0.93.gtf annotation file downloaded from Ensembl.org. Normalization of read counts, multidimensional scaling, and differential expression analyses were performed using edgeR (Robinson et al., 2010). Hierarchichal and bi-clustering were performed in Matlab. Further details of analyses pertinent to figures is given in the figure legends. TFs were defined by Gene Ontology class 6355 (regulation of transcription, DNA templated). Identification of chicken-mouse transcription factor orthologues (Extended data 3-1) was done in Ensembl Biomart (https://www.ensembl.org/biomart) using the "multi species comparisons" tool.

Immunohistofluorescence

Retrogradely labeled brainstems were immersion fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) at 4°C for 30 minutes (embryonic preparations), or 1 hour (d11 chicken, P1 mouse), then washed in PBS, sequentially incubated in 20% and 30% sucrose to equilibration, embedded in Tissue-Tek O.C.T. embedding compound (Sakura, Torrance, CA), frozen in liquid nitrogen, and sectioned transversely at 14 µm using a cryostat. Sections were stored at -20°C or used directly. For immunohistofluorescence, sections were washed once in PBS for 5 minutes, once in 0.1% Tween 20 in tris-buffered saline (TBST) for 5 minutes, blocked in 10% normal donkey serum in TBST (blocking buffer) for 30 minutes at RT, incubated with primary antibodies in blocking buffer overnight at 4°C and washed 3 times 5 min in PBS. They were then incubated with secondary antibodies, and/or fluorophore-conjugated streptavidins, and/or counterstained with 1 µg/mL Hoechst 33342 (Sigma, St. Louis, MO) in TBST for 1 hour at RT, washed 3 times 5 min in PBS, and mounted under coverslips in 1:1 PBS:glycerol or gelatin:H_2O:glycerol 7g:42ml:50ml. Primary antibodies used are listed in Table 1. Secondary antibodies and fluorophore-conjugated streptavidins were obtained from Jackson Immuno Research and Thermo Fisher Scientific, with secondary antibodies dilluted 1:1 000 and streptavidins 1:500.

eNeuro Accepted Manuscript

Imaging and quantification of transcription factor expression

For assessment of colocalization and for quantification of transcription factor (TF) expression in VS neurons, confocal Z-stacks of 2 µm optical sections were acquired from every other transverse section throughout the level of the VS groups (typically 10-15 sections) with Zeiss LMS510 meta, Zeiss LSM700 or Zeiss LSM710 microscopes, using N.A. 1.3 oil or N.A 1.2 water 40x objectives. Colocalization was determined manually, with care taken not to count VS neurons that lacked nuclear Hoechst staining (false negative), and avoiding false positives arising from stacked cells in the z-axis (colocalization in xy plane but not z-axis). This was facilitated by a customized ImageJ (Schneider et al., 2012) macro, which converted manually thresholded colocalized Hoechst and retrograde labeling signals to outlines, which were then superimposed onto contrast-enhanced immunostained images, from which positive and negative neurons for each TF could be counted.

Because of variability in their expression levels, Onecut1, 2 and 3 were quantified in the mouse LVST group both in terms of numbers of neurons and immunostaining intensity. This was done by measuring fluorescent signal intensity (average pixel value) in every retrogradely labeled VS neuron in every fourth transverse section per embryo. Positive versus negative immunostaining was determined using a per embryo threshold value, setting a value that distinguished two distinct populations, corresponding to negative or very weakly stained neurons, from more intensely stained neurons.

Counts and x-y-z coordinates of all neurons immunostained for a given transcription factor were obtained from at least 3 embryos per transcription factor at each stage. Coordinates were plotted as 2D scatterplots and min-max normalized histograms for each cardianal axis using MATLAB Release 2015b, The MathWorks, Inc. (Natick, Massachusetts, United States).

CNS-wide triple and quadruple transcription factor colocalization analysis, and 3D reconstruction

To determine the location and uniqueness of TF signatures, large-scale 3D reconstructions covering the entire brainstem, and (where necessary) portions of the spinal cord and cerebrum, were generated from

eNeuro Accepted Manuscript

images acquired with a Axio Scan Z1 microscope (Zeiss, Oberkochen, Germany), or a Pannoramic Midi microscope (3D histech, Budapest, Hungary) with N.A. 0.8 air 20x objectives. Colocalization for each combination of transcription factors was evaluated in 14 µm immunostained cryosections, in at least 3 individual embryos. Inspection and counting was done in either the left or the right half of at least every 6th transverse section of the brainstem, and where necessary; every 6th parasagittal section of the cerebrum, and at least 3 transversal sections each from the lumbar, cervical and thoracic spinal cord. Identification of neurons with triple- or quadruple-colocalized transcription factors was done by thresholding individual imaging channels either manually or using the built-in automatic local threshold plugin in ImageJ. The resulting masks were applied to successive AND operations with the ImageJ image calculator until a single mask of potential colocalized transcription factor pixels was generated. This mask was used as an aid in locating colocalized transcription factors, but care was taken to verify the validity of each thresholded channel against the original image, and manual inspection was employed where thresholding failed or was uncertain. 3D scatterplots with hindbrain outlines were generated using ImageJ to extract coordinates of colocalized staining and tissue contours from transverse sections. These were then plotted using MATLAB Release 2015b, The MathWorks, Inc., (Natick, Massachusetts, United States), with the alphashape and 3d scatterplot functions. For quantifying colocalization of retrogradely labeled LVST neurons and Phox2b/Esrrg/Maf, or r5-cMVST neruons and Evx2/Esrrg/Maf, only Phox2b (LVST marker) or Evx2 (r5-cMVST marker) positive neurons were considered, in order to avoid counting neurons that lacked nuclei (false negatives).

Statistical Table

Statistical analyses were carried out in the R software version 3.3.3 (https://www.r-project.org/) and MATLAB Release 2015b, The MathWorks, Inc. (Natick, Massachusetts, United States). Two-sample Kolmogorov-Smirnov tests were calculated with the MATLAB function kstest2. See the RNAseq data analysis paragraph in Materials and Methods for additional details on RNAseq data analysis.

eNeuro Accepted Manuscript

Data structure	Parameter tested	Type of test	P value	Figure
RNAseq data passing filter (counts per million > 1 in at least 3 samples), grouped by sample type	Gene counts	Gene by gene ANOVA-like differential abundance analysis to test for differences between any sample groups using edgeR	FDR < 0.1	Used as filter criteria in 3, and 2-1A,B . Specific values listed in 2-2, 2-3. For 2F,G a more stringet threshold of FDR < 5e-4 was used.
Two independent samples: r5 and r4 cMVST cell population, pooled normalized data from n=3 animals.	Esrrg fluorescence intensity	Two-sample Kolmogorov-Smirnov test	3.6e-7	7C

306

eNeuro Accepted Manuscript

Results

A) The LVST and cMVST are hodologically distinct neuron groups with limited spatial overlap

We first characterized the spatial relationships of the LVST and cMVST neuron groups by retrograde labeling and confocal microscopy. Differential retrograde labeling of the LVST and cMVST with conjugated dextran amines in the mouse embryo never led to double-labeled neurons (Figure 1), demonstrating that the LVST and cMVST neuron groups are hodologically distinct, as previously demonstrated in the chicken embryo (Diaz et al., 2003). However, there was some spatial overlap, in which about 20% of cMVST neurons were located within the domain of the LVST group (Figure 1). This highlights the need for an approach employing retrograde labeling and cell sorting to purify these neuron groups prior to transcriptomic characterization.

B) Global RNAseq profiles of LVST and cMVST neuron groups reveal differential transcript abundance across numerous genes, including those encoding transcription factors

Having determined the appropriate early stages for manual cell sorting following retrograde labeling (mouse E13.5, chicken d7.5 - see Materials and Methods) we performed RNAseq on samples from the LVST group, the cMVST group, and control groups positioned medially or caudally to the VS groups (Figure 2A). We selected these control samples to distinguish VS-enriched transcripts from those commonly expressed at similar anteroposterior (AP) or dorsoventral (DV) locations (Figure 2A). Unsupervised clustering algorithms, blinded to labeling strategy, generated groupings that ultimately aligned with the labeled hodological identity or their regional origins (caudal control and medial control) (Figure 2B-E). Differential expression analysis identified significantly differentially expressed transcripts across retrogradely labeled neuron and control sample groups (Figure 2F, G, Figure 3, Extended data 2-1, 2-2, 2-3).

Clustergrams of differentially expressed genes highlight several inverse patterns between groups (Figure 2F, G). The vestibulospinal groups as a whole show an inverse pattern compared to the medial control

eNeuro Accepted Manuscript

group, as expected from their different dorsolateral origins within the same rhombomeres (r4-5). The relationship to the caudal control group is more complex, likely because the caudal control covered multiple (three to four) rhombomeres and also comprised far fewer cells (manually collected) than the medial control (bulk tissue). Most importantly, some genes, including those encoding particular transcription factors (TFs), show an inverse pattern between the LVST and the cMVST, being exclusive to or enriched in one or the other (Figure 2F, G, Figure 3, Extended data 2-1, 2-2, 2-3).

A primary goal of our study was to identify candidate genes that might function as key regulators of VS neuron subtype identity, thus our main focus in further analysis of the LVST and cMVST neuron group transcriptomes was on TF-encoding genes. Nevertheless, we noted that several non-TF genes were highly differentially expressed between the LVST and cMVST neuron groups or between these and controls (Extended data 2-1). In the mouse, the LVST samples showed higher levels of *Islr2* (Linx, Immunoglobulin Superfamily Containing Leucine Rich Repeat 2), *Cdh22* (Cadherin 22), *Cbln1* and *Cbln4* (Cerebellin 1, and 4), and *Calb2* (Calretinin) transcripts as compared to the cMVST, while the cMVST sample showed greater abundance of *Sst (Somatostatin), Sema5a (Semaphorin 5a),* and *Slc32a1 (Vgat)* RNAs. Examples of transcripts enriched in both VS groups compared to control were *Tll1* (Tolloid Like 1 (metalloprotease)), *Ntng2* (Netrin G2), *Rbp1* (retinol Binding Protein 1), and *Cryba2* (Crystallin Beta A2 (Eye Lens Structural Protein)) (Extended data 2-1C). Some of these showed similar relative differences in transcript levels in chicken (Extended data 2-1D).

C) Conserved expression of common and exclusive VS group transcription factors

To identify candidate TFs for further investigation, we considered RNA abundance against control tissue (Figure 3, Extended data 2-2, 2-3), the absolute level of each particular TF-encoding transcript (Figure 3, Extended data 2-2, 2-3), the degree to which group-specific expression in mouse and chicken were conserved (Extended data 2-2, 2-3, 3-1), and by consulting the literature. Lhx1, Lhx5, Foxp2, Evx2, Onecut1, Onecut2, Onecut3, Esrrg, Maf, Phox2b, Lbx1, and Pou3f1 were all selected for further analysis by immunohistofluorescence, based on these criteria and the availability of functional antibodies. Evx1, Casz1,

eNeuro Accepted Manuscript

and Myc were also considered, but available antibodies stained most neurons in the hindbrain at embryonic stages, or had excessive background signal, and thus were not pursued further. Available antibodies against Evx2, Onecut2, Onecut3, and Pou3f1 were not functional in chicken embryonic tissue.

Immunostaining revealed TFs that were expressed at the protein level in the VS groups at the stages we examined (E13.5 and E15 in mouse embryo, and d7.5 and d9 in chicken embryo). Qualitative assessment of TF expression showed that LVST neurons were immunopositive for Phox2b, Lbx1, Esrrg and Maf, immunonegative for Lhx1+5, Evx2 and Foxp2, and that subpopulations were immunopositive for Onecut1/2/3 and Pou3f1 (Figure 4A). By contrast, cMVST neurons were immunopositive for Lhx1+5, Esrrg, Maf and Onecut1/2/3, immunonegative for Phox2b, and subpopulations were immunopositive for Lbx1 (mouse only), Pou3f1, Evx2, and Foxp2 (Figure 4A). Detailed quantification revealed consistency in this pattern across stage and species, but the absence of an Lbx1+ domain in the chicken cMVST, and fewer Esrrg+ cells in chicken versus mouse LVST, were notable exceptions (Figure 4B, C).

While most of the cell type-specific immunostaining patterns we examined correlated with the transcript-level differences identified by RNAseq (Table 2), one exception was Lbx1 expression. We did not detect *Lbx1* mRNA by RNAseq in the mouse cMVST, or in any of the chicken samples (Table 2), despite positive immunostaining for Lbx1 protein in a subpopulation of the mouse cMVST neurons (Figure 4A, B), and in nearly all of the chicken LVST neurons (Figure 4C). Interestingly, mouse *Lbx1* transcript levels were an order of magnitude lower in the LVST group (36 ± 30 transcripts per million (TPM)), compared to Phox2b (678 ± 126 TPM), Maf (495 ± 165 TPM), and Esrrg (753 ± 124 TPM), despite all of these TFs being positively immunostained in nearly all LVST neurons. Thus low transcript abundance, potentially outside the range of sensitivity of our RNA-seq approach, may nonetheless be sufficient for translation of the encoded protein as demonstrated in the case of *Lbx1* in some of our sorted neuron pools.

eNeuro Accepted Manuscript

D) A restricted, conserved TF signature that uniquely defines the LVST group within the r4 lineage

Immunostaining with the TF panel obtained for the LVST group showed that some TFs were expressed in all LVST neurons, whereas some were expressed in only part of the LVST group. In the mouse, Phox2b and Lbx1 were expressed in virtually all LVST neurons, Esrrg and Maf were expressed in 90-95% of LVST neurons, whereas Onecut 1, 2 and 3 and Pou3f1 were expressed in smaller subpopulations of LVST neurons (Figure 4 A, B). Results were similar in chicken, with Phox2b and Lbx1 expressed in virtually all LVST neurons and Maf in 85-95% of LVST neurons. However, Esrrg was expressed in only about 70-80% of LVST neurons, contrasting the 90-95% observed in mouse.

Spatial restriction of the 4-TF signature. Noting that 4 TFs (Phox2b/Lbx1/Maf/Esrrg) were expressed in most of the LVST neurons, we then asked how restricted is the expression of this signature in the developing CNS. We assessed this with quadruple combinatorial immunostaining in the E13.5 mouse. Since Phox2b, and consequently the 4-TF combination, is not expressed outside of the region spanning from the midbrain through the thoracic spinal cord (Pattyn et al., 1997), we first confirmed the cessation of Phox2b immunostaining at these levels and then limited our analysis to that range. We found that cells expressing the 4-TF combination were restricted to a short stretch of the dorsolateral hindbrain. Being limited by our available fluorophore and microscope capacity to simultaneously visualize four fluorophores in any given section, we stained for retrogradely labeled LVST neurons in alternating sections. This indicated that the 4-TF combination was restricted to the LVST group, plus a smaller domain just mediocaudal to the LVST group (Figure 5A).

The 4-TF signature cannot be reduced to a 3-TF signature. To determine whether any 3-TF combinations among the 4 had similar spatial restriction, we assessed each separately in serial sections spanning the midbrain, hindbrain and cervical and thoracic spinal cord. This showed that within the hindbrain, only one 3-TF combination (Maf/Esrrg/Phox2b) had the same spatial restriction as the 4-TF combination (compare Figure 5E to Figure 5A). Other 3-TF combinations were each expressed by additional specific outlying

eNeuro Accepted Manuscript

neuron groups within the hindbrain (Figure 5B-D). However, the Maf/Esrrg/Phox2b combination was also expressed by a group of neurons lying medially within the midbrain at the rostral end of the Phox2b CNS domain (Figure 5E). Thus, only the 4-TF signature Maf/Esrrg/Phox2b/Lbx1 appeared to be expressed exclusively by the LVST group and the small vicinal mediocaudal group.

Use of the 3-TF signature as a proxy for the 4-TF signature within the hindbrain. The equivalence within the hindbrain (excluding the midbrain) of the immunostaining pattern for the 3-TF combination Maf/Esrrg/Phox2b and the 4-TF signature allowed us to use the former as a proxy for the latter in subsequent experiments. We could then directly assess the expression of the 3-TF combination in serial sections in which the LVST group was also retrogradely labeled, instead of indirectly in adjacent sections as done above. This showed unequivocally that the 4-TF signature was restricted to the LVST group and the smaller vicinal mediocaudal neuron group at E13.5 (Figure 6A). We repeated the experiment in the E15.5 mouse hindbrain with identical results (Table 3). At E13.5, 86% and 90% of LVST neurons expressed this 4-TF combination in the two preparations assessed in this way (the remaining minority lacking either Maf or Esrrg, compare with Figure 4B).

The 4-TF signature is unique to the LVST within the r4 lineage. We have not yet determined the identity of the vicinal mediocaudal neuron group, except to confirm that it was not retrogradely labeled from the spinal cord (Figure 6A), nor have we identified alternative TF signatures that differentiate it from the LVST group. We noted, however, that a consistent feature of this neuron group was its relatively weak Maf immunostaining compared to the LVST group (Figure 6C, E). Its more caudal location also prompted us to assess whether the two groups might have different rhombomeric origins. To do this, we utilized *b1r4-Cre x tdTomato* mice (Di Bonito et al., 2013) in which Cre expression and activity is restricted to r4 by a *HoxB1* r4-enhancer element, and tdTomato expression is thereby restricted to cells descendant from r4 (see Methods and Materials). In these mice, we found that among those neurons expressing the 3-TF combination, only the LVST neuron group derived from r4-limited *HoxB1*-expressing progenitors at E14.5

eNeuro Accepted Manuscript

(Figure 6B-E). Thus, the 4-TF signature Maf/Esrrg/Phox2b/Lbx1 uniquely defines the LVST group within the r4 lineage.

The 4-TF signature emerges at an early stage. Having identified a lineage-specific TF signature that uniquely defines the LVST neuron group, we then asked how early during mouse hindbrain development this TF signature arises (Table 3). At E9.5, Phox2b immunostaining was evident in progenitors in the ventricular zone, but there were virtually no Maf, Esrrg or Lbx1 immunopositive cells in the hindbrain. At E11.5, we observed triple-positive Maf/Esrrg/Phox2b neurons (n=2 preparations) and Esrrg/Phox2b/Lbx1 neurons (n=2 preparations) overlapping with retrograde labeling of the nascent LVST group. We found that the Maf/Esrrg/Phox2b triple-positive neurons clustered in a single group in the hindbrain, and 20-50% of these were retrogradely labeled. E11.5 is during the period when LVST axons are growing towards the spinal cord, so some LVST neurons are expected to not be retrogradely labeled because their axons have not yet reached the tracer application site at C1. The percentage of Maf/Esrrg/Phox2b triple-positive neurons within the presumptive LVST domain that were retrogradely labeled from C1 increased to about 70% and 95% at E13.5 and E15.5, respectively.

E) TFs that define LVST neuron subpopulations

Next, we asked whether any of the TFs that were expressed by only a fraction of the LVST group were spatially compartmentalized within the group. We did this by generating 3D reconstructions of the immunostained LVST neurons (Extended data 4-1, 4-2). The examples described below of TFs with differential regional expression suggest a partitioning of LVST neurons into anatomical or functional subpopulations.

Essrg in chicken. As already noted, in the mouse, virtually all LVST neurons expressed Esrrg. By contrast, in the chicken, Esrrg was only expressed by 70-80% of LVST neurons. 3D reconstructions showed that Esrrg-

eNeuro Accepted Manuscript

451 negative neurons in the chicken formed a coherent subpopulation located dorsolaterally within the LVST
452 group (Extended data 4-2).

453 ***Pou3f1 in mouse*.** About 60% of LVST neurons were Pou3f1+ in the mouse. At E13.5 in the mouse there
454 was no obvious spatial pattern in the Pou3f1 expression pattern: Pou3f1+ and Pou3f1- LVST neurons were
455 intermingled. By E15.5, however, the Pou3f1+ subpopulation occupied the dorsolateral, rostral part of the
456 LVST (Extended data 4-1). In the chicken, *POU3F1* is listed as a pseudogene in the
457 Gallus_gallus.Gallus_gallus-5.0.93.gtf annotation file downloaded from Ensembl. In accordance with this
458 we could not detect POU3F1 by immunohistofluorescence (nor by RNAseq).

459 ***Onecut TFs*.** Similar proportions (about 40%) of mouse LVST neurons expressed Onecut1, 2 and 3, and
460 these were all located more ventrocaudally within the LVST (Extended data 4-1). In the chicken, for which
461 we had antibodies only against Onecut1, the picture was similar: about 60% of LVST neurons expressed
462 Onecut1, and these were located more caudally within the LVST (Extended data 4-2).

463 By immunostaining for Onecut1+2 or Onecut2+3 concurrently, we determined that only a minority of LVST
464 neurons in the mouse (around 5%) was single-positive for one of the two Onecut factors (Extended data 4-
465 5). From this we deduced that most LVST neurons express either all Onecut factors, or none of them.
466 Onecuts 1, 2 and 3 displayed a wide range of immunostaining intensities in mouse LVST neurons. By
467 measuring this within individual LVST neurons we determined that Onecut2 and Onecut3 immunostaining
468 intensity correlated highly (E13.5: $R^2 = 0.93 \pm 0.01$, E15.5: $R^2 = 0.88 \pm 0.06$), whereas Onecut1 and 2
469 correlated less (E13.5: $R^2 = 0.56 \pm 0.14$, E15.5: $R^2 = 0.52 \pm 0.14$). A characteristic feature of the mouse LVST
470 group at E15.5 was a rostral and very dorsal subpopulation of LVST neurons that consistently included only
471 a few Onecut1/2/3+ neurons, and essentially no Esrrg-negative or Pou3f1-negative neurons (Extended
472 data 4-1).

473

474 **F) TF combinations that distinguish r4- versus r5-derived cMVST neurons**

eNeuro Accepted Manuscript

In contrast to the LVST group, which derives exclusively from r4, the cMVST group derives from both r4 and r5, with the major part derived from r5. Our initial assessment of TF transcript expression indicated that virtually all cMVST neurons in the mouse express Lhx1/5 and Onecut 1, 2 and 3 at both E13.5 and E15.5 (Figure 4A, B). However, Onecut3 stained most cMVST neurons only very weakly, except in the rostral portion of the cMVST, where all Onecut factors exhibited stronger immunostaining than in the caudal portion (Extended data 4-6). Other TFs were expressed in fractions of the cMVST neuron population (in some cases in a stage-dependent manner). Large proportions of cMVST neurons expressed Esrrg (about 90%), Maf (80-95%), Pou3f1 (40-80%) and Evx2 (80-85%) transcripts. Smaller proportions expressed Lbx1 (about 20%) and Foxp2 (10-30%).

A similar picture was found in the chicken, with virtually all cMVST neurons expressing Lhx1/5 and Esrrg, a very large proportion expressing Maf or Onecut1 (about 90%), and a smaller proportion expressing Foxp2 (about 25%). In contrast to the mouse, essentially no cMVST neurons retrogradely labeled in the chicken at d7.5 or d9 expressed Lbx1 (Figure 4C).

The above observations suggested that some TFs are common to all cMVST neurons, whereas others might be differentially expressed in the minor r4- and major r5-derived portions of the cMVST. We tested this directly by combining immunohistofluorescence and retrograde labeling in *b1r4-Cre x tdTomato* mice, which we did at E16.5 to ensure substantial retrograde labeling of the cMVST. Lbx1 and Evx2 had regionally restricted immunostaining patterns within the cMVST with a strong relationship to rhombomeric origin: Lbx1+ cMVST neurons were exclusively derived from r4, and Evx2+ cMVST neurons were primarily derived from r5 (Figure 7A, Extended data 4-3). Only about 10% of r4-cMVST neurons expressed Evx2, whereas virtually all r5-cMVST neurons expressed Evx2 but not Lbx1 (Figure 7A).

These results reveal TF combinations that distinguish the r4- and r5-derived portions of the cMVST, with the presence or absence of Lbx1 expression providing a clear distinction of the 2 portions. Thus, the TF combination Lhx1+5/Onecut1,2,3/Lbx1 is restricted to the r4-derived portion of the cMVST, whereas the

eNeuro Accepted Manuscript

499 TF combination Lhx1+5/Onecut1,2,3/Evx2 in the absence of Lbx1 is restricted to the r5-derived part of the
500 cMVST.

501 Whether these TF combinations are similarly restricted to r4- and r5-cMVST neurons in the chicken was
502 not determined, due to a lack of key antibodies (such as Evx2) and to the relative difficulty in selectively
503 fate-mapping the r4 and r5 lineages in the chicken embryo, which requires additional approaches such as
504 the generation of quail-chicken chimeras (see Diaz et al. (1998)). However, we note that there may be
505 species-specific differences in cMVST TF profiles, since we found essentially no Lbx1+ cMVST neurons in
506 the chicken, despite this being a defining feature of the r4-derived portion of the cMVST in the mouse. On
507 the other hand, this could be due to immaturity of the r4-cMVST subgroup in the chicken at the stages
508 studied (see Discussion).

509

510 **G) TFs that define additional cMVST neuron subpopulations**.

511 cMVST neurons that were immunopositive for each of Maf, Esrrg, Foxp2 and Pou3f1 were also
512 differentially distributed, albeit less strikingly than for Evx2 and Lbx1, and they also exhibited some
513 dynamic changes (Extended data 4-3, 4-4).

514 ***Maf in mouse***. In the mouse, Maf-negative cMVST neurons (24%) were found mainly rostrally at E13.5,
515 suggesting that they derived from r4. However, by E15.5 there were almost no Maf-negative cMVST
516 neurons at all (5%), and these were found predominantly in the caudal region. By E16.5 about 95% of all
517 cMVST neurons were Maf+, with no bias towards either rhombomere (Figure 7A). Thus, Maf expression
518 may be restricted to the r5 lineage initially, but expands beyond that lineage during subsequent
519 development. We found no regional pattern of Maf expression in the chicken cMVST.

520 ***Esrrg in mouse***. For Esrrg, weakly immunopositive and immunonegative neurons were found mostly in the
521 rostral end of the cMVST group, intermingled with strongly immunopositive neurons. Costaining for Esrrg
522 and Evx2 (Evx2 being used as a proxy for r5 origin – see above) at E13.5 and E15.5 showed that about half

eNeuro Accepted Manuscript

of the Evx2-negative cMVST neurons were also Esrrg-negative (E13.5: 39% ± 24% SD, E15.5: 53% ± 5% SD.). By E16.5, however, about 90% of the r4-cMVST neurons were Esrrg+, although the immunostaining was weak in many of them (Figure 7). Fluorescence intensity measurements confirmed that Esrrg immunostaining was weaker in r4-cMVST neurons compared to r5-cMVST neurons (Figure 7B-E). We found no regional pattern of Esrrg expression in the chicken cMVST.

Foxp2. In the mouse, Foxp2+ neurons were rare in the rostral portion of the cMVST at E13.5 and E15.5, whereas the caudal portion contained intermingled Foxp2+ and Foxp2-negative cMVST neurons. The number of Foxp2+ cMVST neurons in the caudal portion of the cMVST declined noticeably from E13.5 to E15.5 (Extended data 4-3). In the chicken, we found a slight dorsolateral bias in the distribution of Foxp2+ cMVST neurons (Extended data 4-4).

Pou3f1 in mouse. In the mouse, Pou3f1 was expressed in both the rostral and caudal portions of the cMVST group in about equal proportions at E13.5, but at E15.5 there was a bias towards large numbers of Pou3f1+ neurons towards the mediocaudal pole of the cMVST group (Extended data 4-3). There was a precipitous drop in portion of Pou3f1+ cMVST neurons from E13.5 to E15.5 (Figure 4B).

H) A restricted, conserved TF signature that defines the r5-cMVST group

Spatial restriction of a 4-TF signature. Having identified TF combinations that distinguish the r4- and r5-derived subpopulations of the cMVST in the mouse, we then asked whether any TF combination represents a conserved TF signature that is restricted to either of these portions of the cMVST group. Here we focus on the r5-cMVST neurons.

Preliminary experiments assessing various 4-TF combinations in r5-cMVST neurons led us to the combination Lhx1+5/Evx2/Maf/Esrrg. Immunofluoresecence assessment from lumbar spinal cord to cortex in the E13.5 mouse showed that the only cells that expressed this TF combination were in the location of the cMVST group and extending medially from this group towards the midline (Figure 8, Table 4).

eNeuro Accepted Manuscript

We could not make the same assessment of the r5-cMVST TF signature in the chicken due to the lack of a functioning Evx2 antibody. Nevertheless, we had already determined that virtually all cMVST neurons in the chicken express Lhx1+5, Maf and Esrrg, as well as Onecut1. Together with the fact that Evx2 transcripts were detected in chicken cMVST RNA samples (Figure 3B, Table 2), and that no Lbx1 immunopositive neurons were detected in the chicken cMVST, it appears that the Lhx1+5/Evx2/Maf/Esrrg TF signature of r5-cMVST neurons is conserved in the two species.

The 4-TF signature cannot be reduced to a 3-TF signature, but one 3-TF signature can be used as a proxy.

To enable a direct demonstration through retrograde labeling of the expression of this TF signature by the r5-cMVST neurons, we then assessed the possible 3-TF combinations among Lhx1+5, Evx2, Maf and Esrrg to determine if any of these could be used as a proxy for the 4-TF signature. We found that each 3-TF combination was expressed by additional neuron groups within the hindbrain. However, the additional neuron groups expressing the Evx2/Maf/Esrrg combination were well separated from the cMVST group and the contiguous medially extending neuron population (Figure 8), allowing us to use this 3-TF combination as a proxy. We combined Evx2/Maf/Esrrg immunostaining with retrograde labeling of the cMVST at E15.5 to definitively assess overlap with the r5-cMVST neurons. This demonstrated that the contiguous medially extending neuron population was not retrogradely labeled and thus did not project to the spinal cord, whereas about 50% (38, 53 and 66% in 3 separate preparations) of neurons in the cMVST domain was retrogradely labeled (Figure 9). E15.5 is during the period when cMVST axons are growing towards the spinal cord, so some cMVST neurons are likely not retrogradely labeled because their axons have not yet reached C1. Most non-retrogradely labeled Evx2/Maf/Esrrg immunopositive neurons were positioned within the cMVST domain, intermingled with retrogradely labeled cMVST neurons (Figure 9). The proportion of r5-cMVST neurons expressing the Evx2/Maf/Esrrg signature was 89% (the remainder lacking either Maf or Esrrg, compare with Figure 4).

We have not yet determined the identity of the neurons that extend medially, nor have we identified alternative TF signatures that differentiate them from the cMVST neurons.

eNeuro Accepted Manuscript

572 ***Developmental appearance of the 4-TF signature***. Having established that Lhx1+5/Evx2/Maf/Esrrg
573 expression was restricted to the r5-cMVST neuron group and its contiguous, medially extending neighbor
574 group, we set out to assess how early this 4-TF signature appears during development. Quadruple-positive
575 r5-cMVST neurons were not present at E9.5 (since there is yet no Maf or Esrrg expression – see description
576 of LVST-related TF development), but 10 such neurons were seen in a preparation at E11.5 (Table 4), 8 of
577 which were located laterally (in the presumptive cMVST domain) and 2 of which were found medially,
578 close to the midline. By E13.5 the number of quadruple-positive neurons had increased in both lateral and
579 medial regions to levels seen in retrogradely labeled preparations at E15.5 (Table 4).

580

581 **I) In search of a TF signature for the r4-cMVST group**

582 Identifying a TF signature restricted to the r4-cMVST group posed more of a challenge since only 3 TFs
583 were robust markers of this subgroup: Lbx1, Lhx1+5 and Onecut1/2/3. As mentioned in Section F referring
584 to the E13.5 mouse, Esrrg was only detected in about 50% of r4-cMVST neurons, and at least some r4-
585 cMVST neurons are Maf-negative. We therefore assessed the 4-TF signature Lbx1/Lhx1+5/Maf/Onecut,
586 using Onecut2 as a proxy for Onecut1 and 3, since Onecut family members are usually co-expressed, and
587 our Onecut1 and 3 antibodies were of the same host species as our Lbx1 antibody. Co-immunostaining in
588 the E13.5 mouse hindbrain revealed that the majority of neurons expressing this 4-TF signature were not
589 r4-cMVST neurons, or even cMVST neurons at all (Figure 10A, B). Two distinct populations of neurons
590 were observed, with different rostrocaudal and dorsoventral locations. The largest and more caudal of
591 these was clearly non-vestibular, as it was located ventromedially (Figure 10A, B, orange triangles). The
592 more rostral population contained some r4-cMVST neurons, but the majority were non-cMVST neurons, as
593 they lay well outside the cMVST domain (Figure 10A, B, green circles). We then asked whether these
594 rostral non-cMVST neurons were derived from r4, by immunostaining for Lbx1/Lhx1+5/Onecut2 in the
595 *b1r4-Cre x tdTomato* mouse. If not, then this 4-TF signature might be unique to the r4-cMVST within the r4
596 lineage. Nearly all the Lbx1/Lhx1+5/Onecut2 positive cells in this area were, however, co-labeled for

eNeuro Accepted Manuscript

tdTomato, indicating that they derived from r4 (Figure 10C, D). We conclude that this 4-TF signature, although it labels some r4-cMVST neurons, is not unique to the r4-cMVST, even within r4. We therefore did not assess it in the rest of the CNS.

Discussion

A) Principal findings

We have previously described that the vestibulospinal system comprises three coherent groups of neurons with distinct anatomical locations, developmental origins, projection patterns and functional connectivity (J. C. Glover and Petursdottir, 1988, 1991; Diaz and Glover, 1996; Diaz et al., 1998; Auclair et al., 1999; J. C. Glover, 2000a; Pasqualetti et al., 2007; Kasumacic et al., 2010; Di Bonito et al., 2015; Lambert et al., 2016). Specifically, the two largest of these groups, the LVST and cMVST groups, have respectively dorsolateral versus dorsomedial locations, r4 versus r4+r5 origins, ipsilateral versus contralateral projections, and synaptic targets along the entire length of the spinal cord versus limited to the cervical spinal cord. Although previous studies have documented the expression of Phox2b and Lbx1 in LVST neurons (Schubert et al., 2001; Chen et al., 2012), a more complete assessment of TF profiles in VS neurons has been lacking. Here we use RNAseq to provide the first comprehensive post-mitotic transcriptomes of the LVST and cMVST neuron groups, in two species representing two different vertebrate classes. We further demonstrate protein expression of a number of TFs by immunohistofluorescence, and identify conserved sets of TFs that distinguish the LVST group from the cMVST group, and the r4-derived portion of the cMVST group from the r5-derived portion. Using combinatorial TF immunostaining and inspection along the entire length of the CNS, we identify two specific TF signatures that are restricted to respectively the LVST group and the r5-cMVST group, together with a smaller group of nearby neurons in each case. Finally, we identify additional TFs that define subpopulations within the LVST and cMVST neuron groups. Taken together, these results demonstrate a molecular underpinning that correlates with and likely contributes

eNeuro Accepted Manuscript

621 to the anatomical and functional characteristics that distinguish these brainstem-to-spinal cord projection
622 neuron groups.
623
624 **B) Methodological considerations**
625 Neurons acquire cell-specific phenotypes through transcriptional multi-step cascades during development,
626 and transcriptional control relies heavily on differential TF expression (Jessell, 2000; Hobert, 2008; Ooi and
627 Wood, 2008) (Kessaris et al., 2014) (Achim et al., 2014; Arendt et al., 2016). To ascertain which TFs are
628 expressed in nascent post-mitotic VS neurons, we employed RNAseq on manually sorted, retrogradely
629 labeled neurons, which ensured selective isolation of the relatively few neurons involved (Hempel et al.,
630 2007; Okaty et al., 2011). By investigating the VS transcriptome in both mouse and chicken, we could hone
631 in on phylogenetically conserved transcriptional signatures.
632
633 TFs that specify the identity of neuronal lineages are often rapidly downregulated in postmitotic neurons,
634 which typically begin to express new sets of cell-type-specific TFs (Achim et al., 2014). Here, we had to
635 retrogradely label postmitotic VS neurons to identify and isolate them. This restricted the time window in
636 which we could assess transcriptomic profiles. Chen et al. (2012), using a Hoxb1-GFP mouse, observed
637 putative LVST axons extending towards the spinal cord by E10.5. Here we found that the LVST can be
638 retrogradely labeled from the spinal cord by E11.5 and d4-d5 (HH24-26), and the cMVST by E13.5 and
639 d7.5-d8.5 (HH30-32), in the mouse and chicken embryo respectively. We could obtain sufficient numbers
640 of cMVST neurons only by retrogradely labeling from the mid-medulla at E13.5/d7.5 (see Materials and
641 Methods). To minimize sample variation within species due to developmental differences, we thus chose
642 E13.5 and d7.5 for harvesting both the LVST neurons (labeled from C1) and the cMVST neurons (labeled
643 from mid-medulla) neurons. Based on the above indications of the earliest time of axon outgrowth, and
644 presuming that axon outgrowth begins shortly after the neurons are born, our RNAseq data represents
645 expression at least 3 days after birth for the LVST neurons. Since the cMVST neurons appear to develop
646 about 2 days later than the LVST, we estimate that the RNAseq data represents expression at least 1 day

eNeuro Accepted Manuscript

after birth for the cMVST neurons. All of the TFs that make up the VS group-specific signatures were expressed at both E13.5/d7.5 and E15.5/d9, and as early as E11.5 in the mouse, indicating that the signatures are valid over a broad temporal range. The extent to which these TFs or others that we have captured by RNAseq are involved in lineage specification versus cell differentiation (or both) remains an open question.

C) Commonality and exclusivity in vestibulospinal TF signatures

Phox2b versus Lhx1+5 as discriminants of the LVST and cMVST groups

Phox2b and Lhx1+5 are mutually exclusive, discriminating the LVST from the cMVST. Interestingly, forced expression of *Phox2* genes downregulates Lhx1+5 in the spinal cord *in vivo* and also suppresses the growth of commissural axons (Hirsch et al., 2007; Pla et al., 2008), suggesting that repressive interactions between these two TFs may be involved in specifying the LVST from the cMVST.

Phox2b has been described as a pan-visceral homeodomain transcription factor, since most Phox2b-expressing neurons are associated with visceral function, e.g. branchial and visceral motoneurons (Pattyn et al., 2000), and first- and second-order visceral sensory neurons (D'Autreaux et al., 2011). In the adult rat, however, some neurons with no relationship to autonomic function express Phox2b (Kang et al., 2007). Our finding that all LVST neurons express Phox2b at embryonic stages is further evidence of *Phox2b* expression outside of traditionally defined visceral neurons, raising interesting questions about the developmental role of Phox2b, the definition of visceral function, and the evolution of the LVST phenotype.

Lhx1 and Lhx5 are expressed in a number of different cell types throughout the CNS. The antibody we and many other studies have used does not discriminate between the two, but evidence suggests that they are often co-expressed (Cepeda-Nieto et al., 2005; Moreno et al., 2005; Pillai et al., 2007). Lhx1+5 and other Lim homeodomain TFs have been shown to be essential for proper differentiation, axon guidance, and neurotransmitter phenotype of diverse neuron types, and Lhx1+5 expressed by some reticulospinal

eNeuro Accepted Manuscript

672 neurons (Kania et al., 2000; Cepeda-Nieto et al., 2005; Pillai et al., 2007; Zhao et al., 2007; Brohl et al., 2008;
673 Kohl et al., 2015), although a relationship between these and the cMVST group is not immediately obvious.
674
675 ***Lbx1 and Evx2 discriminate r4- and r5-derived VS neurons***
676 Lbx1 was expressed in all LVST neurons in both mouse and chicken and 90% of neurons in the r4-cMVST
677 group in the mouse. In the chicken, in which r4- versus r5-derived neurons could not be distinguished
678 because we did not include rhombomeric fate-mapping, we found no Lbx1+ cMVST neurons. Since r4-
679 derived neurons are known to exist in the chick (Diaz et al., 1998), we believe that this discrepancy relates
680 to the timing of their axon outgrowth to the spinal cord, that is, that we have not retrogradely labeled late
681 enough to detect them. Lbx1 is expressed by a diverse collection of CNS neurons, categorized by some
682 studies as having primarily somatic functions, including somatosensory association interneurons in the
683 dorsal spinal cord (Gross et al., 2002) and somatosensory relay neurons in the trigeminal nucleus (Sieber et
684 al., 2007).
685 The r5-cMVST group did not express Lbx1, but Evx2. Evx2 is expressed postmitotically in V0 interneurons in
686 the spinal cord and in more extensive neuron populations in the hindbrain and midbrain (Moran-Rivard et
687 al., 2001; Inamata and Shirasaki, 2014). Evx2-expressing neurons are primarily, but not exclusively,
688 commissural (Lu et al., 2015), which is generally consistent with the cMVST phenotype.
689
690 ***Maf and Esrrg are VS-common TFs***
691 Among the TFs we studied, Maf and Esrrg are expressed by the vast majority of both LVST and cMVST
692 neurons. These TFs are therefore pivotal for distinguishing VS neurons from other neurons that express
693 combinations of Phox2b, Lbx1, Evx2 and Lhx1+5. Exceptions were the somewhat lower proportions of
694 Esrrg+ neurons in the dorsolateral part of the LVST group in chicken and in the r4-cMVST group in mouse.
695 Both Maf and Esrrg play critical roles in specifying post-mitotic cells types, both within and outside the
696 nervous system (Maf: (Wende et al., 2012; Chuan and Zhi-Min, 2015) Esrrg: (Susens et al., 2000; Friese et
697 al., 2009; Patthey et al., 2016; Yoshihara et al., 2016)). Maf belongs to the activated protein-1 superfamily

eNeuro Accepted Manuscript

698 (Chuan and Zhi-Min, 2015). Its paralog MafB is selectively expressed in r5 and r6 in mice and zebrafish,
699 highlighting the relationship of Maf genes to hindbrain development (Moens and Prince, 2002; Giudicelli et
700 al., 2003). Esrrg is a constitutive active nuclear hormone receptor, with no known physiological activating
701 ligand (Huppunen et al., 2004). Expression of Esrrg and its invertebrate homolog is also rhombomere-
702 specific in developing zebrafish and amphioxus (Bardet et al., 2005).
703
704 **D) CNS-wide restriction of identified VS TF signatures**
705 Increasing efforts are being made to identify TF signatures for functionally identifiable neuron groups
706 within the motor system. Recent evidence suggests a much greater diversity in the functional organization
707 of brainstem-to-spinal cord projections than was previously recognized (reviewed in (Perreault and Glover,
708 2013; Brownstone and Chopek, 2018), but to date there have been few attempts to codify this diversity in
709 terms of TF signatures (Cepeda-Nieto et al., 2005; Chen et al., 2012; Bretzner and Brownstone, 2013). One
710 important question that is not often addressed is whether TF signatures are restricted to a neuron group of
711 interest or also expressed elsewhere in the CNS.
712 Through CNS-wide immunofluorescence mapping in the mouse, we were able to assess the degree to
713 which the TF signatures we identified are specific to the LVST and cMVST neuron groups. The combination
714 of Phox2b/Lbx1/Maf/Esrrg was restricted to the LVST group and a smaller group of neurons that was
715 slightly caudal to the LVST group, had weaker levels of Maf expression, and did not originate from r4. The
716 combination of Evx2/Lhx1+5/Maf/Esrrg was restricted to the r5-cMVST group and to more medial neurons
717 at the same rostrocaudal level. In neither case were the vicinal non-LVST or non-r5-cMVST neurons
718 retrogradely labeled from the spinal cord at the developmental stages examined. Whether this means they
719 never project to the spinal cord, or are in fact nascent bulbospinal neurons whose axons have not yet
720 reached the spinal cord, remains to be determined. Despite the fact that they also are expressed by
721 additional neurons near the LVST and cMVST groups, it is notable that these TF signatures are otherwise
722 unique within the CNS at the stages examined.

eNeuro Accepted Manuscript

In contrast to the LVST and r5-cMVST neuron groups, we have not yet found a TF signature that is similarly restricted to the r4-cMVST group. The 3-TF signature that we found that provides full coverage of the r4-cMVST group, Lbx1/Lhx1+5/Onecut2, was expressed by many other cell groups within the CNS. Even when assessed only within the r4-domain, many non-cMVST neurons within r4 expressed this signature. Further work is required to determine whether additional TFs distinguish the LVST and cMVST from the smaller neighboring non-VST neurons. Nevertheless, the degree of restriction exhibited makes these 4-TF signatures useful markers for the groups and should facilitate further investigations into vestibulospinal development and function.

E) Proposed progenitor origins and evolutionary implications

The differential expression of TF signatures by the LVST and cMVST are related to rhombomeric origin, and thus to rhombomere-specific TFs involved in anteroposterior-patterning. Indeed, Hoxb1 is a critical factor in specifying the LVST and r4-cMVST phenotypes; in the absence of Hoxb1 expression these groups do not develop (Di Bonito et al., 2015). Which r5-related Hox TF(s) is(are) similarly critical for the specification of the r5-cMVST group remains to be determined, but is likely to be among Hoxa3/b3/d3 (Keynes and Krumlauf, 1994).

Although largely composed of postmitotic TFs, the signatures indicate a composite origin of the vestibulospinal system with respect to dorsoventral neural progenitor domains. Expression of Phox2b and Lbx1 in the hindbrain is specific for the dB2 progenitor domain (Hernandez-Miranda et al., 2017) indicating that this is the origin of the LVST group, as previously proposed by Chen et al. (2012). The dB2 progenitor domain is unique to the hindbrain, suggesting that there is no LVST neuron group homolog elsewhere. By contrast, the expression of Lbx1 and Lhx1+5 in the r4-cMVST neurons suggests a possible origin from either the pdB1, the pdB4 or the late pdBLa progenitor domains (Hernandez-Miranda et al., 2017). Currently unpublished work from our laboratories indicates that none of the cMVST neurons derive from Ascl1-expressing progenitors (Glover and Dymecki (unpublished observations), suggesting an origin from the pdB4 domain.

eNeuro Accepted Manuscript

749 The r5-cMVST group TF signature suggests yet another DV origin, given the inclusion of Evx2 and the
750 exclusion of Lbx1. Evx2 is indicative of a p0 progenitor origin, which would place the r5-cMVST in the
751 unusual position of deriving from a ventral as opposed to a dorsal progenitor domain.
752 That these diverse lineages converge on a vestibulospinal phenotype raises interesting questions about the
753 evolution of the vestibulospinal system. Give that they derive from progenitor domains that are found
754 throughout the hindbrain-spinal cord axis, it would seem that the r4- and r5-cMVST groups represent a
755 situation in which spinal interneuron populations have been co-opted and repurposed for vestibulospinal
756 function. Since the cMVST is primarily involved in vestibulo-collic reflexes (that is, reflexive movements of
757 the head about the neck), this co-option would most likely have occurred first with the advent of a
758 movable neck, that is, later than fish and amphibians. By contrast, the LVST group is among vestibulospinal
759 neurons uniquely related to Phox2b, and projects to the entire length of the spinal cord in both limbed and
760 non-limbed vertebrates. Given that Phox2 genes predate the chordates, and that balance-related control
761 of both trunk and limb musculature is important throughout the vertebrate radiation, this could mean that
762 the LVST is the primordial vestibulospinal projection, which has been supplemented by the cMVST (and the
763 iMVST) later in evolution.
764
765 **F) VS group subpopulations defined by additional TFs**
766 The cMVST group is composed of 2 subgroups based on rhombomeric origin. Both the LVST group and the
767 cMVST group are likely to exhibit further heterogeneity, given that they target various spinal segments and
768 neuron subtypes (Wilson and Yoshida, 1969; Uchino and Kushiro, 2011; Kasumacic et al., 2015), and
769 receive various types of vestibular afferent input (Boyle and Pompeiano, 1981; Uchino and Kushiro, 2011).
770 Expression of TFs that define subpopulations within these groups was therefore expected.
771 We found that Lbx1, Evx2, Onecut1/2/3, Foxp2, Pou3f1, and to a certain degree Maf and Esrrg, all defined
772 subpopulations of neurons within at least one VS group. How these expression patterns relate to
773 functional differences among VS neurons remains to be determined. Some predictions can be made,
774 however. Several studies have shown in mammals that respectively rostral versus caudal LVST neurons

eNeuro Accepted Manuscript

target cervical versus lumbar spinal levels (Brodal, 1963; Shamboul, 1980; Esposito et al., 2014). This could be related to our observation that Pouf31 is expressed in more rostral LVST neurons, whereas Onecut1/2/3 are expressed in more caudal LVST neurons. Similarly, Foxp2 (which was absent in the LVST) was expressed by more neurons in the caudal and dorsolateral portions, respectively, of the mouse and chicken cMVST neuron group.

We note that such intrinsic patterning within VS neuron groups could be explained by transitory developmental states, rather than permanent functional subdivisions, although some examples (such as the rostrocaudally biased Onecut and Foxp2 expression patterns) are stable during the developmental period we have studied. Clearly, further work is needed to determine the relevance of these TFs for functional diversity within the VS groups.

G) Significance of the work

The molecular mechanisms that specify hodological and anatomical subdivisions within brainstem-to-spinal cord projection neurons are poorly understood. Assessment of TF expression profiles provides information essential for addressing this question. Here we show that TF signatures conserved between mammals and birds can be defined for specific, functionally identifiable VS projection neuron groups. This lends support to the idea that descending systems for communication between the brainstem and spinal cord are set up by a genetic blueprint established early in vertebrate evolution, in this case at least 300 million years ago, when the avian and mammalian lineages diverged.

The identification of specific TF signatures provides opportunities in several directions, including to ascertain evolutionary relationships and changes, to unravel functional heterogeneity, to facilitate molecular manipulations, to elucidate molecular programs of differentiation and ascertain terminal selector genes, and to generate specific types of vestibulospinal neurons from stem or progenitor cells *in vitro* for research and medical purposes.

eNeuro Accepted Manuscript

800 **REFERENCES**

801 Achim K, Salminen M, Partanen J (2014) Mechanisms regulating GABAergic neuron development. Cellular
802 and molecular life sciences : CMLS 71(8): 1395-1415.
803 Akaike T (1983) Neuronal organization of the vestibulospinal system in the cat. Brain Research 259(2): 217-
804 227.
805 Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, Pavlicev M, Schlosser G, Widder S,
806 Laubichler MD, Wagner GP (2016) The origin and evolution of cell types. Nature reviews Genetics
807 17(12): 744-757.
808 Auclair F, Marchand R, Glover JC (1999) Regional patterning of reticulospinal and vestibulospinal neurons
809 in the hindbrain of mouse and rat embryos. The Journal of Comparative Neurology 411(2): 288-300.
810 Bardet PL, Schubert M, Horard B, Holland LZ, Laudet V, Holland ND, Vanacker JM (2005) Expression of
811 estrogen-receptor related receptors in amphioxus and zebrafish: implications for the evolution of
812 posterior brain segmentation at the invertebrate-to-vertebrate transition. Evolution &
813 development 7(3): 223-233.
814 Boyle R, Pompeiano O (1981) Convergence and interaction of neck and macular vestibular inputs on
815 vestibulospinal neurons. Journal of neurophysiology 45(5): 852-868.
816 Bretzner F, Brownstone RM (2013) Lhx3-Chx10 reticulospinal neurons in locomotor circuits. The Journal of
817 neuroscience : the official journal of the Society for Neuroscience 33(37): 14681-14692.
818 Brodal A (1963) ANATOMICAL OBSERVATIONS ON THE VESTIBULAR NUCLEI, WITH SPECIAL REFERENCE TO
819 THEIR RELATIONS TO THE SPINAL CORD AND THE CEREBELLUM. Acta oto-laryngologica
820 Supplementum 192: Suppl 192:124+.
821 Brohl D, Strehle M, Wende H, Hori K, Bormuth I, Nave KA, Muller T, Birchmeier C (2008) A transcriptional
822 network coordinately determines transmitter and peptidergic fate in the dorsal spinal cord. Dev
823 Biol 322(2): 381-393.
824 Brownstone RM, Chopek JW (2018) Reticulospinal Systems for Tuning Motor Commands. Frontiers in
825 Neural Circuits 12: 30.

eNeuro Accepted Manuscript

826 Cepeda-Nieto AC, Pfaff SL, Varela-Echavarria A (2005) Homeodomain transcription factors in the
827 development of subsets of hindbrain reticulospinal neurons. Molecular and cellular neurosciences
828 28(1): 30-41.
829 Chen Y, Takano-Maruyama M, Fritzsch B, Gaufo GO (2012) Hoxb1 controls anteroposterior identity of
830 vestibular projection neurons. PloS one 7(4): e34762.
831 Chuan Z, Zhi-Min G (2015) Multiple functions of Maf in the regulation of cellular development and
832 differentiation. Diabetes/Metabolism Research and Reviews 31(8): 773-778.
833 D'Autreaux F, Coppola E, Hirsch MR, Birchmeier C, Brunet JF (2011) Homeoprotein Phox2b commands a
834 somatic-to-visceral switch in cranial sensory pathways. Proceedings of the National Academy of
835 Sciences of the United States of America 108(50): 20018-20023.
836 Di Bonito M, Boulland JL, Krezel W, Setti E, Studer M, Glover JC (2015) Loss of Projections, Functional
837 Compensation, and Residual Deficits in the Mammalian Vestibulospinal System of Hoxb1-Deficient
838 Mice. eNeuro 2(6).
839 Di Bonito M, Narita Y, Avallone B, Sequino L, Mancuso M, Andolfi G, Franze AM, Puelles L, Rijli FM, Studer
840 M (2013) Assembly of the auditory circuitry by a Hox genetic network in the mouse brainstem.
841 PLoS Genet 9(2): e1003249.
842 Diaz C, Glover JC (1996) Appropriate pattern formation following regulative regeneration in the hindbrain
843 neural tube. Development (Cambridge, England) 122(10): 3095-3105.
844 Diaz C, Glover JC, Puelles L, Bjaalie JG (2003) The relationship between hodological and cytoarchitectonic
845 organization in the vestibular complex of the 11-day chicken embryo. J Comp Neurol 457(1): 87-
846 105.
847 Diaz C, Puelles L, Marin F, Glover JC (1998) The relationship between rhombomeres and vestibular neuron
848 populations as assessed in quail-chicken chimeras. Dev Biol 202(1): 14-28.
849 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR:
850 ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) 29(1): 15-21.

eNeuro Accepted Manuscript

851 Donevan AH, Fleming FL, Rose PK (1992) Morphology of single vestibulospinal collaterals in the upper
852 cervical spinal cord of the cat: I. Collaterals originating from axons in the ventromedial funiculus
853 contralateral to their cells of origin. J Comp Neurol 322(3): 325-342.
854 Espana A, Clotman F (2012) Onecut transcription factors are required for the second phase of
855 development of the A13 dopaminergic nucleus in the mouse. J Comp Neurol 520(7): 1424-1441.
856 Esposito MS, Capelli P, Arber S (2014) Brainstem nucleus MdV mediates skilled forelimb motor tasks.
857 Nature 508(7496): 351-356.
858 Friese A, Kaltschmidt JA, Ladle DR, Sigrist M, Jessell TM, Arber S (2009) Gamma and alpha motor neurons
859 distinguished by expression of transcription factor Err3. Proceedings of the National Academy of
860 Sciences of the United States of America 106(32): 13588-13593.
861 Giudicelli F, Gilardi-Hebenstreit P, Mechta-Grigoriou F, Poquet C, Charnay P (2003) Novel activities of Mafb
862 underlie its dual role in hindbrain segmentation and regional specification. Dev Biol 253(1): 150-
863 162.
864 Glover J (1995) Retrograde and anterograde axonal tracing with fluorescent dextran-amines in the
865 embryonic nervous system. Neurosci Prot 30: 1-13.
866 Glover JC (2000a) Development of specific connectivity between premotor neurons and motoneurons in
867 the brain stem and spinal cord. Physiological reviews 80(2): 615-647.
868 Glover JC (2000b) Neuroepithelial 'compartments' and the specification of vestibular projections. Progress
869 in brain research 124: 3-21.
870 Glover JC, Petursdottir G (1988) Pathway specificity of reticulospinal and vestibulospinal projections in the
871 11-day chicken embryo. J Comp Neurol 270(1): 25-38, 60-21.
872 Glover JC, Petursdottir G (1991) Regional specificity of developing reticulospinal, vestibulospinal, and
873 vestibulo-ocular projections in the chicken embryo. Journal of neurobiology 22(4): 353-376.
874 Grant GR, Farkas MH, Pizarro AD, Lahens NF, Schug J, Brunk BP, Stoeckert CJ, Hogenesch JB, Pierce EA
875 (2011) Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper
876 (RUM). Bioinformatics (Oxford, England) 27(18): 2518-2528.

eNeuro Accepted Manuscript

877 Grillner S, Hongo T, Lund S (1970) The vestibulospinal tract. Effects on alpha-motoneurones in the
878 lumbosacral spinal cord in the cat. Experimental brain research 10(1): 94-120.
879 Gross MK, Dottori M, Goulding M (2002) Lbx1 specifies somatosensory association interneurons in the
880 dorsal spinal cord. Neuron 34(4): 535-549.
881 Hamburger V, Hamilton HL (1992) A series of normal stages in the development of the chick embryo. 1951.
882 Developmental dynamics : an official publication of the American Association of Anatomists 195(4):
883 231-272.
884 Hempel CM, Sugino K, Nelson SB (2007) A manual method for the purification of fluorescently labeled
885 neurons from the mammalian brain. Nat Protocols 2(11): 2924-2929.
886 Hernandez-Miranda LR, Müller T, Birchmeier C (2017) The dorsal spinal cord and hindbrain: From
887 developmental mechanisms to functional circuits. Developmental Biology 432(1): 34-42.
888 Hirsch M-R, Glover JC, Dufour HD, Brunet J-F, Goridis C (2007) Forced expression of Phox2 homeodomain
889 transcription factors induces a branchio-visceromotor axonal phenotype. Developmental Biology
890 303(2): 687-702.
891 Hobert O (2008) Gene regulation by transcription factors and microRNAs. Science 319(5871): 1785-1786.
892 Huppunen J, Wohlfahrt G, Aarnisalo P (2004) Requirements for transcriptional regulation by the orphan
893 nuclear receptor ERRgamma. Molecular and cellular endocrinology 219(1-2): 151-160.
894 Inamata Y, Shirasaki R (2014) Dbx1 triggers crucial molecular programs required for midline crossing by
895 midbrain commissural axons. Development (Cambridge, England) 141(6): 1260-1271.
896 Jessell TM (2000) Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat
897 Rev Genet 1(1): 20-29.
898 Kang BJ, Chang DA, Mackay DD, West GH, Moreira TS, Takakura AC, Gwilt JM, Guyenet PG, Stornetta RL
899 (2007) Central nervous system distribution of the transcription factor Phox2b in the adult rat. J
900 Comp Neurol 503(5): 627-641.
901 Kania A, Johnson RL, Jessell TM (2000) Coordinate roles for LIM homeobox genes in directing the
902 dorsoventral trajectory of motor axons in the vertebrate limb. Cell 102(2): 161-173.

eNeuro Accepted Manuscript

903 Kasumacic N, Glover JC, Perreault MC (2010) Segmental patterns of vestibular-mediated synaptic inputs to
904 axial and limb motoneurons in the neonatal mouse assessed by optical recording. J Physiol 588(Pt
905 24): 4905-4925.
906 Kasumacic N, Lambert FM, Coulon P, Bras H, Vinay L, Perreault MC, Glover JC (2015) Segmental
907 organization of vestibulospinal inputs to spinal interneurons mediating crossed activation of
908 thoracolumbar motoneurons in the neonatal mouse. The Journal of neuroscience : the official
909 journal of the Society for Neuroscience 35(21): 8158-8169.
910 Kessaris N, Magno L, Rubin AN, Oliveira MG (2014) Genetic programs controlling cortical interneuron fate.
911 Current opinion in neurobiology 26: 79-87.
912 Keynes R, Krumlauf R (1994) Hox genes and regionalization of the nervous system. Annu Rev Neurosci 17:
913 109-132.
914 Kohl A, Marquardt T, Klar A, Sela-Donenfeld D (2015) Control of axon guidance and neurotransmitter
915 phenotype of dB1 hindbrain interneurons by Lim-HD code. The Journal of neuroscience : the
916 official journal of the Society for Neuroscience 35(6): 2596-2611.
917 Lambert FM, Bras H, Cardoit L, Vinay L, Coulon P, Glover JC (2016) Early postnatal maturation in
918 vestibulospinal pathways involved in neck and forelimb motor control. Developmental
919 neurobiology 76(10): 1061-1077.
920 Lu DC, Niu T, Alaynick WA (2015) Molecular and cellular development of spinal cord locomotor circuitry.
921 Frontiers in Molecular Neuroscience 8(25).
922 Moens CB, Prince VE (2002) Constructing the hindbrain: insights from the zebrafish. Developmental
923 dynamics : an official publication of the American Association of Anatomists 224(1): 1-17.
924 Moran-Rivard L, Kagawa T, Saueressig H, Gross MK, Burrill J, Goulding M (2001) Evx1 Is a Postmitotic
925 Determinant of V0 Interneuron Identity in the Spinal Cord. Neuron 29(2): 385-399.
926 Moreno N, Bachy I, Retaux S, Gonzalez A (2005) LIM-homeodomain genes as territory markers in the
927 brainstem of adult and developing Xenopus laevis. J Comp Neurol 485(3): 240-254.

eNeuro Accepted Manuscript

928 Muller T, Brohmann H, Pierani A, Heppenstall PA, Lewin GR, Jessell TM, Birchmeier C (2002) The
929 homeodomain factor lbx1 distinguishes two major programs of neuronal differentiation in the
930 dorsal spinal cord. Neuron 34(4): 551-562.
931 Murray AJ, Croce K, Belton T, Akay T, Jessell TM (2018) Balance Control Mediated by Vestibular Circuits
932 Directing Limb Extension or Antagonist Muscle Co-activation. Cell reports 22(5): 1325-1338.
933 Okaty BW, Sugino K, Nelson SB (2011) A quantitative comparison of cell-type-specific microarray gene
934 expression profiling methods in the mouse brain. PloS one 6(1): e16493.
935 Ooi L, Wood IC (2008) Regulation of gene expression in the nervous system. The Biochemical journal
936 414(3): 327-341.
937 Pasqualetti M, Diaz C, Renaud JS, Rijli FM, Glover JC (2007) Fate-mapping the mammalian hindbrain:
938 segmental origins of vestibular projection neurons assessed using rhombomere-specific Hoxa2
939 enhancer elements in the mouse embryo. The Journal of neuroscience : the official journal of the
940 Society for Neuroscience 27(36): 9670-9681.
941 Patthey C, Clifford H, Haerty W, Ponting CP, Shimeld SM, Begbie J (2016) Identification of molecular
942 signatures specific for distinct cranial sensory ganglia in the developing chick. Neural development
943 11: 3.
944 Pattyn A, Hirsch M, Goridis C, Brunet JF (2000) Control of hindbrain motor neuron differentiation by the
945 homeobox gene Phox2b. Development (Cambridge, England) 127(7): 1349-1358.
946 Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF (1997) Expression and interactions of the two closely
947 related homeobox genes Phox2a and Phox2b during neurogenesis. Development (Cambridge,
948 England) 124(20): 4065-4075.
949 Perreault MC, Glover JC (2013) Glutamatergic reticulospinal neurons in the mouse: developmental origins,
950 axon projections, and functional connectivity. Annals of the New York Academy of Sciences 1279:
951 80-89.
952 Peterson BW, Coulter JD (1977) A new long spinal projection from the vestibular nuclei in the cat. Brain Res
953 122(2): 351-356.

eNeuro Accepted Manuscript

954 Peterson BW, Maunz RA, Fukushima K (1978) Properties of a new vestibulospinal projection, the caudal
955 vestibulospinal tract. Experimental brain research 32(2): 287-292.
956 Pierreux CE, Vanhorenbeeck V, Jacquemin P, Lemaigre FP, Rousseau GG (2004) The transcription factor
957 hepatocyte nuclear factor-6/Onecut-1 controls the expression of its paralog Onecut-3 in
958 developing mouse endoderm. The Journal of biological chemistry 279(49): 51298-51304.
959 Pillai A, Mansouri A, Behringer R, Westphal H, Goulding M (2007) Lhx1 and Lhx5 maintain the inhibitory-
960 neurotransmitter status of interneurons in the dorsal spinal cord. Development (Cambridge,
961 England) 134(2): 357-366.
962 Pla P, Hirsch MR, Le Crom S, Reiprich S, Harley VR, Goridis C (2008) Identification of Phox2b-regulated
963 genes by expression profiling of cranial motoneuron precursors. Neural development 3: 14.
964 Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression
965 analysis of digital gene expression data. Bioinformatics (Oxford, England) 26(1): 139-140.
966 Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat
967 Methods 9(7): 671-675.
968 Schubert FR, Dietrich S, Mootoosamy RC, Chapman SC, Lumsden A (2001) Lbx1 marks a subset of
969 interneurons in chick hindbrain and spinal cord. Mechanisms of development 101(1-2): 181-185.
970 Shamboul KM (1980) Lumbosacral predominance of vestibulospinal fibre projection in the rat. J Comp
971 Neurol 192(3): 519-530.
972 Shinoda Y, Sugiuchi Y, Izawa Y, Hata Y (2006) Long descending motor tract axons and their control of neck
973 and axial muscles. Progress in brain research 151: 527-563.
974 Sieber MA, Storm R, Martinez-de-la-Torre M, Muller T, Wende H, Reuter K, Vasyutina E, Birchmeier C (2007)
975 Lbx1 acts as a selector gene in the fate determination of somatosensory and viscerosensory relay
976 neurons in the hindbrain. The Journal of neuroscience : the official journal of the Society for
977 Neuroscience 27(18): 4902-4909.
978 Straka H, Baker R, Gilland E (2001) Rhombomeric organization of vestibular pathways in larval frogs.
979 437(1): 42-55.

eNeuro Accepted Manuscript

980 Studer M, Popperl H, Marshall H, Kuroiwa A, Krumlauf R (1994) Role of a conserved retinoic acid response
981 element in rhombomere restriction of Hoxb-1. Science 265(5179): 1728-1732.
982 Susens U, Hermans-Borgmeyer I, Borgmeyer U (2000) Alternative splicing and expression of the mouse
983 estrogen receptor-related receptor gamma. Biochemical and biophysical research communications
984 267(2): 532-535.
985 Suwa H, Gilland E, Baker R (1996) Segmental Organization of Vestibular and Reticular Projections to Spinal
986 and Oculomotor Nuclei in the Zebrafish and Goldfish. The Biological bulletin 191(2): 257-259.
987 Tsuchida T, Ensini M, Morton SB, Baldassare M, Edlund T, Jessell TM, Pfaff SL (1994) Topographic
988 organization of embryonic motor neurons defined by expression of LIM homeobox genes. Cell
989 79(6): 957-970.
990 Uchino Y, Kushiro K (2011) Differences between otolith- and semicircular canal-activated neural circuitry in
991 the vestibular system. Neuroscience research 71(4): 315-327.
992 Wende H, Lechner SG, Birchmeier C (2012) The transcription factor c-Maf in sensory neuron development.
993 Transcription 3(6): 285-289.
994 Wilson VJ, Yoshida M (1969) Comparison of effects of stimulation of Deiters' nucleus and medial
995 longitudinal fasciculus on neck, forelimb, and hindlimb motoneurons. Journal of neurophysiology
996 32(5): 743-758.
997 Yoshihara E, Wei Z, Lin CS, Fang S, Ahmadian M, Kida Y, Tseng T, Dai Y, Yu RT, Liddle C, Atkins AR, Downes
998 M, Evans RM (2016) ERRgamma Is Required for the Metabolic Maturation of Therapeutically
999 Functional Glucose-Responsive beta Cells. Cell metabolism 23(4): 622-634.
1000 Zhao Y, Kwan KM, Mailloux CM, Lee WK, Grinberg A, Wurst W, Behringer RR, Westphal H (2007) LIM-
1001 homeodomain proteins Lhx1 and Lhx5, and their cofactor Ldb1, control Purkinje cell differentiation
1002 in the developing cerebellum. Proceedings of the National Academy of Sciences of the United
1003 States of America 104(32): 13182-13186.

1004

eNeuro Accepted Manuscript

1005

1006

eNeuro Accepted Manuscript

Figures and tables legends

Table 1. Primary antibodies used. *Antibodies that stained most cells in the hindbrain, or had excessive background signal. Abbreviations: AA, amino acids; DSHB, Developmental Studies Hybridoma Bank; N/A, not available; RRID, Research Resource Identifiers.

Table 2. Immunostaining patterns and RNA levels of transcription factors in vestibulospinal neurons. Mean TPM values (± SD) of TFs validated by immunohistofluorescence, showing a strong correlation of VS group RNA levels and immunostaining. A value of "0" indicates no detection by RNAseq. N/A = not assessed by immunohistofluorescence due to lack of appropriate antibodies.

Table 3. Developmental appearance of the 3-TF proxy for the LVST neuron group signature. Means (± SD) of Maf/Esrrg/Phox2b triple-positive hindbrain neurons at different locations (with indicated co-expression of either Lbx1 and/or Hoxb1) at different developmental stages in the mouse embryo. N/A = not assessed.

Table 4. Number of Lhx1+5/Evx2/Maf/Esrrg quadruple-positive neurons at different developmental stages in the mouse. Counts of quadruple positive cells ± standard deviation in different areas of the mouse CNS. *Counts represent entire CNS. Only the hindbrain was assessed at other stages. **Preparations at E15.5 were not co-immunostained for Lhx1+5 because they included retrograde labeling of the cMVST; here, counts represent Evx2/Maf/Esrrg immunostained neurons in and around the cMVST domain only (other brainstem areas were not assess). See figure 9.

Figure 1. **Spatial relationship between LVST and cMVST neuron groups**. Differential retrograde labeling in the E13.5 mouse of LVST with RDA (pseudocolored magenta, applied unilaterally at C1), and cMVST with BDA/FDA (pseudocolored green, applied contralaterally at mid-medulla; see inset in (B). Structures visible in the green channel have been superpositioned on structures visible in the magenta channel for clarity. (A) Overlay of 17 serial transverse sections spanning the rostrocaudal extent of the LVST/cMVST groups. Dotted white outlines show the section perimeter (outer) and the principal domains of the LVST and cMVST groups. LVST axons (arrow) course medioventrally to the LVST group. In this preparation, only 26 of

eNeuro Accepted Manuscript

44 cMVST neurons that appear to be within the domain of the LVST group in the image projection are in fact intermingled with LVST neurons; the remainder actually lies outside the LVST group domain. (B) Overlay of 9 frontal sections spanning the dorsoventral extent of the LVST/cMVST groups. LVST axons (arrow) course mediocaudally to the LVST group. Arrowheads indicate cMVST axons crossing the midline. In the inset, the dashed box shows the location of the full image, and magenta and green indicate dextran amine injection sites. VIII = eighth cranial nerve entry site. Scale bar 200 µm. Extended data 1-1: Retrograde labeling from mid-medulla does not label outside of the cMVST group.

Figure 2. Mouse and chicken vestibulospinal neuron groups cluster separately based on whole transcriptome gene expression profiles. (A) Schematic of E13.5 mouse or d7.5 chicken hindbrain showing anatomical locations of RDA injections, control lesions, and the retrogradely labeled VS neuron groups. LVST and cMVST samples were dissociated and individual fluorescent cells sorted. Caudal control samples consisted of sorted non-fluorescent cells from a region immediately caudal to the LVST group (only collected in the mouse). Medial control samples consisted of tissue pieces at the same anteroposterior level as the VS neuron groups, laterally delimited by the retrogradely labeled LVST and cMVST groups. (B, C) Hierarchical clustering plots showing correlation distance, and incorporating all genes that pass filter (based on cpm > 1 in at least three samples, one isoform per gene), the data are log2(Transcripts per million (TPM) +1), distance metric is 1-Pearson correlation, linkage method is 'ward'. (D, E) Multidimensional scaling plots using the top 2000 highest variance genes (mouse, D), or top 500 highest variance genes (chicken, E), performed in edgeR. (F, G) Bi-clustering of differentially expressed genes (FDR < .0005 ANOVA-like analysis and log2 fold difference > 1 between at least two sample groups) across samples in mouse (F) and chicken (G), performed in Matlab using the clustergram function. Boxes labeled I and II highlight genes that are more abundant, respectively, in LVST than in cMVST samples, and vice versa. The thick black lines separate genes with an overall inverse expression pattern between the medial control group and the vestibulospinal groups as a whole. Extended data 2-1: RNA levels and fold changes for all transcripts in LVST versus cMVST, normalized to control tissue. Extended data 2-2 (.xlsx): TPM values and

eNeuro Accepted Manuscript

ANOVA-like analysis for transcription factors from mouse RNAseq data. Extended data 2-3 (.xlsx): TPM values and ANOVA-like analysis for transcription factors from chicken RNAseq data.

Figure 3. RNA levels and fold changes for transcription factors in LVST versus cMVST neurons, normalized to control tissue. Top row (A) shows mouse and bottom row (B) shows chicken data. (A,B) Left panels show names of genes with significant differential expression (FDR < .01, ANOVA-like analysis) and high fold change between groups. Gene names of those TFs subsequently validated by immunohistofluorescence indicated in red (except for chicken ONECUT2, ONECUT3 and EVX2, for which antibodies were only functional in mouse). Right panels show only the latter gene names. Chicken genes prefixed with ENSGALG and leading zeros have been shortened to digits only. Immunohistofluorescence for the TFs marked with an asterisk (Evx1*, Casz1* and Myc*) stained most hindbrain cells above background level, and these genes were therefore not analyzed further. The x- and y-axes show respectively increasing levels of expression in the LVST and cMVST groups relative to control. The red diagonal indicates unity, and the color scale represents log2(TPM+1) for cMVST for points above the red diagonal, and for LVST below the red diagonal, i.e. the VS group with the highest TPM value for each gene determines the point color. Genes above the upper blue diagonal have a >4-fold increase of TPM+1 values for cMVST versus LVST, and those below the lower blue diagonal have a >4-fold increase of TPM+1 values for LVST versus cMVST. For mouse n = 2529 TFs, for chicken n = 620 TFs. Extended data 3-1: RNA level comparison of identified mouse/chicken TF orthologs.

Figure 4. Immunohistofluorescent staining for transcription factors in LVST and cMVST neurons. A) Confocal images of transverse sections through the level of the VS neurons in an E15.5 mouse. Retrogradely labeled LVST neurons (top row) and cMVST neurons (bottom row) shown in green, and immunostaining for TFs in colors as indicated below each column. Overlapping LVST and Lhx1+5 pixels (yellow) in the LVST/Lhx1+5 panel do not co-localize in the z-axis, and are thus false positives (see Methods and Materials). Scalebar 100 µm. Dorsal up, lateral left. (B, C) Percentage of retrogradely labeled LVST or cMVST neurons immunopositive for the indicated TFs in E13.5 and E15.5 mouse embryos (B) and d7.5 and d9 chicken embryos (C). Error bars represent standard deviations.

eNeuro Accepted Manuscript

Extended data 4-1: Histograms and scatterplots showing the spatial distribution of mouse LVST neurons expressing the indicated TFs

Extended data 4-2: Histograms and scatterplots showing the spatial distribution of chicken LVST neurons expressing the indicated TFs.

Extended data 4-3: Histograms and scatterplots showing the spatial distribution of mouse cMVST neurons expressing the indicated TFs.

Extended data 4-4: Histograms and scatterplots showing the spatial distribution of chicken cMVST neurons expressing the indicated TFs

Extended data 4-5: Percentage of mouse LVST neurons immunopositive for a single Onecut transcription factor.

Extended data 4-6: Differential Onecut TF staining intensity in the cMVST neuron group at different rostro-caudal levels.

Figure 5. Spatial restriction of the Maf/Esrrg/Phox2b/Lbx1 LVST neuron group signature compared to different 3-way combinations thereof. (A-D) Each panel shows a 3D reconstruction of the E13.5 mouse hindbrain with the neurons (n = number) co-expressing the indicated set of TFs shown as green circles. Neurons are plotted from every third serial transverse section. (A) Most neurons expressing Maf/Esrrg/Phox2b/Lbx1 lie within the LVST neuron group domain (by comparison with adjacent serial sections with retrograde labeling), but a few lie in a more medioventral domain (arrow indicates this smaller, mediolateral vicinal neuron group). (B-D) Neurons expressing the indicated 3-TF signatures are found in the LVST group domain and in additional domains within the hindbrain. (E) 3D reconstruction of the E13.5 mouse hindbrain and midbrain, with neurons co-expressing Maf, Esrrg and Phox2b indicated as green circles within the hindbrain and blue triangles within the midbrain. Neurons expressing this 3-TF signature were found in the domains of the LVST group, the smaller vicinal mediocaudal neuron group (arrow) and a neuron group in the midbrain near the midline (blue triangles). Neurons are plotted from every sixth serial transverse section.

eNeuro Accepted Manuscript

Figure 6. The LVST neuron group and the vicinal mediocaudal neuron group sharing the LVST TF signature derive from different rhombomeric lineages. (A) Expression of Maf/Esrrg/Phox2b confirmed in LVST neurons by direct combination with retrograde BDA labeling in the E13.5 mouse hindbrain. Maf/Esrrg/Phox2b triple-positive neurons were restricted to two clusters, the large LVST neuron group and the smaller, vicinal mediocaudal neuron group, seen in frontal (left) and lateral (right) views. Red circles = Maf+ Esrrg+ Phox2b+ BDA+ neurons (n=306), blue triangles = Maf+ Esrrg+ Phox2b+ BDA- neurons (n=59). Cells from every 3rd section in this preparation counted. Dashed line in A indicates the midline. (B,C,D,E) Transverse sections from an E14.5 b1r4-Cre x tdTomato mouse, triple immunostained for Maf, Esrrg and Phox2b, at the level of the vicinal mediocaudal neuron group (B, C) and the LVST neuron group (D, E). C,E show separate confocal images of each fluorescent channel from within the dashed boxes shown in B and D, with red dots indicating cells co-labeled for Maf/Esrrg/Phox2b. tdTomato staining is absent in the vicinal mediocaudal neuron group (C), indicating its origin from outside of r4. The anti-Maf signal is not shown in B or D for clarity. (B,C,D,E) Dorsal up, lateral left, midline right.

Figure 7. cMVST TFs are differentially expressed in r4- and r5-derived cMVST neurons. (A) Percentage of r4-derived (Hoxb1+) versus r5-derived (Hoxb1-) cMVST neurons immunostained for Lbx1, Evx2, Esrrg or Maf (n=2 for Maf, n=3 for others). Lbx1 and Evx2 immunostaining discriminates r4-cMVST and r5-cMVST neurons. (B) Histogram of Esrrg fluorescence intensity (in normalized arbitrary units, a.u.) in individual r4-cMVST neurons (white bars), r5-cMVST neurons (light grey bars) and Esrrg-negative cells (dark grey bar, comprising nearby non-immunostained nuclei), assessed in n=7, 8 and 9 sections from each of n=3 different hindbrains. Note that r4-cMVST neurons are substantially fewer than r5-cMVST neurons, as expected from the relative sizes of the two portions of the cMVST. (C) Empirical cumulative probability distributions of Esrrg fluorescence intensity in the same cMVST neurons and Essrg-negative cells as in (B). Note that using cumulative probability distributions in effect normalizes the numbers of r4- and r5-cMVST neurons. The statistical difference between the two distributions was tested using the two-sample Kolmogorov-Smirnov test, with p-value as indicated. (D) Confocal image of cMVST neurons (green), r4-derived cells (magenta) and dual positive r4-cMVST neurons (white). (E) Same field of view as in (D), with

eNeuro Accepted Manuscript

Esrrg-immunostained nuclei (white) and Hoechst stained nuclei (red outlines). In D and E, vertical arrows indicate r4-cMVST neurons (mostly weak Esrrg staining), horizontal arrows r5-cMVST neurons (mostly strong Esrrg staining). All quantification was done on sections from E16.5 *b1r4-Cre x tdTomato* mice. Ventral left, lateral down. Optical slice thickness 2 µm. Scale bar 20 µm.

Figure 8. Expression of the Lhx1+5/Evx2/Maf/Esrrg TF combination in the E13.5 mouse is restricted to two clusters of neurons. Only two neuron groups in the mouse CNS co-expressed Lhx1+5/Evx2/Maf/Esrrg, indicated by green circles (cMVST group) and blue triangles (medially extending neighboring neuron population). Retrogradely labeled LVST neurons (magenta triangles) are plotted from alternating serial sections from the same preparation for orientation. The other neuron groups shown (black circles, red circles, turquoise triangles) are Evx2/Maf/Esrrg triple-positive but Lhx1+5-negative. Outline in right panel drawn from the level of the cMVST group. Scale bar left: 500 µm, right: 200 µm.

Figure 9. Overlap of retrograde labeling and Evx2/Maf/Esrrg-immunostaining in the region of the cMVST group. (A, B) Locations of laterally located Evx2/Maf/Esrrg-expressing cells in a E15.5 mouse of overlapping with contralateral retrograde labeling (white circles, n=87) or without (green circles, n=41). Medially located Evx2/Maf/Esrrg-expressing neurons with no retrograde labeling shown as blue triangles (n=42). Neurons from every 3rd section throughout the rostro caudal extent of the cMVST shown. A) Transverse section through the level of the cMVST outlined. B) Side view of the rostro caudal extent of the cMVST.

Figure 10. **Hindbrain cells expressing Lbx1/Lhx1+5/Onecut2/Maf are not restricted to the r4-cMVST group.** A) 3D map and B) transverse projection map of cells co-expressing Lbx1/Lhx1+5/Maf/Onecut2 in the E13.5 mouse hindbrain, with retrogradely labeled LVST neurons (white circles) plotted from adjacent interleaved sections for orientation. Green circles and orange triangles represent Lbx1/Lhx1+5/Maf/Onecut2+ cells located respectively at the same rostrocaudal level but predominantly ventral to the cMVST and LVST, and at a more caudal and ventromedial location. The dashed line roughly delineates the location of vestibulospinal neurons (LVST + cMVST), and contains only a few quadruple-immunolabeled cells. Cells are plotted from every sixth serial transverse section; in B these are projected

eNeuro Accepted Manuscript

1158 onto a single plane. C) 3D map and D) transverse projection map of cells co-expressing
1159 Lbx1/Lhx1+5/Onecut2/tdTomato in the E14.5 *b1r4-Cre x tdTomato* mouse hindbrain, with retrogradely
1160 labeled cMVST neurons (white triangles) plotted from adjacent interleaved sections for orientation. Green
1161 circles represent Lbx1/Lhx1+5/Onecut2/tdTomato+ cells. The 3D reconstruction in A is limited
1162 rostrocaudally to the level of the quadruple-immunolabeled cells and the cMVST neurons. The dashed line
1163 roughly delineates the location of vestibulospinal neurons (LVST + cMVST), and contains 11 quadruple-
1164 immunolabeled cells that are likely to be r4-cMVST neurons. Typically 20% of the cMVST group as a whole
1165 (45 neurons in this reconstruction) are r4-cMVST neurons. This corresponds well with the counts here (45
1166 * 0.2 = 9). Cells are plotted from every fourth serial transverse section; in D these are projected onto a
1167 single plane.

1168 **Extended data legends:**

1169 **Figure 1-1. Retrograde labeling from mid-medulla does not label outside of the cMVST group.** cMVST
1170 group in the d11 chicken embryo retrogradely labeled from C1 (A), and mid-medulla (B). cMVST group in
1171 the P1 mouse retrogradely labeled from C1 (C), and mid-medulla (D). Scale bar 200 µm.

1172 **Figure 2-1. RNA levels and fold changes for all transcripts in LVST versus cMVST, normalized to control**
1173 **tissue.** All detected transcripts in mouse (A) and chicken (B) from RNAseq data (13844 mouse and 11982
1174 chicken genes). A, B) Legend as in Figure 3. C) Transcript levels of select highly differentially expressed non-
1175 TF genes in individual mouse RNAseq samples. D) Transcript levels of same genes as in (C) for individual
1176 chicken RNAseq samples.

1177 **Figure 2-2 (.xlsx). TPM values and ANOVA-like analysis for transcription factors from mouse RNAseq data.**
1178 Values calculated with ANOVA-like differential abundance analysis performed in in edgeR. Transcription
1179 factor inclusion criteria based Gene Ontology class 6355 (regulation of transcription, DNA templated) and
1180 counts per million > 1 in at least three samples. Columns legend: A) assigned gene index, B) gene name, C)
1181 RefSeq ID, D) transcript length in nucleotides, E) likelihood ratio statistics, F) two-sided p-value, G) false
1182 discovery rate adjusted p-value, H) LVST mean TPM value, I) cMVST mean TPM value, J) mean caudal

eNeuro Accepted Manuscript

control sample TPM value, K) mean medial control sample TPM value, L-Q) individual LVST sample TPM values, R-T) individual cMVST sample TPM values, U-X) individual caudal control sample TPM values, Y-AB) individual medial control sample TPM values.

Figure 2-3 (.xlsx). TPM values and ANOVA-like analysis for transcription factors from chicken RNAseq data. Values calculated with ANOVA-like differential abundance analysis performed in in edgeR. Transcription factor inclusion criteria based Gene Ontology class 6355 (regulation of transcription, DNA templated) and counts per million > 1 in at least three samples. Columns legend: A) assigned gene index, B) gene name, C) ensembl ID, D) transcript length in nucleotides, E) likelihood ratio statistics, F) two-sided p-value, G) false discovery rate adjusted p-value, H) LVST mean TPM value, I) cMVST mean TPM value, J) mean medial control sample TPM value, K-N) individual LVST sample TPM values, O-R) individual cMVST sample TPM values, S-V) individual medial control sample TPM values.

Figure 3-1. RNA level comparison of identified mouse/chicken TF orthologs. Plots show identified TFs that shared orthologues in mouse and chicken, plotted for Log2(TPM + 1) values, with mouse values on the x-axis and chicken values on the y-axis, for the LVST groups (A) and cMVST groups (B). Mouse-distinct TFs are defined as having Log2(TPM + 1) value >4 in mouse, and <1 in chicken, and vice versa. TF gene names (only mouse names used) shown in red indicate TFs whose expression was validated by immunohistofluorescence. Diagonal line shows the linear regression line with slope, intercept and r2 values shown at top left in each plot.

Figure 4-1. Histograms and scatterplots showing the spatial distribution of mouse LVST neurons expressing the indicated TFs. Neurons are indicated as either immunopositive (blue) or immunonegative (red). Histograms show averages and standard deviations along the indicated axes. Scatterplots show projections in the indicated planes from single, representative preparations for each TF and stage.

Figure 4-2. Histograms and scatterplots showing the spatial distribution of chicken LVST neurons expressing the indicated TFs. Legend as Figure 4-1.

eNeuro Accepted Manuscript

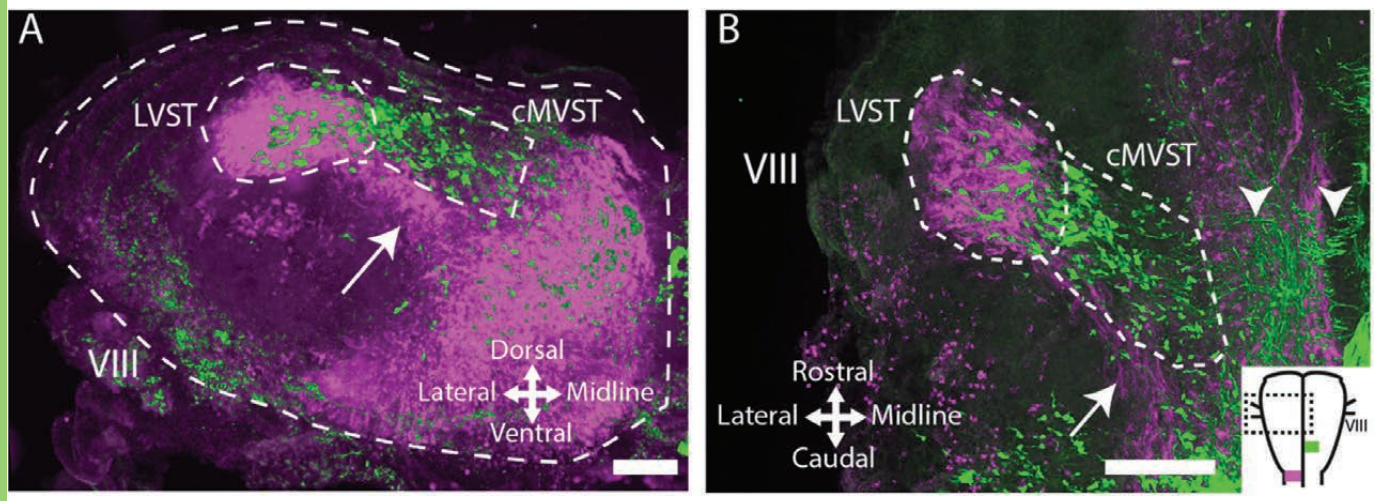
Figure 4-3. Histograms and scatterplots showing the spatial distribution of mouse cMVST neurons expressing the indicated TFs. Legend as figure 4-1.

Figure 4-4. Histograms and scatterplots showing the spatial distribution of chicken cMVST neurons expressing the indicated TFs. Legend as figure 4-1.

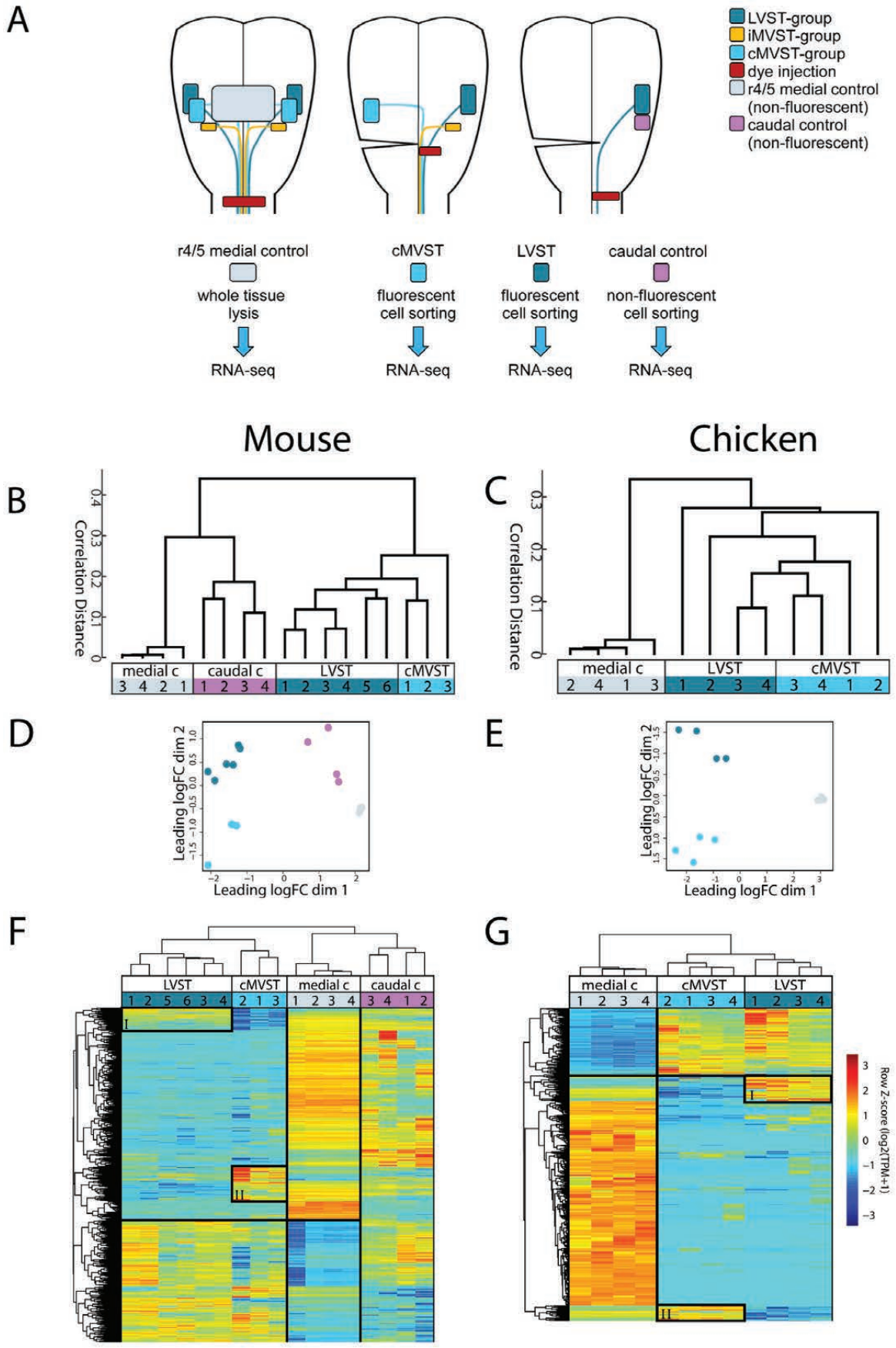
Figure 4-5. Percentage of mouse LVST neurons immunopositive for a single Onecut transcription factor. Percentage ± standard deviation of mouse LVST neurons at indicated stages that were singly immunopositive for one (of two) Onecut factors. Abbreviation: Oc – Onecut.

Figure 4-6. Differential Onecut TF staining intensity in the cMVST neuron group at different rostro-caudal levels. (A-L) Confocal images of transverse sections through the E13.5 mouse cMVST, retrogradely labeled with BDA from the mid-medulla. Immunostained with antibodies specific for the indicated Onecut TFs. Imaging settings and contrast levels preserved between different rostro-caudal levels. Red dots in A,B,C,G,H,I indicate the locations of corresponding cMVST neurons in D,E,F,J,K,L,M. A-F are from the mid rostrocaudal level of the cMVST group, and G-L are from the rostral level of the cMVST group. Immunostaining for Onecut TFs is stronger in rostral cMVST neurons. Dorsal up, lateral left, medial right. Scale bar 50 μm. Abbreviations: Oc, Onecut.

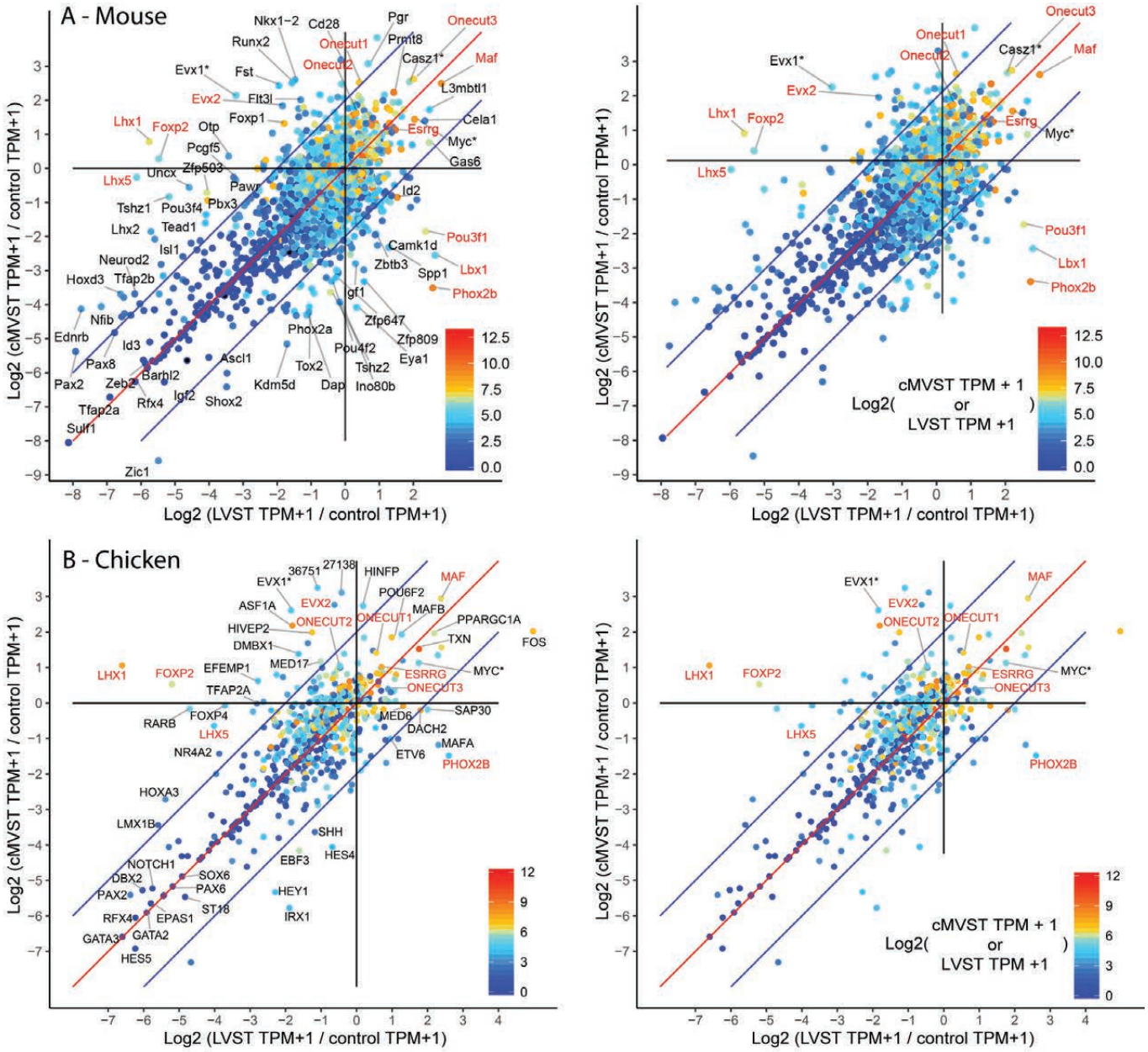
eNeuro Accepted Manuscript



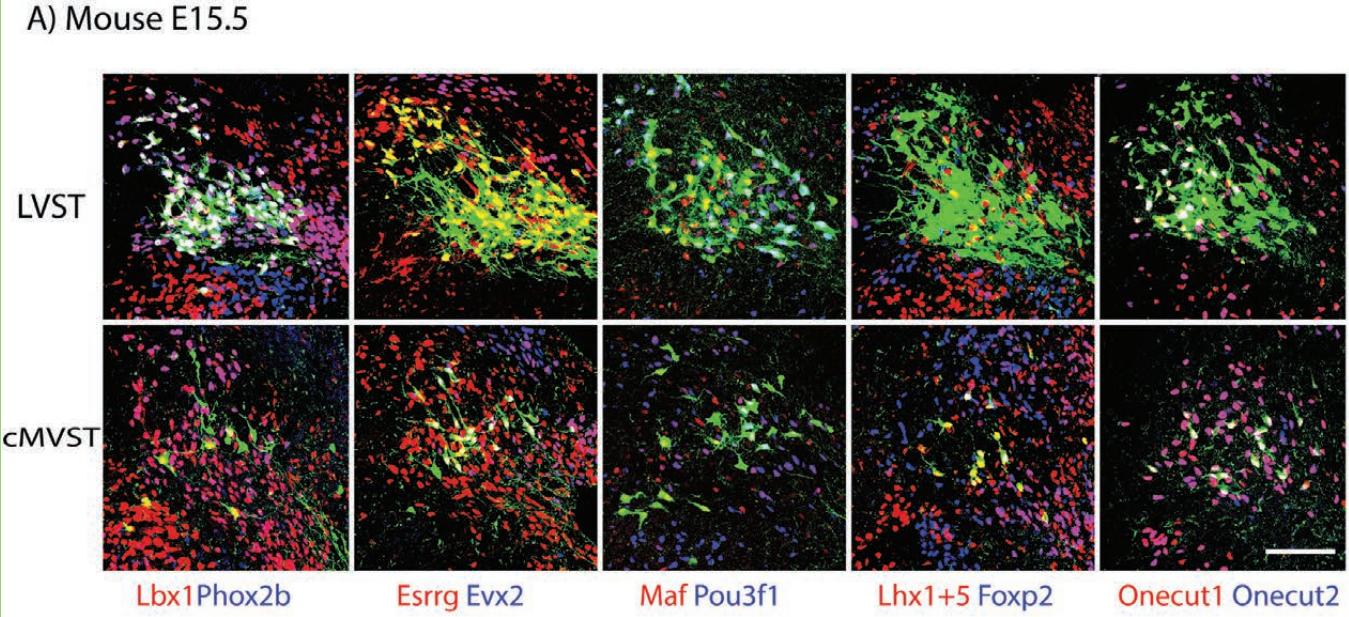
eNeuro Accepted Manuscript

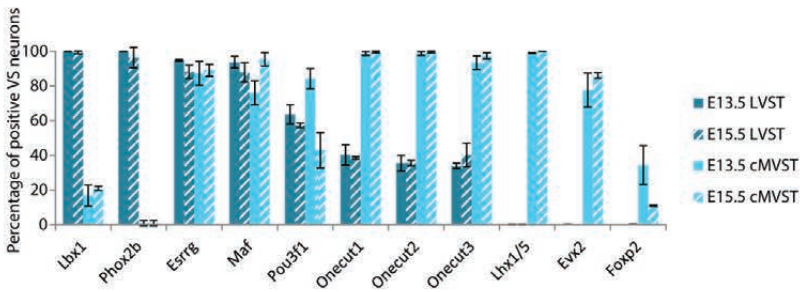


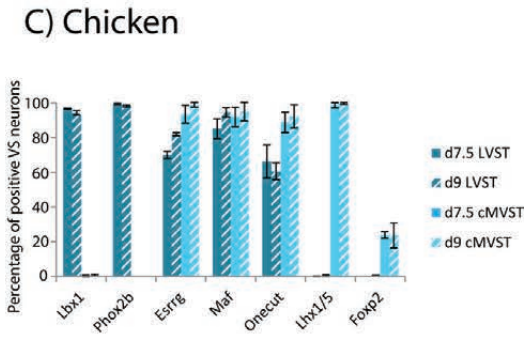
eNeuro Accepted Manuscript



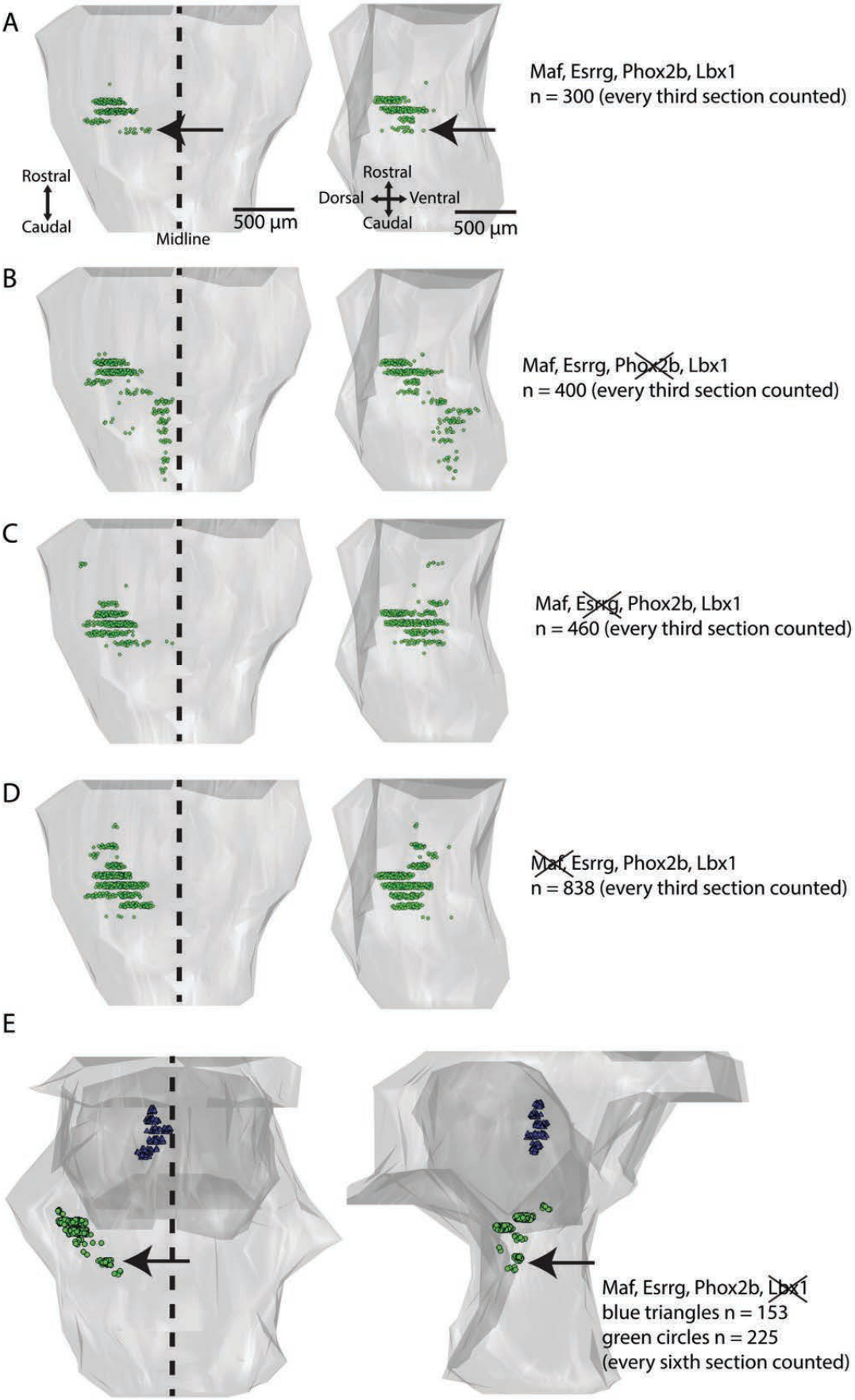
eNeuro Accepted Manuscript



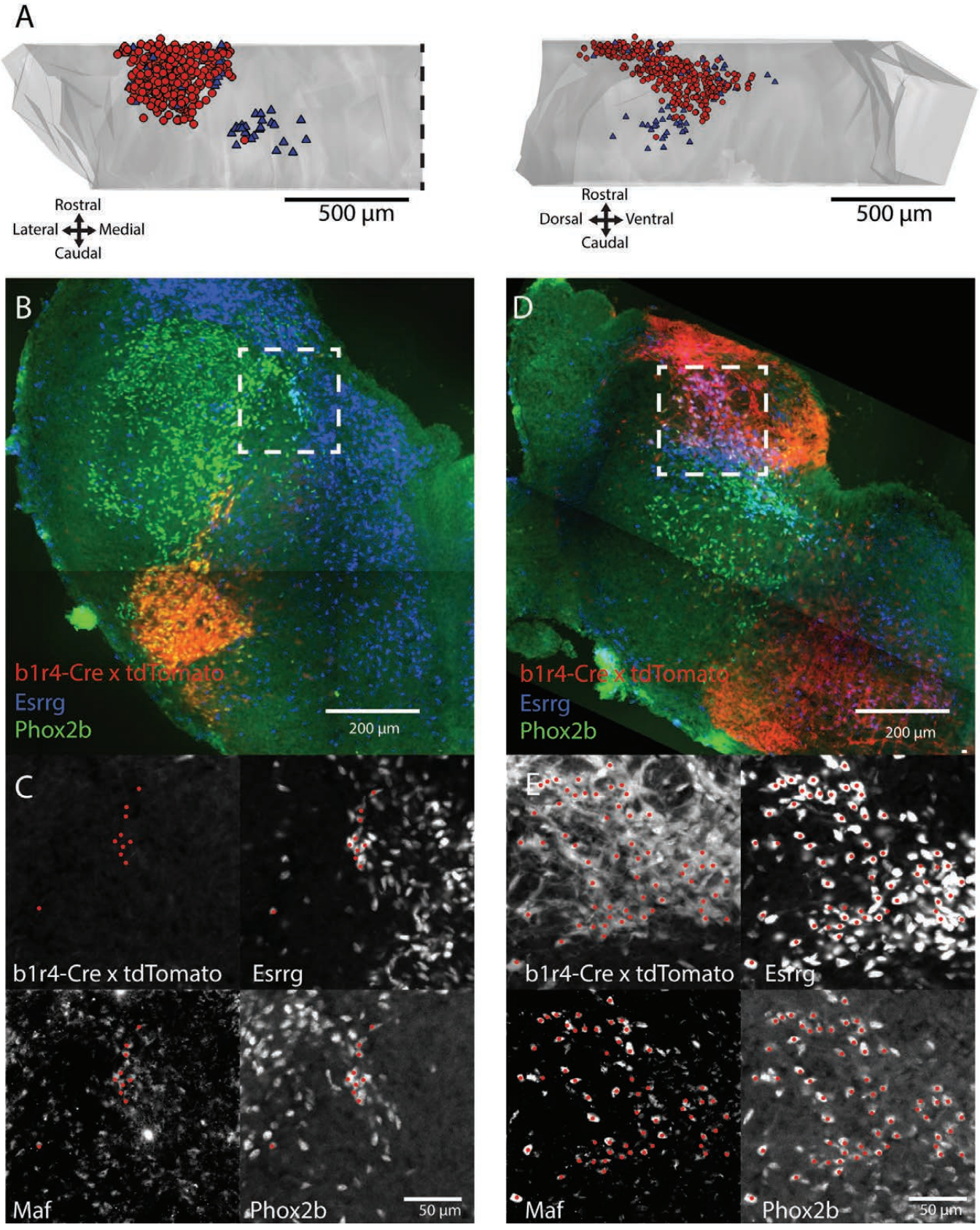




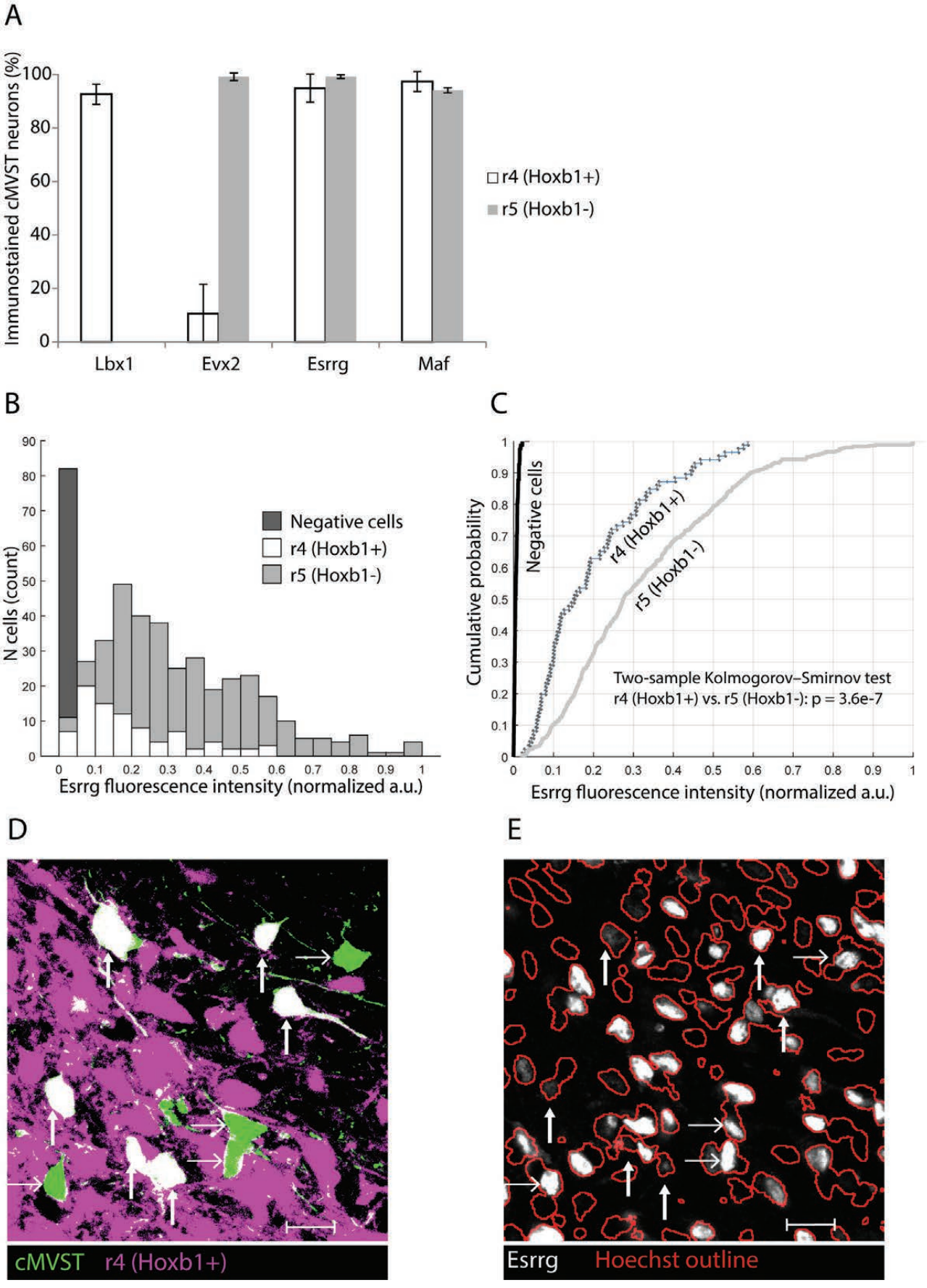
eNeuro Accepted Manuscript



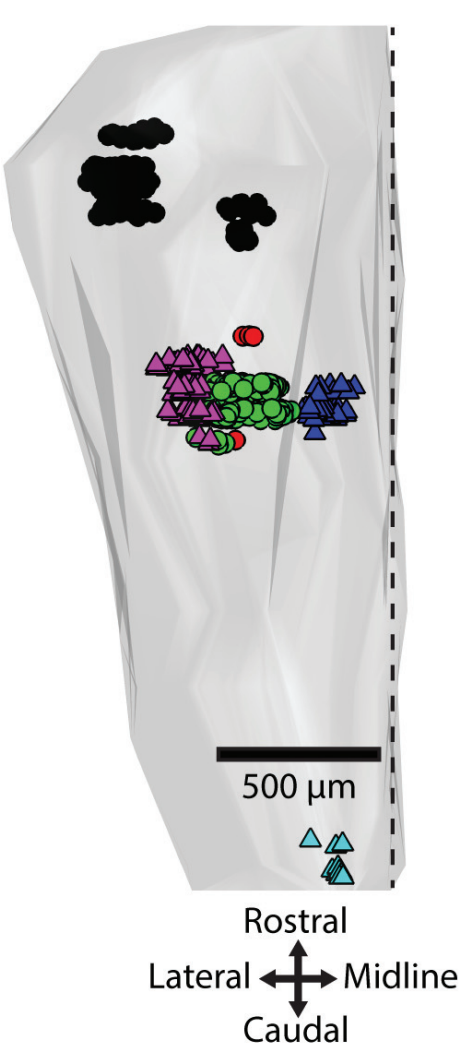
eNeuro Accepted Manuscript

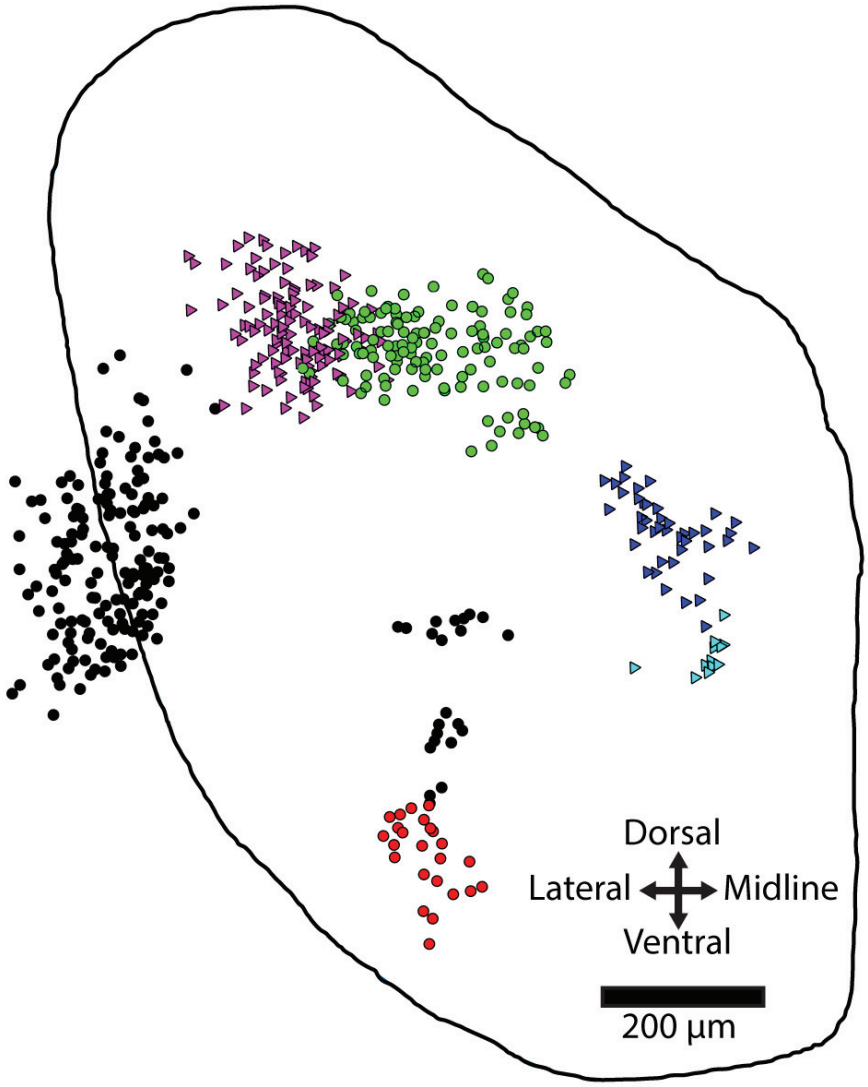


eNeuro Accepted Manuscript

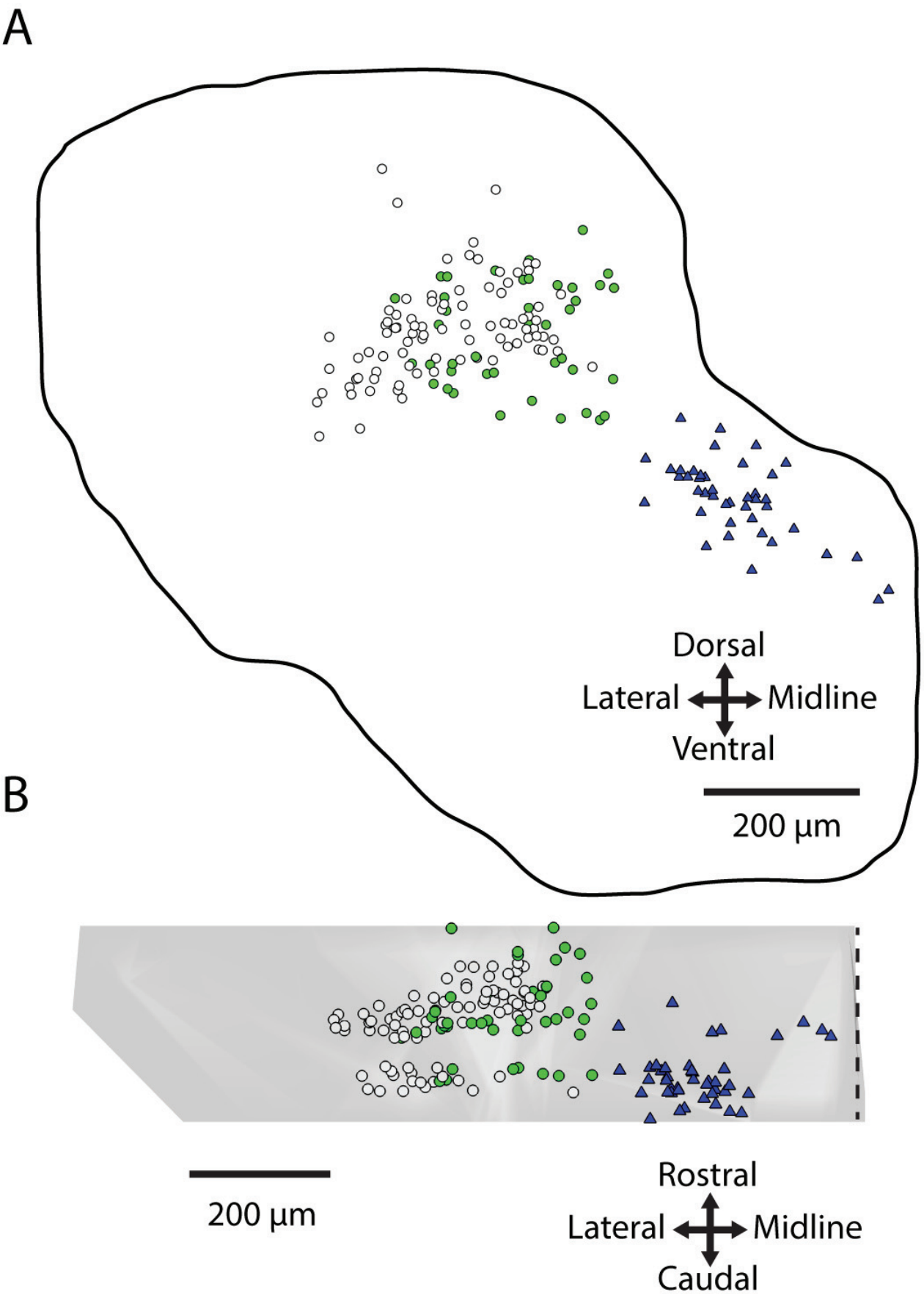


eNeuro Accepted Manuscript

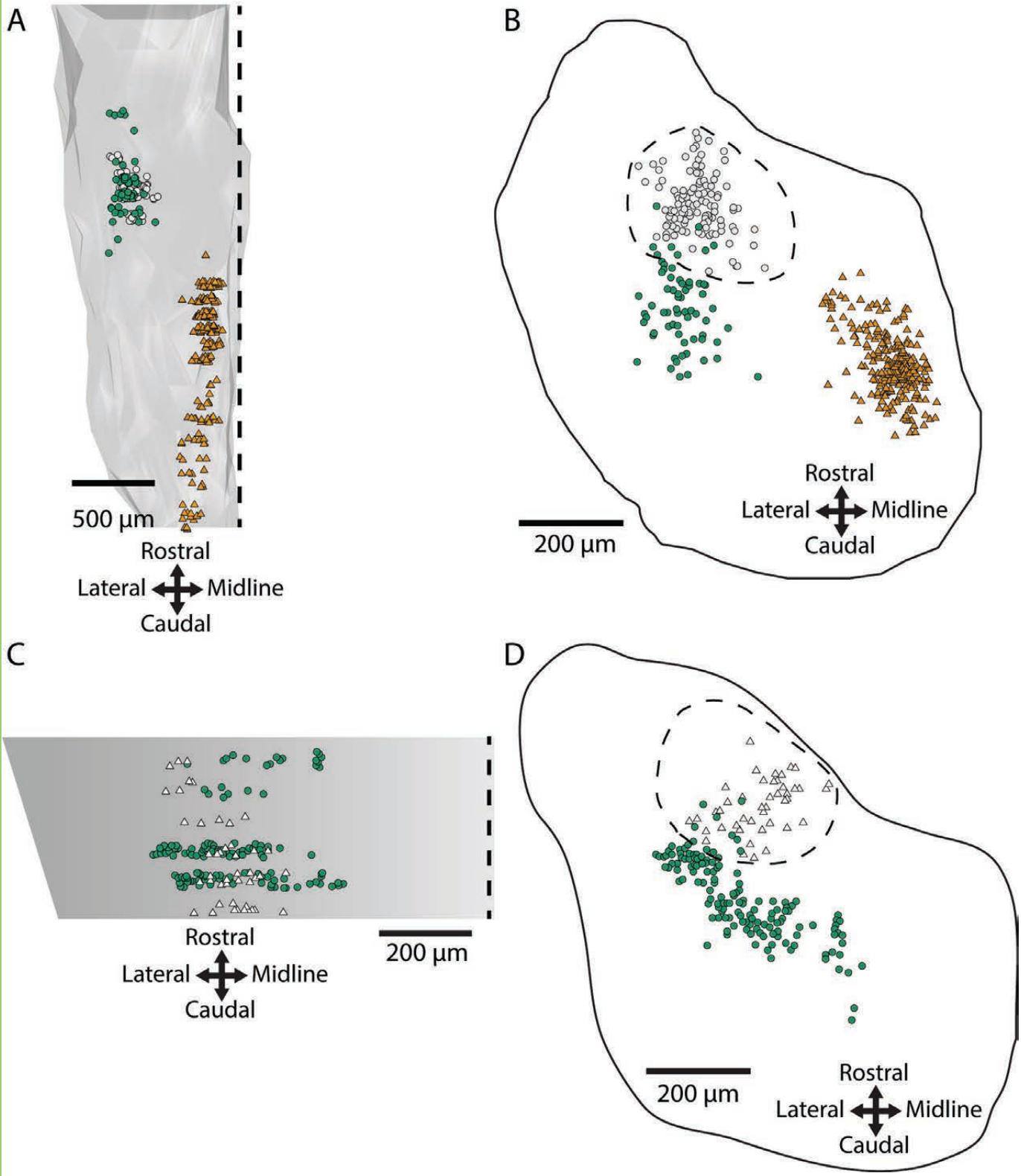




eNeuro Accepted Manuscript



eNeuro Accepted Manuscript



eNeuro Accepted Manuscript

Antigen	Immunogen	Host species:	Reactivity	Source:	Catalog#:	RRID	Dillution	References
Casz1*	Human Casz1 peptide	Rabbit	N/A	Rockland	600-401-B62S	AB_1961496	1:3 000	
Esrrg	AA 2-100 of human Esrrg	Mouse	Mouse, chicken	RnD systems	PP-H6812-00	AB_2100280	1:1 000	
Evx1*	AA 1-192	Rabbit	N/A	Dr. Martyn Goulding	N/A	N/A	1:300	(Moran-Rivard et al., 2001)
Evx2	AA 92-102 of mouse Evx2	Guinea pig	Mouse	Dr. Ryuichi Shirasaki	N/A	N/A	1:6 000	(Inamata and Shirasaki, 2014)
FoxP2	A peptide near N-terminus of human Foxp2	Goat	Mouse, chicken	Santa Cruz	sc-21069	AB_2107124	1:500	
Lbx1	Full length mouse Lbx1	Guinea pig	Mouse, chicken	Dr. Thomas Müller	N/A	N/A	1:30 000	(Muller et al., 2002)
Lhx1/5	AA 1-360 of rat Lhx5	Mouse	Mouse, chicken	DSHB	4F2	AB_531784	1:15	(Tsuchida et al., 1994)
Maf	AA 150-200 of mouse Maf	Rabbit	Mouse, chicken	Bethyl Laboratories	A300-613	N/A	1:2 000	
Myc*	AA 408-439 of human Myc	Mouse	N/A	DSHB	9E 10	AB_2266850	N/A	
Myc*	Full length human Myc	Rabbit	N/A	Millipore	06-340	AB_310106	1:1 000	
Myc*	AA ~1-100 of human Myc	Rabbit	N/A	Abcam	ab32072 (Y69)	AB_731658	1:18 000	
Onecut1	Mix of AA 11-53 and 63-81 of mouse Onecut1	Guinea pig	Mouse	Dr. Frédéric Clotman	N/A	N/A	1:5 000	(Espana and Clotman, 2012)
Onecut1	AA 11-110 of human Onecut1	Rabbit	Mouse, chicken	Santa Cruz	sc-13050	AB_2251852	1:300	
Onecut2	AA 185-326 of human Onecut2	Sheep	Mouse	R&D systems	AF6294	AB_10640365	1:500	
Onecut3	AA 23-333 of mouse Onecut3	Guinea pig	Mouse	Dr. Frédéric Clotman	N/A	N/A	1:6 000	(Pierreux et al., 2004)
Phox2b	A peptide near N-terminus of	Goat	Mouse	Santa Cruz	sc-13224	AB_2251852	1:1 000	

eNeuro Accepted Manuscript

	human Phox2b							
Phox2b	N/A	Rabbit	Chicken (mouse untested)	Dr. Jean-François Brunet	N/A	N/A	1:20 000	Unpublished
Pou3f1	A peptide near C-terminus of human Pou3f1	Goat	Mouse	Santa Cruz	sc-11661	AB_2268536	1:500	

eNeuro Accepted Manuscript

Gene	Species	Groups immunostained	LVST TPM	cMVST TPM
Phox2b	Mouse	LVST	678 ± 126	9 ± 3
	Chicken	LVST	16 ± 20	0 ± 0
Lbx1	Mouse	LVST + cMVST	36 ± 30	0
	Chicken	LVST	0 ± 0	0 ± 0
Maf	Mouse	LVST + cMVST	495 ± 165	390 ± 125
	Chicken	LVST + cMVST	55 ± 35	82 ± 23
Esrrg	Mouse	LVST + cMVST	753 ± 124	588 ± 272
	Chicken	LVST + cMVST	143 ± 106	177 ± 85
Pou3f1	Mouse	LVST + cMVST	103 ± 41	5 ± 2
	Chicken	N/A	0 ± 0	0 ± 0
Onecut1	Mouse	LVST + cMVST	40 ± 23	175 ± 75
	Chicken	LVST + cMVST	50 ± 28	92 ± 49
Onecut2	Mouse	LVST + cMVST	112 ± 20	275 ± 128
	Chicken	N/A	2 ± 3	8 ± 6
Onecut3	Mouse	LVST + cMVST	87 ± 61	131 ± 99
	Chicken	N/A	10 ± 10	6 ± 5
Lhx1	Mouse	cMVST	0	115 ± 35
	Chicken	cMVST	0 ± 0	200 ± 121
Lhx5	Mouse	cMVST	0	57 ± 19
	Chicken	cMVST	1 ± 2	18 ± 14
Foxp2	Mouse	cMVST	0	59 ± 27
	Chicken	cMVST	0 ± 0	65 ± 15
Evx2	Mouse	cMVST	0	12 ± 10
	Chicken	N/A	0 ± 0	9 ± 8

eNeuro Accepted Manuscript

Stage	Mediocaudal vicinal group (Lbx1+, Hoxb1-)	LVST neuron group (Lbx1+, Hoxb1+)	Midbrain group (Lbx1-, Hoxb1-)	N preparations	Sections assessed per preparation
E9.5	0	0	0	3	every 4th
E11.5	0, 0	152, 59	N/A	2	every 4th
E13.5	31 ± 4	203 ± 41	145 ± 13	3	every 6th
E15.5	31 ± 13	186 ± 66	N/A	3	every 6th

eNeuro Accepted Manuscript

Stage	r5-cMVST lateral	Medial extension	N preparations	Sections assessed per preparation
E9.5	0	0	3	every 4th
E11.5	8	2	1	every 4th
E13.5*	85 ± 25	25 ± 21	3	every 6th
E15.5**	98 ± 21	25 ± 9	3	every 6th