Multimodal analysis of cell types in a hypothalamic node controlling social behavior in mice

Thesis by

Dong-Wook Kim

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended December 5, 2019)

Dong-Wook Kim
ORCID: 0000-0002-5497-5853

ACKNOWLEDGEMENTS

Like many other previous PhD students in biology, it was really a long journey full of twists and turns, to finally get to the finish line. First and foremost, I deeply thank my advisor, David J Anderson, for his strong supports, encouragements, and patience throughout my PhD period. The breadth and depth of his knowledge always guided me in the right directions and kept me focused on the core questions that I wanted to address. Especially, given the fact that our every scientific discussion was fruitful, I was very fortunate enough to have you as my PI.

I also would like to thank my other PhD thesis committee members, for their valuable discussions and commitment to my project: Lior Patchter, for his insight on our data analysis; Matthew Thomson and Yuki Oka, for their initial help to establish my 10x experiments. Furthermore, my main thesis project was purely the result of an entire team effort. First of all, I would like to thank all the people in the Allen Institute for Brain Science, for their efforts putting into our paper until the end: Hongkui Zeng and Bosiljka Tasic, for their supervising all my SMART-seq experiments and critical comments on the manuscript; Tae Kyung Kim and Thuc Nghi Nguyen, for their helps to sample and library preparations; Zizhen Yao and Lucas T. Graybuck, for their significant contributions to analyzing and visualizing our data. In addition, I would like to thank the Long Cai's lab members, Noushin Koulena, Nico Pierson, and Sheel Shah, for all their efforts on my seqFISH experiments.

Fortunately, I have met and been surrounded by many fascinating people on campus during my PhD, and eventually learned how important it would be to have great colleagues for conducting my research. All of my pre- and current Anderson lab members should be on the top of the list: Especially, Hyosang Lee, my former desk-mate, for teaching me all the basic lab skills as well as his scientific and professional advice; Liching Lo, my next lab-neighbor sitting behind me, for her encouragement and relentless scientific curiosity; Haijiang Cai, for teaching me his brilliant acute brain slice preparation and E-phys skills; Allan Wong, for his

initial help to set up my two-photon Ca^{2+} imaging rig in slices; Ryan Remedios and Prabhat Kunwar, for helping me acclimate to the Anderson lab; Jung-Sook Chang, for her contributions to the behavior annotations and cell counting; Gian Mancuso, for administrative assistance; Celine Chiu, for lab management; Helen Huang, for mouse genotyping; Tomomi Karigo and George Mountoufaris, for all productive scientific discussions; Brian Duistermars and Brady Weissbourd, for their many good questions during my lab meetings, Vivian Chiu and Zeynep Turan, for their vibrant energy all the time; Ke Ding, for his silly but relaxed jokes; And Kiichi Watanabe and Yonil Jung, my old lab-mates, for all the countless conversions we had each other. Plus, I would thank my Korean friends at Caltech for all the moments and memories we have shared together: Peter Hyung-Kook Lee, Jongmin Kim, Sungwook Woo, Seyoon Kim, Jeesoon Choi, Kiyoul Yang, Mooseok Jang, Jungwoo Kim, Hyun Ji Yun, Ho-Il Ji, Jungmin Kim, Junwon Choi, Sarah Yunmi Lee, Jounhong Ryan Cho, Kyu Hyun Lee, Yongwhi Kim, Dongmin Kang, Jaeho Lee, Sangjun Lee, Hyung Chan Jo, Sangjin Yoo, Minji Jang, Byungkuk Lim, and others.

In addition, I would like to thank all of my off-campus friends for their supports and helps to get through my PhD together: Jeeyun Baik, Hyun-Jung Kim, Joonho Back, Seungjong Lee, Hongju Jeon, Minji Lim, Sungwoo Park, Christina S Kang, Sung-Hwan Kim, Jae Hyun Bae, Yong Sun Kim, Ahrang Hong, HongSuk Chun, Chae Suk Lim, and all members in Korean Salvation Army in L.A.

Lastly, this long journey would not have been possible without the endless supports and sacrifices of my family. My dad and mom always showed their unconditional love and dedication to me; my sister and brother-in-law took care of all family-related things smoothly (and thank to my twin cousins too); my wife and son always remind me the meaning of my life. Also, I would like to thank my in-laws, for their remote baby-cares (via Facetime) and praying every day.

ABSTRACT

The advent and recent advances of single-cell RNA sequencing (scRNA-seq) have yielded transformative insights into our understanding of cellular diversity in the central nervous system (CNS) with unprecedented detail. However, due to current experimental and computational limitations on defining transcriptomic cell types (T-types) and the multiple phenotypic features of cell types in the CNS, an integrative and multimodal approach should be required for the comprehensive classification of cell types.

To this end, performing multimodal analysis of scRNA-seq in hypothalamus would be very beneficial in that hypothalamus, controlling homeostatic and innate survival behaviors which known to be highly conserved across a wide range of species and encoded in hard-wired brain circuits, is likely to display the more straightforward relationship between transcriptomic identity, axonal projections, and behavioral activation, respectively. In my dissertation, I have been focused on the cell type characterizations of a hypothalamic node controlling innate social behavior in mice, the ventrolateral subdivision of the ventromedial hypothalamus (VMHvl). VMHvl only contains ~4,000 neurons per hemisphere in mice but due to its behavioral, anatomical, and molecular heterogeneity, which T-types in VMHvl are related to connectivity and behavioral function is largely unknown.

In Chapter II, I described my main thesis work to perform scRNA-seq in VMHvl using two independent platforms: SMART-seq2 (~4,500 neurons sequenced) and 10x (~78,000 neurons sequenced). Specifically, 17 joint VMHvl T-types including several sexually dimorphic clusters were identified by canonical correlation analysis (CCA) in Seurat, and the majority of them were validated by multiplexed single-molecule FISH (seqFISH). Correspondence between transcriptomic identity, and axonal projections or behavioral activation, respectively, was also investigated. Immediate early gene analysis identified T-types exhibiting preferential responses to intruder males versus females but only rare examples of behavior-specific activation. Unexpectedly, many VMHvl T-types comprise a mixed population of neurons with different projection target preferences. Overall our

analysis revealed that, surprisingly, few VMHvl T-types exhibit a clear correspondence with behavior-specific activation and connectivity.

In Chapter III, I will discuss about future directions for a deeper and better understanding of VMHvl cell types. Briefly, my previous data from whole-cell patch clamp recording in VMHvl slices suggested that there were at least 4 distinct electrophysiological cell types (E-types). Additionally, two distinct neuromodulatory effects on VMHvl were observed (persistently activated by vasopressin/oxytocin vs. silenced by nitric oxide) by monitoring populational activities using two-photon Ca^{2+} imaging in slices. Based on the results from the first part and combined with advanced molecular techniques (e.g. Patch-seq and CRISPR-Cas9), we can further dissect out the cellular diversity in VMHvl and their functional implications.

PUBLISHED CONTENT AND CONTRIBUTIONS

Kim, D-.W. et al. (2019). “Multimodal analysis of cell types in a hypothalamic node controlling social behavior”. In: *Cell* 179.3, pp.713–728. doi: 10.1016/j.cell.2019.09.020.

D-.W.K participated in the experimental designs, collected and analyzed data, prepared relevant figures, and participated in the writing of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND/OR TABLES

Chapter II

Chapter III

NOMENCLATURE

Act-seq. single-cell RNA-sequencing following induced neuronal activity

AUROC. Area Under Receiver-Operator Characteristics curve.

catFISH. cellular Compartment Analysis of Temporal activity by FISH

CCA. Canonical Correlation Analysis.

ChR2. Channelrhodopsin-2.

CI. Close Investigation.

dPAG. dorsal Periaqueductal gray.

Esr1. Estrogen Receptor Type 1.

FACS. Fluorescence Activated Cell Sorting.

FISH. Fluorescence *In Situ* Hybridization.

HGNC. HUGO Genome Nomenclature Committee.

IEG. Immediate Early Gene.

lPAG. lateral Periaqueductal gray.

MERFISH. Multiplexed error-robust FISH

MPOA. Medial preoptic area

Retro-seq. single-cell RNA-sequencing following retrograde labeling

scRNA-seq. single-cell RNA-sequencing.

seqFISH. sequential multiplexed smFISH.

smFISH. single-molecule FISH.

t-SNE. t-distributed stochastic neighbor embedding.

T-types. Transcriptomic “leaf” clusters.

VMH. Ventromedial hypothalamus.

VMHdm/c. Ventromedial and central subdivision of ventromedial hypothalamus.

VMHvl. Ventrolateral subdivision of ventromedial hypothalamus.

UMAP. Uniform manifold approximation and projection.

C h a p t e r I

INTRODUCTION

1.1 Cell type characterization using scRNA-seq and previous efforts to find its biological correspondence

Since the first attempts made by Ramón y Cajal to characterize morphological cell types (Ramón y Cajal, 1909), it has been one of central themes in neuroscience to systematically understand the cellular diversity in the brain and investigate the functional roles of the identified cell types (Jorgenson et al., 2015). While the existence of distinct brain cell types has been discovered based on various criteria including electrophysiological (Llinas, 1988), molecular (Nelson et al., 2006), morphological (Seung and Sümbül, 2014), and integrative approaches (Ascoli et al., 2008; Migliore and Shepherd, 2005), the more comprehensive pictures of cell taxonomy and their functional roles have been still needed to be achieved.

The advances of single-cell RNA-sequencing (scRNA-seq) technologies enable us to get access to the enriched information of transcriptomic cell types (T-types) from either a given area of central nervous system (CNS) (Darmanis et al., 2015; Macosko et al., 2015; Tasic et al., 2016) or even an entire organism (Saunders et al., 2018; Zeisel et al., 2018) in a genome-wide, high-throughput, and unbiased manner. However, given the facts that 1) cell types in CNS should be defined by many other cellular phenotypic features than transcriptomics and 2) the stringency and criteria of clustering analysis for scRNA-seq data are somewhat subjective (lumping vs. splitting) partly due to technical limitations of the methods (Tasic, 2018), the question about how such diverse T-types are related to other facets of cell identity, such as functional, anatomical, and physiological properties, follows immediately.

There were some previous attempts to perform multimodal analysis for better understanding of identified T-types. For examples, a technique called Patch-Seq for acquiring

electrophysiological, morphological, and transcriptomic information from the same cell by using a patch-clamp pipette was recently introduced (Cadwell et al., 2016; Fuzik et al., 2016). In spite of their technical limitations, including 1) a low number of total transcripts captured per cell, 2) high likelihood of sample contaminations from off-target sources of mRNA, and 3) relatively low-throughput and labor-intensive, it was successfully applied to this tri-modal data acquisition pipeline in the primary visual cortex recently (Gouwens et al., 2019). In addition, there are at least two distinct methods to map the projection information onto T-types: in the studies using Retro-seq (Tasic et al., 2016; Tasic et al., 2018), retrogradely-labeled single-cells were isolated by fluorescence activated cell sorting (FACS) and scRNA-seq was followed, whereas in case of BARseq (Barcoded Anatomy Resolved by sequencing), previous MAPseq (Multiplexed Analysis of Projections by Sequencing; a high-throughput technique labeling neurons in a given area with random unique RNA sequences (called barcodes) to trace their projection patterns by bulk RNA sequencing; Kebschull et al., 2016) was combined with *in-situ* sequencing or multiplexed smFISH to link multiple projectional information at a single-cell resolution with the expression patterns of marker genes from scRNA-seq (Chen et al., 2019).

In parallel, there were broadly three different approaches trying to investigate the functional relevance of identified T-types during certain behaviors or sensory stimulations. By using a new FosTRAP transgenic mouse line, TRAP2 (Guenthner et al., 2013), authors could efficiently and specifically label thirst-activated neurons in the median preoptic nucleus (MnPO), and then run scRNA-seq after isolating labeled-cells via FACS (Allen et al., 2017) to identify the transcriptomic profiles of these activated populations. Although this method has a great potential combined with different effector viruses expressing either a optogenetic opsin or genetically encoded Ca^{2+}/voltage indicator on $TRAP^{+}$ populations as illustrated in this paper, it has some caveats in that 1) it requires breeding the transgenic mouse line, 2) TRAP efficiency and sensitivity would be varied from region to region, and 3) independent additional scRNA-seq experiments are needed to get the T-type information of $TRAP^{-}$ populations in a given region. In a second approach, scRNA-seq was combined with multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) to identify

behaviorally-relevant T-types (Moffitt et al., 2018). Specifically, they first conducted scRNA-seq in the preoptic region (POR), and then MERFISH was performed with the immediate early gene (IEG) *c-fos* and ~150 previously identified marker genes to see which T-types were preferentially activated during different social behaviors, in a sex-specific manner. However, as the TRAP case, this paper only relied on *c-fos* as a neural activity marker. Thus, it would not be feasible to be adapted in other brain areas where *c-fos* is not efficiently induced by neural activity or its background expressions are too high so that other IEGs would be needed instead. Lastly, a technique called Act-seq (Wu et al., 2017; Hrvatin et al., 2018) combines a high-throughput droplet-based scRNA-seq with a new single-cell preparation protocol that minimize artificially evoked transcriptional perturbations. Therefore, it enables the sensitive detection of gene expression changes by acute behaviors or stimulations, and thus by using endogenous inductions of IEGs as activity markers, we can map functional information onto T-types.

Nevertheless, except a few exceptions, one of which Retro-seq and morphological reconstructions were combined with *in vivo* recordings and opto-tagging showing which two distinct descending projection cell types are differently involved in motor control by incorporating multi-modal information (Economo et al., 2018), dual modality convergences are common so far even in a handful of cases. Therefore, it is extremely important to do a next multimodal approach for figuring out how transcriptomic heterogeneity is mapped onto connectivity and behavioral function.

1.2 Functional, anatomical, and molecular heterogeneity in VMHvl

The ventromedial hypothalamus (VMH) is a nucleus located at the mediobasal hypothalamus bilaterally, and has distinct cytoarchitectures revealed by various histological methods including Nissl staining (e.g. elliptical shape of cell nucleus surrounded by a cell-poor/fiber-rich zone). Since the first VMH lesion study in rats showed dramatic increases of feeding and body weight (Hetherington and Ranson, 1942), VMH was recognized as one of brain satiety centers regulating body weight and energy metabolism (reviewed in King, 2006). Among three anatomic subdivisions of VMH (dorsomedial (VMHdm), central (VMHc), and ventrolateral (VMHvl), respectively), VMHdm was more focused on these metabolic roles because one of VMHdm-specific transcription factors, Steroidogenic factor-1, *Sf-1* (*NR5A1*) was crucial for controlling energy homeostasis (Klöckener, et al., 2011). In contrast, VMHvl was known for being associated with innate social behaviors, such as conspecific aggression and mating, and although it only contains ~4,000 neurons per hemisphere in mice, it is increasingly clear that VMHvl are behaviorally, anatomically, and molecularly heterogeneous (reviewed in Anderson, 2016; Hashikawa et al., 2017b).

Firstly, since the seminal discovery showing the electrical activation of the lateral hypothalamus in cats could evoke a strong aggressive response (Hess and Brügger, 1943), the follow-up studies extensively tried to delineate this hypothalamic area in which micro-stimulation could trigger eliciting attacks in multiple other species, and eventually a so-called hypothalamic attack area (HAA) was broadly identified (Hrabovszky et al., 2005; also reviewed in Siegel et al., 1999). However, due to technical limitations (e.g. co-activations of fibers-of-passage and no cell-type specificity), the more precise location of the relevant neurons still remained unclear until the advent of optogenetics which allows researchers to precisely manipulate neuronal populations in a cell-type specific manner. The first optogenetic study in VMHvl clearly described its causal role in male aggression (Lin et al., 2011), and two largely overlapped subpopulations in VMHvl, progesterone receptor-expressing (VMHvl^{Pgr}; Yang et al., 2013) and estrogen receptor type 1-expressing neurons (VMHvl^{Esr1}; Lee et al., 2014) were shown to control both male aggression and mating

behavior. And more recently, VMHvl^{Esr1} also plays an essential role in female aggression and mating respectively in two separate subdivisions of VMHvl (Hashikawa et al., 2017a).

There is also accumulating evidence in support of the idea that VMHvl is responsible for metabolism and other social behaviors besides aggression and mating. In females, VMHvl^{Esr1} was known to play a role in controlling energy balance and the body weight maintenance (Musatov et al., 2007), and more recently a subpopulation of VMHvl^{Esr1}, which is 1) marked by NK2 homeobox transcription factor 1 (*Nkx2-1*) and tachykinin 1 (*Tac1*), 2) dedicated to mediating estrogen-dependent female locomotion, and 3) different from other VMHvl^{Esr1} subpopulation controlling female reproductive behaviors, was found (Correa et al., 2015). In addition, some neuronal populations of VMHvl inhibited by glucose would contribute to systemic glucoregulation like the case of VMHdm/c (reviewed in Shimazu and Minokoshi, 2017). Furthermore, VMHvl neurons preferentially activated during social defeat and subsequent fear responses (called social fear) were reported (Silva et al., 2013), and by using activity-dependent labeling/manipulating technique (Sakurai et al., 2016) and optogenetics (Wang et al., 2019), these VMHvl "social fear" populations were distinct from VMHvl^{Esr1} eliciting male-male aggression. Interestingly, VMHvl neurons promoting aggression seeking behavior in male mice was found (Falkner et al., 2016), suggesting that VMHvl could be an important circuit-node encoding an internal state of aggression arousal.

Heterogeneous physiological responses of VMHvl neurons during resident-intruder assays under various conditions were also observed. Specifically, *in vivo* single-unit recordings in VMHvl were conducted to show that VMHvl carries information for the imminence of attack and future attack duration after accounting for the variation in sensory and motor parameters (Falkner et al., 2014). More interestingly, by monitoring populational Ca^{2+} signals in VMHvl^{Esr1} measured by microendoscopic imaging under freely behaving conditions, intruder sex identity was encoded separately in VMHvl^{Esr1} after brief sexual experience with females (Remedios et al., 2017), suggesting that experience-dependent plasticity after sexual experience might be involved in changing neural representations of social information in VMHvl^{Esr1} for appropriate action selection in social contexts.

As other hypothalamic nucleus (reviewed in Swanson, 2000; Swanson, 2005), previous anterograde and retrograde tracing studies revealed complex input and output connectivity patterns in VMHvl (Canteras et al., 1994; Toth et al., 2010). Especially, recent study using various viral-genetic tracing methods nicely illustrated that VMHvl^{Esr1} has high convergence (fan-in) and divergence (fan-out) in their inputs and outputs, respectively, instead of showing a simple 1-1 relationship (Lo et al., 2019). Interestingly, two distinct morphological subpopulations which have been described in VMHvl by classical Golgi staining (Millhouse, 1973a; b) or electrophysiological recordings of antidromic spikes (Sakura and Tada, 1984; Sakura and Akaishi, 1987) were also recapitulated such that the projection-defined subpopulations of VMHvl^{Esr1} (anterior vs. posterior biases in their collateralization targets) showed different patterns of distribution along AP axis (posterior vs. anterior), cell size (small vs. large), and local branching (small vs. large).

Lastly, the distinct expression patterns of various genes including transcription factors, neuropeptides, membrane receptors, and neurotransmitter receptor subunits were observed in VMH/VMHvl by bulk RNA-seq (Kurrasch et al., 2007; Segal et al., 2005) and *in situ* hybridization (reviewed in McClellan et al., 2006). Consistent with these heterogeneous expressions, *in vitro* extracellular slice recording showed that VMHvl cells respond to a variety of neuromodulators (e.g. acetylcholine, norepinephrine, serotonin, and dopamine), neuropeptides (e.g. oxytocin), and steroid sex hormone (e.g. estrogen) by bath applications (reviewed in Hashikawa et al., 2017b).

Taken together, in the following chapter, we applied multimodal analysis (transcriptomics + behavioral activation + projection specificity) to further dissect out these cellular heterogeneity in VMHvl.

1.3 References

Anderson, D.J. (2016). Circuit modules linking internal states and social behaviour in flies and mice. *Nat. Rev. Neurosci.* 17, 692-704.

Allen, W.E., DeNardo, L.A., Chen, M.Z., Liu, C.D., Loh, K.M., Fenno, L.E., Ramakrishnan, C., Deisseroth, K., and Luo, L. (2017). Thirst-associated preoptic neurons encode an aversive motivational drive. *Science* 357, 1149-1155.

Ascoli, G. A., et al. (2008). Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. Nat. Rev. Neurosci. 9, 557–568.

Cadwell, C.R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., Reimer, J., Shen, S., Bethge, M., Tolias, K.F., et al. (2016). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat. Biotechnol.* 34, 199-203.

Canteras, N.S., Simerly, R.B., and Swanson, L.W. (1994). Organization of projections from the ventromedial nucleus of the hypothalamus: A Phaseolus vulgaris-leucoagglutinin study in the rat. *J. Comp. Neurol.* 348, 41–79.

Chen, X., Sun, Y.C., Zhan, H., Kebschull, J.M., Fischer, S., Matho, K., Huang, Z.J., Gillis, J., and Zador, A.M. (2019). High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ Sequencing. *Cell* 179, 772-786 e719.

Correa, S.M., Newstrom, D.W., Warne, J.P., Flandin, P., Cheung, C.C., Lin- Moore, A.T., Pierce, A.A., Xu, A.W., Rubenstein, J.L., and Ingraham, H.A. (2015). An estrogen-responsive module in the ventromedial hypothalamus selectively drives sex-specific activity in females. *Cell Rep.* 10, 62–74.

Darmanis, S., et al. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. USA*, 7285-7290.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. *Nature* 563, 79–84.

Falkner, A.L., Dollar, P., Perona, P., Anderson, D.J., and Lin, D. (2014). Decoding ventromedial hypothalamic neural activity during male mouse aggression. *J. Neurosci.* 34, 5971–5984.

Falkner, A.L., Grosenick, L., Davidson, T.J., Deisseroth, K., and Lin, D. (2016). Hypothalamic control of male aggression-seeking behavior. *Nat. Neurosci.* 19, 596-604.

Fuzik, J., Zeisel, A., Mate, Z., Calvigioni, D., Yanagawa, Y., Szabo, G., Linnarsson, S., and Harkany, T. (2016). Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* 34, 175-183.

Gouwens, N.W., Sorensen, S.A., Berg, J., Lee, C., Jarsky, T., Ting, J., Sunkin, S.M., Feng, D., Anastassiou, C.A., Barkan, E., et al. (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat Neurosci* 22, 1182-1195.

Guenthner, C.J., Miyamichi, K., Yang, H.H., Heller, H.C., and Luo, L. (2013). Permanent genetic access to transiently active neurons via TRAP: targeted recombination in active populations. *Neuron* 78, 773-784.

Hashikawa, K., Hashikawa, Y., Tremblay, R., Zhang, J., Feng, J.E., Sabol, A., Piper, W.T., Lee, H., Rudy, B., and Lin, D. (2017a). Esr1(+) cells in the ventromedial hypothalamus control female aggression. *Nat. Neurosci.* 20, 1580-1590.

Hashikawa, Y., Hashikawa, K., Falkner, A.L., and Lin, D. (2017b). Ventromedial Hypothalamus and the Generation of Aggression. *Front. Syst. Neurosci.* 11, 94.

Hess, W.R. and Brügger, M. (1943). Das subkortikale Zentrum der affecktiven Abwehr-reaktion. *Helv. Physiol. Acta.* 1, 33–52 (in German).

Hetherington, A.W., and Ranson, S.W. (1942). The relation of various hypothalamic lesions to adiposity in the rat. *J. Comp. Neurol.* 76, 475-499.

Hrabovszky, E., Halasz, J., Meelis, W., Kruk, M.R., Liposits, Z., and Haller, J. (2005). Neurochemical characterization of hypothalamic neurons involved in attack behavior: glutamatergic dominance and co-expression of thyrotropin-releasing hormone in a subset of glutamatergic neurons. *Neuroscience* 133, 657-666.

Hrvatin, S., Hochbaum, D.R., Nagy, M.A., Cicconet, M., Robertson, K., Cheadle, L., Zilionis, R., Ratner, A., Borges-Monroy, R., Klein, A.M., et al. (2018). Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci.* 21, 120-129.

Jorgenson, L.A., Newsome, W.T., Anderson, D.J., Bargmann, C.I., Brown, E.N., Deisseroth, K., Donoghue, J.P., Hudson, K.L., Ling, G.S., MacLeish, P.R., et al. (2015). The BRAIN Initiative: developing technology to catalyse neuroscience discovery. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20140164.

Kebschull, J.M., Garcia da Silva, P., Reid, A.P., Peikon, I.D., Albeanu, D.F., and Zador, A.M. (2016). High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. *Neuron* 91, 975-987.

King, B.M. (2006). The rise, fall, and resurrection of the ventromedial hypothalamus in the regulation of feeding behavior and body weight. *Physiol. Behav.* 87, 221–244.

Klöckener, T., Hess, S., Belgardt, B.F., Paeger, L., Verhagen, L.A., Husch, A., Sohn, J.W., Hampel, B., Dhillon, H., Zigman, J.M., et al. (2011). High-fat feeding promotes obesity via insulin receptor/PI3K-dependent inhibition of SF-1 VMH neurons. *Nat. Neurosci.* 14, 911–918.

Kurrasch, D.M., Cheung, C.C., Lee, F.Y., Tran, P.V., Hata, K., and Ingraham, H.A. (2007). The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. *J. Neurosci.* 27, 13624–13634.

Lee, H., Kim, D.W., Remedios, R., Anthony, T.E., Chang, A., Madisen, L., Zeng, H., and Anderson, D.J. (2014). Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. *Nature* 509, 627-632.

Lin, D., Boyle, M.P., Dollar, P., Lee, H., Lein, E.S., Perona, P., and Anderson, D.J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. *Nature* 470, 221-226.

Llinas, R.R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242, 1654–1664.

Lo, L., Yao, S., Kim, D.W., Cetin, A., Harris, J., Zeng, H., Anderson, D.J., and Weissbourd, B. (2019). Connectional architecture of a mouse hypothalamic circuit node controlling social behavior. *Proc. Natl. Acad. Sci. USA* 116, 7503–7512.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214.

Masland, R.H. (2004). Neuronal cell types. *Curr Biol,* 14, 497-500

McClellan, K.M., Parker, K.L., and Tobet, S. (2006). Development of the ventromedial nucleus of the hypothalamus. *Front. Neuroendocrinol.* 27, 193-209.

Migliore, M., and Shepherd, G.M. (2005). An integrated approach to classifying neuronal phenotypes. *Nat. Rev. Neurosci*. 6, 810–818.

Millhouse, O.E. (1973a). Certain ventromedial hypothalamic afferents. *Brain Res.* 55, 89–105.

Millhouse, O.E. (1973b). The organization of the ventromedial hypothalamic nucleus. *Brain Res.* 55, 71–87.

Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* 362, eaau5324.

Musatov, S., Chen, W., Pfaff, D.W., Mobbs, C.V., Yang, X.J., Clegg, D.J., Kaplitt, M.G., and Ogawa, S. (2007). Silencing of estrogen receptor alpha in the ventromedial nucleus of hypothalamus leads to metabolic syndrome. *Proc. Natl. Acad. Sci. USA* 104, 2501–2506.

Nelson, S.B., Sugino, K., and Hempel, C.M. (2006). The problem of neuronal cell types: a physiological genomics approach. *Trends Neurosci.* 29, 339–345.

Ramón y Cajal, S. (1909). Histology of the nervous system of man and vertebrates, English language edn (Oxford University Press).

Remedios, R., Kennedy, A., Zelikowsky, M., Grewe, B.F., Schnitzer, M.J., and Anderson, D.J. (2017). Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex. *Nature* 550, 388–392.

Sakuma, Y., and Akaishi, T. (1987). Cell size, projection path, and localization of estrogen-sensitive neurons in the rat ventromedial hypothalamus. *J. Neurophysiol.*

Sakuma, Y., and Tada, K. (1984). Evidence that two sizes of ventromedial hypothalamic neurones project to the mesencephalic central grey matter in rats. *J. Physiol.* 349, 287-97.

Sakurai, K., Zhao, S., Takatoh, J., Rodriguez, E., Lu, J., Leavitt, A.D., Fu, M., Han, B.-X., and Wang, F. (2016). Capturing and Manipulating Activated Neuronal Ensembles with CANE Delineates a Hypothalamic Social-Fear Circuit. *Neuron* 92, 739–753.

Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., et al. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* 174, 1015–1030 e1016.

Segal, J.P., Stallings, N.R., Lee, C.E., Zhao, L., Socci, N., Viale, A., Harris, T.M., Soares, M.B., Childs, G., Elmquist, J.K., et al. (2005). Use of laser-capture microdissection for the identification of marker genes for the ventromedial hypothalamic nucleus. *J. Neurosci.* 25, 4181-4188.

Seung, H.S., and Sümbül, U. (2014). Neuronal cell types and connectivity: lessons from the retina. *Neuron* 83, 1262–1272.

Siegel, A., Roeling, T.A., Gregg, T.R., and Kruk, M.R. (1999). Neuropharmacology of brain-stimulation-evoked aggression. *Neurosci. Biobehav. Rev.* 23, 359-389.

Silva, B.A., Mattucci, C., Krzywkowski, P., Murana, E., Illarionova, A., Grinevich, V., Canteras, N.S., Ragozzino, D., and Gross, C.T. (2013). Independent hypothalamic circuits for social and predator fear. *Nat. Neurosci.* 16, 1731–1733.

Swanson, L.W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Res*, 886, 113–164.

Swanson, L.W. (2005). Anatomy of the soul as reflected in the cerebral hemispheres: Neural circuits underlying voluntary control of basic motivated behaviors. *J. Comp. Neurol.* 493, 122–131.

Tasic, B. (2018). Single cell transcriptomics in neuroscience: cell classification and beyond. *Curr. Opin. Neurobiol*. 50, 242–249.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci*. 19, 335–346.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature* 563, 72–78.

Toth, M., Fuzesi, T., Halasz, J., Tulogdi, A., and Haller, J. (2010). Neural inputs of the hypothalamic "aggression area" in the rat. *Behav. Brain Res.* 215, 7–20.

Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. *Neuron* 96, 313-329 e316.

Yang, C.F., Chiang, M.C., Gray, D.C., Prabhakaran, M., Alvarado, M., Juntti, S.A., Unger, E.K., Wells, J.A., and Shah, N.M. (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* 153, 896-909.

Zeisel, A., Hochgerner, H., Lonnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molec- ular Architecture of the Mouse Nervous System. *Cell* 174, 999–1014 e1022.

Chapter II

MULTIMODAL ANALYSIS OF CELL TYPES IN A HYPOTHALAMIC NODE CONTROLLING SOCIAL BEHAVIOR

Kim, D-.W. *et al*. (2019). "Multimodal analysis of cell types in a hypothalamic node controlling social behavior". In: *Cell* 179.3, pp.713–728. doi: 10.1016/j.cell.2019.09.020.

2.1 Summary

The ventrolateral subdivision of the ventromedial hypothalamus (VMHvl) contains ~4,000 neurons that project to multiple targets and control innate social behaviors including aggression and mounting. However, the number of cell types in VMHvl and their relationship to connectivity and behavioral function are unknown. We performed single-cell RNA sequencing using two independent platforms—SMART-seq (~4,500 neurons) and 10x (~78,000 neurons)—and investigated correspondence between transcriptomic identity and axonal projections or behavioral activation, respectively. Canonical correlation analysis (CCA) identified 17 transcriptomic types (T-types), including several sexually dimorphic clusters, the majority of which were validated by seqFISH. Immediate early gene analysis identified T-types exhibiting preferential responses to intruder males versus females but only rare examples of behavior-specific activation. Unexpectedly, many VMHvl T-types comprise a mixed population of neurons with different projection target preferences. Overall our analysis revealed that, surprisingly, few VMHvl T-types exhibit a clear correspondence with behavior-specific activation and connectivity.

2.2 Introduction

Cell types are a fundamental unit of organization and specificity in multicellular organisms. An understanding of cellular diversity in the brain is critical for studies of neural function and dysfunction (Jorgenson et al., 2015). Using single-cell RNA sequencing (scRNA-seq), recent surveys have estimated up to ~600 different transcriptomic cell types (T-types) in the mouse brain (Saunders et al., 2018; Zeisel et al., 2018). Such diversity immediately poses the ‘‘correspondence problem:’’ how is transcriptomic heterogeneity related to other facets of neuronal identity, such as connectivity and physiology (Tasic, 2018; Zeng and Sanes, 2017)? While such correspondence is well established for retinal cell types (Macosko et al., 2015; Seung and Sümbül, 2014), it is not yet clear whether this principle extends to the central brain. For example, out of over a hundred cortical T-types (Tasic et al., 2018), only two have been shown to project to different subcortical targets and to exert distinct functions in motor control (Economo et al., 2018).

The hypothalamus is an evolutionarily ancient collection of deep subcortical nuclei that control homeostatic and innate ‘‘survival behaviors’’ and associated motivational states (reviewed in Luiten et al., 1987; Saper and Lowell, 2014; Sternson, 2013). The instinctive nature of these functions suggests they might be controlled by specific T-types with genetically specified connectivity. Initial scRNA-seq studies have revealed evidence of extensive hypothalamic cell diversity (2–35 cell types/mm3 tissue sampled) (Campbell et al., 2017; Chen et al., 2017; Romanov et al., 2017), but the behavioral relevance of such cell types was not examined. A recent study of the preoptic region (POR), a large area (~20% of total hypothalamus volume) containing ~20 distinct subdivisions (nuclei), revealed ~70 T-types (~3–4 T-types per nucleus). Multiplexed error-robust fluorescence in situ hybridization (MERFISH) (Moffitt et al., 2016) experiments using the immediate early gene (IEG) c-fos (Greenberg and Ziff, 1984; Morgan et al., 1987) indicated that some T-types were preferentially activated during a particular social behavior (Moffitt et al., 2018). However, the relationship between transcriptomic identity and axonal projections (Kohl et al., 2018) was not investigated.

The ventromedial hypothalamus, ventrolateral subdivision (VMHvl) occupies ~0.5% of hypothalamus volume and contains ~4,000 primarily glutamatergic neurons that collectively control social behaviors, including aggression, as well as metabolism (reviewed in Chen and Hong, 2018; Hashikawa et al., 2017b; Kennedy et al., 2014; Krause and Ingraham, 2017). Calcium imaging of estrogen receptor type 1-expressing VMHvl (VMHvl^{Esr1}) neurons during social behavior revealed that population activity represents intruder sex identity and behavior (Remedios et al., 2017), while anatomic analysis indicated that VMHvl^{Esr1} neurons project to multiple (~30) downstream targets (Lo et al., 2019). Cellular subpopulations in VMHvl have been identified using morphology (Millhouse, 1973a; b), molecular markers (Correa et al., 2015; Xu et al., 2012), and bulk RNA-seq (Hashikawa et al., 2017a), but the behavioral function and connectivity of these populations were not established.

We have carried out scRNA-seq of VMHvl neurons at high sampling density (2×10^5 neurons sequenced/mm^3 of tissue), using two independent platforms: SMART-seq v4 (Picelli et al., 2013) and the 10x genomics droplet-based platform (Zheng et al., 2017). Our results identify 17 T-types in this subnucleus alone, a density (~160 T-types/mm^3 tissue sampled) 3- to 68-fold higher than reported in previous studies (Chen et al., 2017; Mickelsen et al., 2019; Moffitt et al., 2018; Romanov et al., 2017). Investigation of correspondence between anatomy, behavioral activation, and projection specificity using seqFISH (Eng et al., 2019; Shah et al., 2016), Act-seq (Wu et al., 2017), and Retro-seq (Tasic et al., 2018) indicated that, with a few notable exceptions, most VMHvl T-types do not map to specific behaviors and project to specific targets.

2.3 Results

A Census of VMHvl Transcriptomic Cell Types Using SMART-seq

Initially, we used SMART-seq (v4, Clontech) to derive a census for VMHvl cell types. Because of its small size (Figure S1A) and the difficulty of manually dissecting it free from surrounding tissue, we pre-selected for neurons in this nucleus using transgenic Cre lines, fluorescent Cre-dependent reporters, and fluorescence activated cell sorting (FACS; Figure S1B) (Tasic et al., 2016; Tasic et al., 2018). This approach yielded ~100 cells per animal on average. We used several Cre lines (see Methods; Figures S1C–S1E); *Nr5a1-Cre* was particularly useful because it afforded labeling of ~75% of all cells in adult VMHvl (Figures S1G and S1H; *Nr5a1* is transiently expressed in most VMH cells during development but is downregulated in adult VMHvl [Cheung et al., 2013]). We collected 4,037 cells from 48 males and 537 cells from six females (Figure S1C) under home cage or a variety of behavioral conditions (Figures S1D and S1E). In total, 4,574 cells were sequenced from 54 mice, of which 4,473 were neurons (Figure 1C).

Clustering using the released version of scratch.hicat (Tasic et al., 2018) yielded 46 transcriptomic ‘‘leaf’’ clusters, or transcriptomic cell types (T-types), at the terminal hierarchical level (Figure 1A), six of them non-neuronal and the remainder neuronal (Figure 1C). A similar number of clusters was identified using two other independent methods: Ward’s hierarchical clustering (Ward, 1963) or graph-based clustering using Seurat (Butler et al., 2018) (Figures S2D and S2E). We used differentially expressed genes (DEGs) and prior knowledge to group neuronal T-types into several broad categories with different predicted anatomic locations. VMH has a core-and-shell organization, such that cells within the core are mostly glutamatergic and express *Fezf1*, *Adcyap1*, *Gda*, *Nrgn*, *Cbln1*, *Lmo3*, *Nr5a1*, *C1ql2*, and *Rreb1* (Kurrasch et al., 2007), while those in the surrounding shell are mostly GABAergic (Choi et al., 2005; Hahn et al., 2019) and express *Gad1*, *Gad2*, *Tmem176a*, *Tmem176b*, *Six3*, *Dlk1*, and *Ecel1* (Figure 1B, upper). Only 13 of 40 neuronal T-types (n = 611 out of 4,473) expressed *Gad1*, *Gad2*, and *Slc32a1* (*Vgat*), reflecting the fact that our Cre driver lines were chosen to label cells within VMH (Figures 1C, S1F, and S1G).

Within VMH, two sub-populations could be defined: those in VMHc or anterior VMH, which expressed *Nr5a1*, *C1ql2*, *Rreb1*, *Six3*, and *Ldb2*; and those in VMHvl (Figure 1B, lower). There were few T-types from dorso-medial VMH (VMHdm), as our microdissection procedure deliberately selected against this region (Figure S1B, yellow outline). Within VMHvl, several classes of neurons could be identified, by expression of *Esr1*, *Dlk1* (Deltalike homolog 1; also called Pre-adipocyte factor 1, or *Pref-1*), and *Satb2* (Figures 1B, S2A, and S2C). Altogether, this analysis identified 17 different T-types predicted to be in VMHvl.

Anatomic Analysis of T-types by seqFISH

To place the T-types identified by SMART-seq in anatomical context with higher resolution, we performed sequential fluorescence in situ hybridization (seqFISH) (Eng et al., 2019; Shah et al., 2016). We applied 58 probes comprising DEGs that marked the different VMHvl SMART-seq clusters and carried out multiple sequential cycles of single-molecule (sm) FISH on sections covering multiple positions along the anterior-posterior axis (Figures 2A1 and 2G). Over 23,000 neurons were semi-automatically segmented within four regions of interest (ROIs) per section (Figure 2A2), and hybridization signals for each probe set were counted within each cell. Segmented cells were distributed across VMHvl, VMHdm/c, and surrounding regions (VMH-out; Figures 2A3 and S3).

We performed independent clustering of the seqFISH data (Figure 2B) and compared the results to the SMART-seq cluster analysis. In general, there was a high correlation between VMHvl SMART-seq clusters and seqFISH clusters (Figure 2C; 85.2% of seqFISH clusters correspond to at least one SMART-seq T-type), as well as of marker gene expression levels between the two datasets (Figures S3F and S3G), suggesting that transcriptomic profiles were not strongly altered by tissue dissociation or FACS. In general, the anatomical locations of cells revealed by seqFISH validated the provisional locations of SMART-seq clusters assigned by marker expression Figures 1 and S3). However, we found several cases where

cells in certain clusters ‘‘violated’’ cytoarchitectonic locations predicted by marker expression: (1) a small number of GABAergic cells assigned to VMH-out clusters by SMART-seq was detected just inside the lateral border of VMHvl (Figures S3A [VMH-out], S3B, and S3C), (2) a small number of cells assigned to VMHdm/c (based on expression of *Nr5a1* and *C1ql2*) were found just across the border in VMHvl (Figures S3A [VMHdm/c], S3D, and S3E), (3) some cells assigned to VMHvl were also detected in the adjacent VMH-out region (Figures S3A [VMHvl], S3D, and S3E). Interestingly, cells in several of the VMHvl seqFISH clusters exhibited different distributions along the anterior-posterior axis, e.g., those in seqFISH cluster #11 (Esr1_2,3) were located more posteriorly, while those in cluster #5 (Nr5a1_4,6) were located more anteriorly (Figure 2D, lower).

Neuronal Projections of VMHvl T-types

Classical neuroanatomical studies have indicated that VMHvl neurons, as a population, project to multiple (>20) brain regions (Canteras et al., 1994; Saper et al., 1976). The relationship between these projections and different VMHvl cell types is not yet clear. Recent studies using viral tracing in *Esr1-Cre* mice have identified two anatomically distinct subsets of VMHvl^{Esr1} neurons that collateralize preferentially to posterior versus anterior targets, such as the dorso-medial periaqueductal gray (dmPAG) versus the medial preoptic area (MPOA), respectively (Lo et al., 2019). To examine the relationship of VMHvl T-types to these projection-specific subsets, we performed scRNA-seq following retrograde labeling from MPOA or dorsal or lateral peri-aqueductal gray (dPAG and lPAG, respectively) (Retro-seq; see Methods; Figure 3A) (Tasic et al., 2018).

These data revealed groups of VMH T-types with clear projection preferences. Consistent with anterograde labeling studies (Kunwar et al., 2015; Oh et al., 2014; Wang et al., 2015), the *Nr5a1*$^{+}$ clusters, which are located primarily in VMHdm/c, showed a significant projection bias to the dPAG (Figure 3B). By contrast, the *Esr1*$^{+}$ and *Dlk1*$^{+}$ classes, which are located in VMHvl, showed a significant projection bias to the MPOA. This is not

surprising, as MPOA is one of the strongest projection targets of VMHvl (Canteras et al., 1994; Lo et al., 2019; Roeling et al., 1994; Yang et al., 2013). Surprisingly, although retrograde labeling in *Esr1-Cre* mice revealed mostly non-overlapping populations of MPOA- or dPAG-projecting VMHvl^{Esr1} neurons (Lo et al., 2019), most *Esr1*$^{+}$ T-types in VMHvl contained cells back labeled from both targets, indicating a probable mixture of cells with different projections. Nevertheless a few T-types exhibited a dPAG projection preference (#28/Esr1_4 and #39/Dlk1_5; Figures 3C and 3D). Cluster #31 (*Nup62cl*$^{+}$) was specifically labeled by injections to lPAG. Some T-types were underrepresented in the retrogradely labeled populations and may correspond to local interneurons or to projection biases not interrogated by our retrograde injection sites (Figure 3D, bars below dashed line).

Linking Transcriptomic Identity to Activity by Act-seq

Electrophysiology, *c-fos* catFISH *in situ* hybridization (Lin et al., 2011), and in vivo calcium imaging (Remedios et al., 2017) studies have revealed that VMHvl in males contains distinct, largely non-overlapping, populations of neurons that are active during social interactions with males versus females. In addition, different patterns of VMHvl neuronal activity are observed during mating versus aggression (Falkner et al., 2014; Hashikawa et al., 2017a; Remedios et al., 2017), as well as during social fear (Sakurai et al., 2016; Silva et al., 2013; Wang et al., 2019). We therefore wished to investigate how VMHvl activity during different social behaviors is related to T-types. Attempts to measure IEG transcript levels using SMART-seq in cells obtained from animals following different social behaviors were unsuccessful due to the low yield of cells per animal after FACS isolation. While this work was in progress, methods for behavioral IEG analysis using droplet-based scRNA-seq were described (Hrvatin et al., 2018; Wu et al., 2017). In order to apply one such method, Act-seq (Wu et al., 2017), to VMHvl, we first characterized cellular diversity in this region using the 10x Genomics droplet-based platform.

We microdissected VMHvl from acute slices, dissociated it using modified protocols designed to maximize cell viability and number of genes detected (A.H.P. and Y.O., in preparation), and immediately subjected the cell suspension to 10x preparation and sequencing without employing FACS. Out of 149,663 cells sequenced (n = 90 mice), 27.7% were VMH neurons (*Slc17a6*$^{+}$, *Fezf1*$^{+}$, and *Adcyap1*$^{+}$), 24.8% were non-VMH neurons (*Slc32a1*$^{+}$, *Gad1*$^{+}$, and *Gad2*$^{+}$), and 47.5% were nonneuronal cells (Figures S4B–S4I). Higher proportions of non-VMH neurons, compared with the SMART-seq dataset, are expected since FACS enrichment for VMH markers was omitted.

Analysis of the VMH neurons revealed 29 T-types and their DEGs (Figures 4E and S4L). Of these, 17 clusters (including one female-specific cluster; see below) were provisionally assigned based on marker expression to VMHvl (Figure 4E, red squares) and the remainder to VMHdm/c (Figure 4E, blue squares). Like SMART-seq, the 10x data revealed that VMHvl neurons fell into three distinct classes defined by *Esr1*, *Satb2*, and *Dlk1* (Figure 4D).

To determine how neural activation during different behaviors maps onto 10x-derived VMHvl T-types, we repeated the 10x analysis using tissue dissected from animals ~1 h after they engaged in one of eight different social behaviors (see Methods; Figures 5A and S4I). Close investigation (CI) was tested using intruders either suspended by the tail or confined in a mesh pencil cup. Initially, for each cluster and each behavior, we compared the percentage of *Fos*$^{+}$ cells and the average fold-change in *Fos* expression versus control (Figure 5B). This analysis revealed that 8 of 16 male VMHvl T-types (50%), in aggregate, displayed a statistically significant *Fos* induction during the behaviors tested. During most behaviors, one to three T-types were activated (with the exception of social fear in single-housed males; seven T-types), and each of those T-types was typically activated during multiple behaviors. However, a few T-types showed more restricted patterns of activation. For example, cluster Nup62cl^{10x} was activated only during social fear, while cluster Dlk1_1^{10x} was only activated during mating (Figure 5B).

In addition to exhibiting higher *Fos* expression during a particular behavior in comparison with home-cage control animals, some T-types also exhibited significant differences in *Fos*

expression between different behaviors (Figure S5A, rows, colored squares). For example, cluster Esr1_7^{10x} was activated to a significantly greater extent during aggression than during male-male (M-M) CI (suspended by tail), social fear (group-housed mice), male-female (M-F) CI, or mating, while cluster Dlk1_1^{10x} was activated at higher levels during mating than during aggression or M-M CI (pencil cup). Seven T-types were activated during social fear in single-housed mice, but only three were activated during social fear in group-housed mice (Figure 5B). This difference is likely to reflect defensive aggression, which is initially displayed by most single- but not group-housed mice.

Previous studies have shown that distinct, partially overlapping subsets of VMHvl neurons are activated in males during fighting versus mating, using both *Fos* cellular compartment analysis of temporal activity by fluorescence *in situ* hybridization (catFISH) (Lin et al., 2011) and calcium imaging of GCaMP6 expressed in *Esr1-Cre* mice (Remedios et al., 2017). While cluster Esr1_7^{10x} was activated during M-M but not M-F social interactions, none of the other *Esr1*$^{+}$ clusters showed significant *Fos* induction during mating relative to controls (Figure 5B). However, statistical correction for multiple comparisons in this analysis might have obscured the relatively weaker *Fos* induction during mating (Lee et al., 2014) in some *Esr1*$^{+}$ clusters. To investigate this possibility, we performed a statistical contrast between combined M-F social interaction tests (mating and two M-F CI tests), M-M interaction tests (aggression and two M-M CI tests), and controls among the different *Esr1*$^{+}$ clusters. Because Act-seq allows simultaneous analysis of multiple IEGs (Wu et al., 2017), we examined *Fos*, *Fosl2*, and *Junb* expression. This analysis indicated that cluster Esr1_5^{10x} exhibited significantly greater *Fosl2* induction than cluster Esr1_7^{10x} following M-F social interactions (Figure 5C, *Fosl2*, blue data points), while conversely Esr1_7^{10x} showed significantly greater *Fosl2* induction than cluster Esr1_5^{10x} following M-M social interactions (Figure 5C, red data points). Cluster Esr1_4^{10x} was activated during both types of interactions. Similar trends were observed for *Fos* and *Junb* expression.

Systematic examination of 139 different IEGs (Bravo, 1990; Wu et al., 2017) revealed different patterns of IEG activation in different clusters during different behaviors (Figures

5D and S5B). Cluster Esr1_7^{10x} exhibited activation of multiple IEGs during aggression and M-M CI tests, cluster Dlk1_1^{10x} showed the most IEGs induced during mating, and cluster Esr1_4^{10x} was activated to varying extents during all eight behaviors (Figure 5D).

We investigated next whether VMHvl T-types activated during social behaviors were also distinguished by other classes of non-activity-dependent DEGs. Interestingly, certain classes of DEGs were almost as effective, as the entire set of annotated mRNAs at classifying these behaviorally relevant clusters (Figure S6A), including the I-set domain-containing group of immunoglobulin (Ig) superfamily cell surface proteins (Figure S6B, mean area under receiver-operator characteristic [AUROC] curve = 0.92 ± 0.05), while others including neuropeptides, Class A orphan G protein-coupled receptors (GPCRs), and transcription factors were less discriminating (Table S3).

Alignment of SMART-seq and 10x Datasets by CCA

Next, we sought to determine whether there was a three-way correspondence between T-type identity, behavioral activation. and projection specificity. As the latter two measures were carried out using different scRNA-seq platforms, we first aligned the SMART-seq and 10x datasets by canonical correlation analysis (CCA; Figure S7) (Butler et al., 2018; Stuart et al., 2019). Co-clustering of the joint dataset revealed 31 CCA VMH-in clusters (Figure S7C). Both the 10x and SMART-seq clusters were relatively well-aligned with the joint CCA clusters, with some exceptions (Figures S7D, S7E, and S7G).

We then computed the hierarchical relationship between and compared the expression of different marker genes among these joint CCA clusters (Figure 6A). 17 glutamatergic CCA clusters were predicted to be in VMHvl and the remainder in VMHdm/c, based on the expression of *Fezf1*, *Adcyap1*, *Slc17a6*, and *Nr5a1* (see above). We found excellent agreement between the cluster markers identified by 10x versus SMART-seq (Figure 6A).

VMHvl neurons again fell into 3 distinct subsets marked by *Esr1*, *Satb2*, and *Dlk1* (Figures 6A and S7H).

Identification of Sexually Dimorphic CCA Clusters in VMHvl

We annotated seven joint CCA clusters as ‘‘*Esr1*$^{+}$’’ (Figures 6A and S7H). One of these seven clusters, which expressed higher levels of Esr1 than the others, was female specific (Tsix_Esr1^{CCA3} Figures 6B, 6C [this cluster was split into two subclusters by SMART-seq], and S7G). Conversely, clusters Esr1_1^{CCA4} and Esr1_2/3^{CCA5} (Figures 6B and 6C) were strongly enriched in males; the latter expressed Moxd1 (Figure 6B), which encodes an enzyme expressed dimorphically in the male hypothalamus (Tsuneoka et al., 2017). To validate these results, we performed seqFISH analysis for additional markers identified by scRNA-seq in males and females using *Esr1*/*Gpc3* or *Gldn* as reference markers (Figures 6B and 6F). The results confirmed the differential expression of these additional markers in these sexually dimorphic cell types and indicated a posterior bias in the anatomic location of their cell somata (Figures 6E and 6F).

A previous study demonstrated that female VMHvl contains two distinct subpopulations of Esr1+ neurons: one, activated during mating, is located in a subregion called VHMpvll; the other, activated during aggression, is located in subregion VMHpvlm (Hashikawa et al., 2017a). Marker genes for these two anatomic subdivisions were identified using bulk RNA-seq (Hashikawa et al., 2017a). Comparison of those marker genes with our scRNA-seq dataset revealed that the female-specific cluster Tsix_Esr1^{CCA3} expressed most of the markers spatially enriched in VMHpvll (mating region) (Figure 6D), while cells in the *Nr5a1*$^{+}$ and *Dlk1*$^{+}$ classes expressed most of the markers enriched in VMHpvlm (aggression region).

Relationship between Behavioral Activation, Projection specificity and Transcriptomic Identity

We next sought to correlate transcriptomic phenotype, behavioral function, and projection bias (Economo et al., 2018; Tasic et al., 2018). Although it was necessary to measure projection bias and behavioral activation using two different scRNA-seq platforms (SMART-seq and 10x, respectively), the correspondence provided by CCA (Stuart et al., 2019) allowed us to correlate the two phenotypic measures. Surprisingly, we did not observe a clear 1-to-1-to-1 relationship between T-type identity, behavioral activation, and projection bias for most clusters (Figures 7A and 7E). A rare exception was T-type Nup62cl, which was selectively activated during social fear (in group-housed mice), and which selectively projected to lPAG (Figure 7A, yellow), a known circuit node controlling freezing behavior in mice (Tovote et al., 2016). Similarly, another social fear-activated T-type, Dlk1_4^{CCA18}, also preferentially projected to lPAG (Figure 7A). In general, there was a positive correlation ($r = 0.66$) between the proportions of cells projecting to lPAG and stronger *Fos* expression induced by social fear (Figures 7C and 7E).

We also correlated preferential activation during M-M versus M-F social interactions with projection bias. A positive correlation was observed between CCA T-types that were preferentially activated during all M-M versus all M-F social interactions (e.g., Esr1_7^{CCA9}; Figure 7A, all M-M), and that exhibited a relative bias for projections to dPAG versus MPOA based on Retro-seq data (Figure 7B; $r = 0.71$). By contrast, cluster Esr1_5^{CCA7}, which exhibited significantly greater Fosl2 inductions following M-F versus M-M social interactions (Figure 7A), did not show any projection bias.

To independently investigate the relationship between social behavior and projection bias for VMHvl^{Esr1} neurons, we performed experiments combining retrograde tracing from dPAG versus MPOA with anti-*Fos* immunostaining of *Esr1*$^{+}$ neurons following aggression, mating, or M-F CI (see Methods). The results of these experiments indicated that VMHvl^{Esr1} cells activated during aggression showed a significant bias for projection to dPAG versus MPOA, whereas VMHvl^{Esr1} cells activated following M-F interactions showed no such bias (Figure 7D).

These latter observations suggested that we should find a CCA cluster preferentially activated during aggression, which contained only dPAG-projecting cells; however, we did not. Cluster Esr1_7^{CCA9} showed strong, specific IEG activation during aggression (Figures 5D and 7A) but corresponded to a SMART-seq cluster (Esr1_4 Smoc2^{SS}) that contains dPAG- as well as MPOA-projecting cells by Retro-seq (Figure 3C). However, Esr1_4 Smoc2^{SS} was mapped by CCA not only to Esr1_7^{CCA9} but also to Esr1_4^{CCA6} (Figure S7G); the latter cluster did not show any preferential activation during aggression versus other social behaviors (Figures 5D and 7A). If the dPAG-projecting and MPOA-projecting cells in cluster Esr1_4^{SS} are indeed distinct (Lo et al., 2019), then the dPAG projection-specific cells in that cluster could correspond to those in the aggression-activated cluster Esr1_7^{CCA9}. If so, then Esr1_7^{CCA9} might indeed represent an aggression-activated T-type that preferentially projects to dPAG and other brainstem structures (Lo et al., 2019). Notably, this cluster was slightly enriched in males versus females (Figure 6B, #9).

2.4 Discussion

The hypothalamus contains ~65 nuclei or regions per hemisphere (Hahn et al., 2019). Earlier scRNA-seq studies identified ~30–60 T-types in the entire hypothalamus (Chen et al., 2017; Romanov et al., 2017). Recently, Moffitt et al. (2018) reported ~70 cell types in the POR, which contains ~20 nuclei (a density of 3–4 cell types per nucleus). Our results reveal ~30 T-types in VMH and 17 T-types in VMHvl alone. Similarly, ~35 T-types were identified in the arcuate nucleus, which controls feeding and is of comparable size to VMH (Campbell et al., 2017). These results suggest that the total number of hypothalamic cell types has been significantly underestimated.

Our results demonstrate good agreement between T-types identified by the 10x and SMART-seq platforms. We identified three major transcriptomic classes in VMHvl. One is defined by *Esr1* (Morrell and Pfaff, 1982), the other two by *Dlk1*/*Pref-1* (Persson-Augner et al., 2014) and *Satb2*, respectively. The *Esr1*$^{+}$ class, which expresses the progesterone receptor (*Pgr*) and controls social behaviors (Lee et al., 2014; Yang et al., 2013; Yang et al., 2017), contains seven different T-types, while the *Dlk1*$^{+}$ class contains six. Three Satb2 T-types and Nup62cl comprise the remaining clusters. While we identified some VMHdm/c T-types, which contaminated our VMHvl dissection and which include *Sf1*/*Nr5a1*$^{+}$ neurons, a comprehensive census of this regions remains to be performed.

Using a combination of previously defined and newly discovered markers, we performed seqFISH on VMHvl in vivo. Independent clustering of the seqFISH data confirmed the cluster diversity revealed by SMART-seq and 10x—only 4 of 27 seqFISH clusters (15%) did not map to any scRNA-seq T-types. Several of the seqFISH clusters exhibited differences in cellular distribution along the A-P axis of VMHvl, consistent with earlier observations (Lo et al., 2019; Wang et al., 2019). Our results are also suggestive of heterogeneity along the M-L axis of VMHvl (Hashikawa et al., 2017a). Thus, some T-types appear spatially restricted within VMHvl.

Male- and Female-Enriched *Esr1*$^{+}$ Neuronal Cell Clusters in VMHvl

Sex-specific neurons are well established in *Drosophila* (Cachero et al., 2010; Yu et al., 2010) and *C. elegans* (Liu and Sternberg, 1995), but have not previously been reported in mice. While quantitative sex differences in gene expression have been described in rodents (Xu et al., 2012), including at the single-cell level (Welch et al., 2019), our results provide initial evidence of sex-specific neuronal T-types in the mammalian CNS. Importantly, these sex-specific T-types are not defined simply by sex-chromosome-encoded sexually dimorphic genes (e.g., *Tsix*, which is also expressed in most non-sexually dimorphic T-types in females). Rather, they are defined by specific patterns of covariation of autosomal gene expression. scRNA-seq analysis at high sampling density is likely to reveal more examples of such rare, sex-specific populations in other sexually dimorphic brain regions (Yang and Shah, 2014).

Foundational studies in the rat have shown that VMHvl^{Esr1} neurons control sexual receptivity (lordosis) in females (Pfaff, 2017; Pfaff and Sakuma, 1979a; b) and were the first to implicate this region and genetic marker in a social behavior. Comparison of our data with bulk RNA-seq data (Hashikawa et al., 2017a) and preliminary Act-seq experiments suggest that the female-specific cluster Tsix_Esr1 may be specifically activated during mating. Functional studies will be required to confirm whether this cell type indeed controls lordosis behavior.

Relationship of Transcriptomic Identity to Projection Heterogeneity in VMHvl

Our results indicate a clear correspondence with projection bias for groups of T-types in VMHdm/c versus those in VMHvl—the former preferentially project to dPAG and lPAG, while the latter project preferentially to the MPOA. This may reflect the fact that VMHvl controls social behaviors, many of which are also controlled by neurons in MPOA (McHenry et al., 2017; Sano et al., 2013; Simerly, 2002; Wei et al., 2018; Wu et al., 2014). In contrast,

VMHdm/c mainly controls anti-predator (non-social) defensive behaviors (Kunwar et al., 2015; Silva et al., 2013; Wang et al., 2015), although one subpopulation in this region controls female sexual receptivity (Ishii et al., 2017).

The correspondence between projection bias and T-types within VMHvl is less clear. We recently identified two non-overlapping subpopulations of VMHvl^{Esr1} neurons which exhibit a strong projection bias to posterior (e.g., dPAG) versus anterior (e.g., MPOA) targets, respectively (Lo et al., 2019). Cells from these two subpopulations also have different cell body sizes and are differentially distributed along the A-P axis. Surprisingly, however, most *Esr1*$^{+}$ T-types contained neurons retrogradely labeled from both dPAG and MPOA, indicating that a single T-type may contain multiple projection-specific types. This may be explained by the transient developmental expression of genes that determine projection specificity as seen in the *Drosophila* antennal lobe (Li et al., 2017). Alternatively, subtle differences in adult gene expression between projection-specific cells within a given T-type may not be detectable by the methods used here.

Functional Heterogeneity among VMHvl Neurons during Social Behaviors

VMHvl neurons have been functionally implicated in the control of aggression, mating, social fear and the encoding of conspecific sex (reviewed in Hashikawa et al., 2017b; Kennedy et al., 2014; Yang and Shah, 2014). In particular, optogenetic activation of VMHvl^{Esr1} neurons can promote mating or aggressive behavior in both males (Lee et al., 2014) and females (Hashikawa et al., 2017a). Our Act-seq data indicated that both distinct and common *Esr1*$^{+}$ T-types were activated during aggression or mating. The common types may control behavioral elements shared by mating and aggression (e.g., CI, chasing). However, calcium imaging of VMHvl^{Esr1} neurons indicated that different populations are activated during male versus female CI; indeed most (50%–60%) of the variance in population activity is explained by intruder sex, with only 10% explained by behavior (Remedios et al., 2017). Consistent with this, distinct T-types were preferentially activated

following M-M versus M-F social interactions rather than following specific behaviors. Nevertheless, we identified one *Esr1*$^{+}$ T-type that was more strongly activated during aggression than during other behaviors. We also identified a *Dlk*$^{+}$ T-type specifically activated during mating and several T-types specifically activated during social fear (Sakurai et al., 2016; Silva et al., 2013; Wang et al., 2019).

A recent study of *Fos* activation in the POR by MERFISH concluded that ‘‘genetically encoded circuits comprised of transcriptionally distinct neuronal cell types control specific hypothalamic functions’’ (Moffitt et al., 2018). The apparent difference in conclusions between that study and the present one may reflect the higher diversity of T-types in VMHvl and/or methodological differences. Alternatively, it is possible that the relationship between cell types and behavior encoding is different for nuclei in the POR, which are largely GABAergic (Moffitt et al., 2018), and VMHvl, which is primarily glutamatergic. Answering this question will require correlating transcriptomic identity with activity on a millisecond time-scale, e.g., using calcium imaging followed by gene-specific labeling (Lovett-Barron et al., 2017).

Importantly, VMHvl neurons (including the *Esr1*$^{+}$ subset) also control metabolic function (Musatov et al., 2007; Xu et al., 2011); reviewed in Krause and Ingraham, 2017), and there is evidence that *Esr1*$^{+}$ cell populations controlling reproductive behavior and metabolism can be genetically distinguished in females (Correa et al., 2015). The relationship between the VMHvl transcriptomic cell types described here and metabolic function remains to be explored, although cells in the *Dlk1*$^{+}$ class are good candidates, given data implicating this gene in obesity in mice and humans (Moon et al., 2002; Wermter et al., 2008).

Conclusion

The data presented here constitute one of the few attempts to link transcriptomic diversity in the CNS to projection specificity and behavioral or physiological function (see also Economo

et al., 2018). With a few rare exceptions, we generally do not observe a clear, 1-to-1-to-1 correspondence between T-types, projection specificity, and behavioral function. It is possible that most VMHvl T-types reflect facets of cell identity (or cell state) distinct from projection specificity and behavioral function. Alternatively, identifying correspondences between T-types, connectivity, and behavioral activation may require methods with higher spatial and temporal resolution than those employed here. Application of such methods should be facilitated by more specific genetic access to VMHvl subtypes using intersectional transgenic models based on the expression of markers identified in this study.

2.5 Main figures

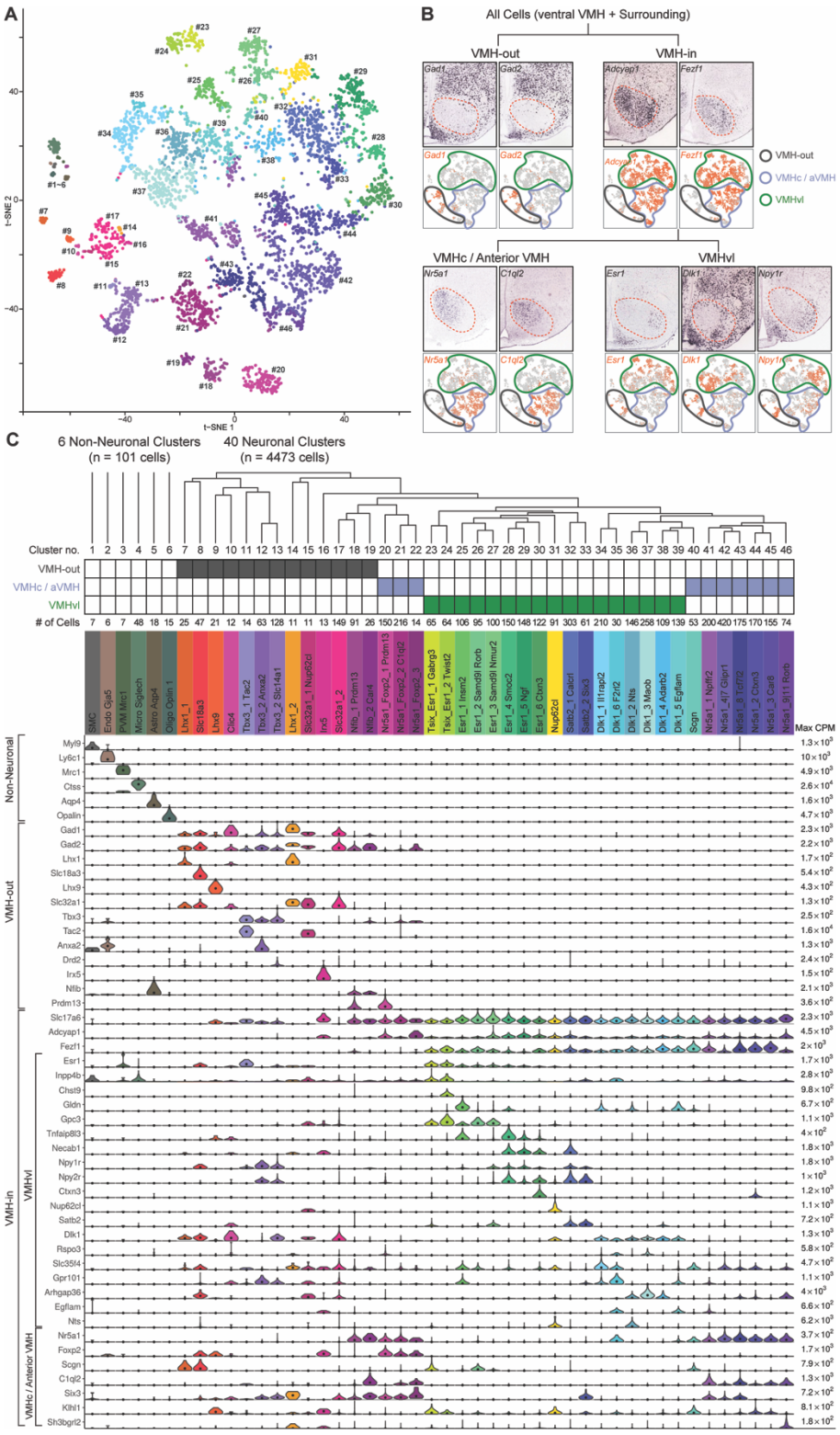

Figure 1. VMHvl Transcriptomic Cell Types Revealed by SMART-seq. (A) Two-dimensional t-distributed stochastic neighbor embedding (t-SNE) plot color-coded by 46 SMART-seq clusters (N = 4,574) from ventral VMH. (B) Expression patterns of major marker genes in VMH (red dashed circle) and surrounding areas by ISH (Allen Mouse Brain Atlas; upper) arranged by anatomical hierarchy; orange points on t-SNE plots (lower) indicate their expression levels. Anatomic location of clusters is outlined on t-SNE plots (VMH-out: black; VMHvl: green; VMHc or anterior VMH: light blue). (C) Violin plots illustrating expression levels of marker genes by cluster; ‘‘max CPM’’ (right), maximum counts per million reads. Dendrogram and matrix show relatedness between clusters and their spatial locations (see B), respectively. See also Figures S1 and S2.

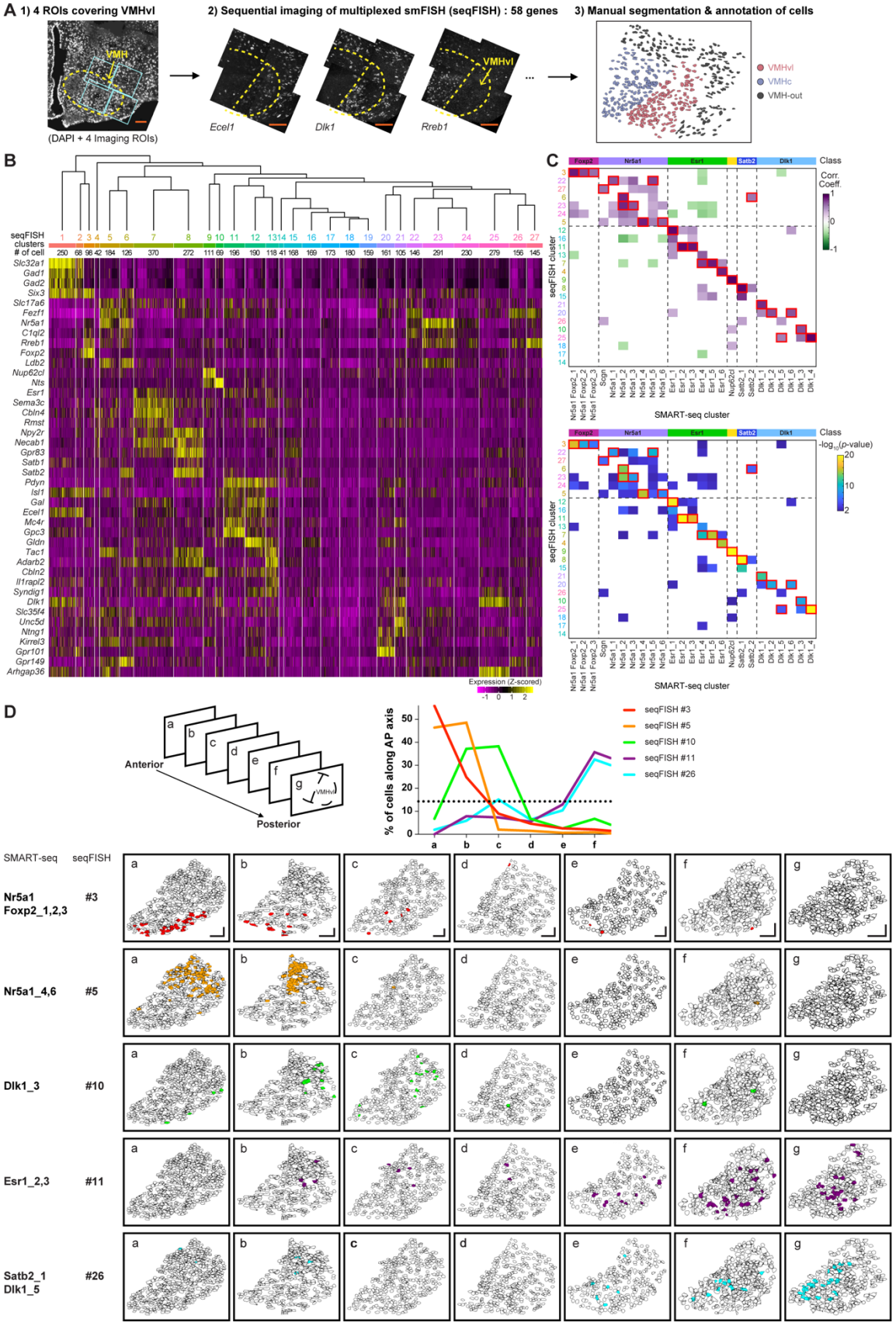
A
1) 4 ROIs covering VMHvl
2) Sequential imaging of multiplexed smFISH (seqFISH) : 58 genes
3) Manual segmentation & annotation of cells
VMH
(DAPI + 4 Imaging ROIs)
Ecel1
Dlk1
Rreb1
VMHvl
VMHvl
VMHc
VMH-out
B
seqFISH clusters
# of cell
Slc32a1
Gad1
Gad2
Six3
Slc17a6
Fezf1
Nr5a1
C1ql2
Rreb1
Foxp2
Ldb2
Nup62cl
Nts
Esr1
Sema3c
Cbln4
Rmst
Npy2r
Necab1
Gpr83
Satb1
Satb2
Pdyn
Isl1
Gal
Ecel1
Mc4r
Gpc3
Gldn
Tac1
Adarb2
Cbln2
Il1rapl2
Syndig1
Dlk1
Slc35f4
Unc5d
Ntng1
Kirrel3
Gpr101
Gpr149
Arhgap36
Expression (Z-scored)
C
Class
Corr. Coeff.
seqFISH cluster
SMART-seq cluster
-log10(p-value)
D
Anterior
Posterior
VMHvl
% of cells along AP axis
seqFISH #3
seqFISH #5
seqFISH #10
seqFISH #11
seqFISH #26
SMART-seq
seqFISH
Nr5a1
Foxp2_1,2,3
#3
Nr5a1_4,6
#5
Dlk1_3
#10
Esr1_2,3
#11
Satb2_1
Dlk1_5
#26

Figure 2. Anatomic Distribution of VMH scRNA-seq Clusters by seqFISH. (A) Schematic of seqFISH procedure in VMH. Light blue solid lines (1) outline ROIs of sequential hybridizations; yellow dashed lines (2) (maximum intensity Z projections) outline VMH and VMHvl. Three major anatomic regions (VMHvl, VMHc, and VMH-out) are color coded (3). Scale bars, 100 μm (inset). (B) Heatmap showing expression level of marker genes (rows) in 27 seqFISH clusters (columns; n = 4,497; VMHvl only). (C) Heatmaps showing correlation between seqFISH and SMART-seq clusters (n = 3,824; VMH only) and their p values (white, $p > 0.05$ or $r < 0.35$; see Methods). seqFISH clusters that are the most strongly correlated with each scRNA-seq cluster are marked by red squares. (D) Spatial distribution along A-P axis of seqFISH clusters (color-filled) showing anterior (#3, #5, #10) or posterior (#11, #26) biases, projected onto all segmented VMHvl cells (a–f). Line plot shows quantification for indicated clusters; black dotted line shows chance distribution. Scale bars, 50 μm (inset). See also Figure S3.

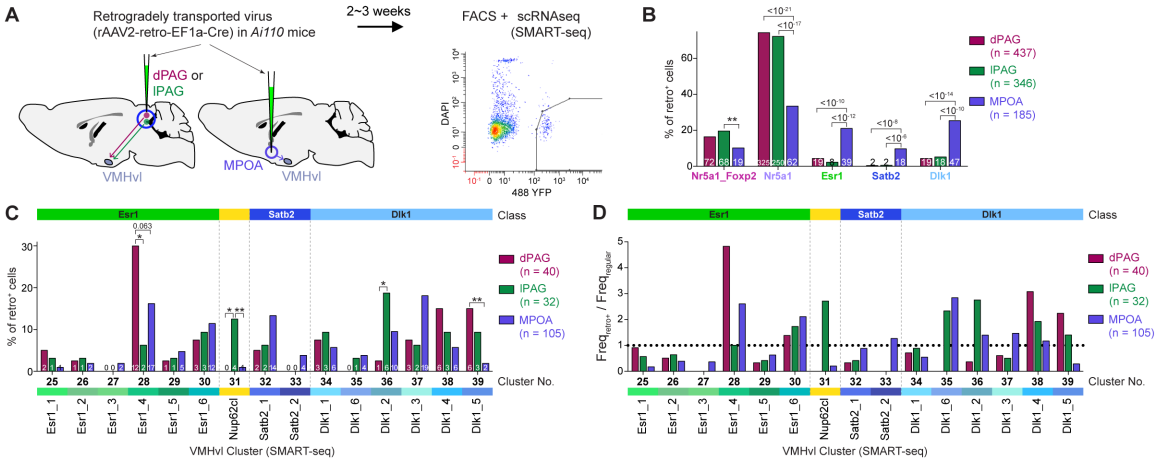

Figure 3. Anterior- vs. Posterior-projection Biases of Neurons in scRNA-seq Clusters. (A) Schematic of Retro-seq procedure (see Methods) from dPAG, lPAG, or MPOA. (B) Bar plot shows the distributions of cells retrogradely labeled from each target for the major VMH classes. (C and D) Percent of retrogradely labeled cells (C) and their relative frequency compared to non-Retro-seq samples (D) in each VMHvl SMART-seq cluster. Numbers of retro^{+} cells sequenced are listed at the bottom of the graphs. *$p < 0.05$, **$p < 0.01$ (Fisher's exact test).

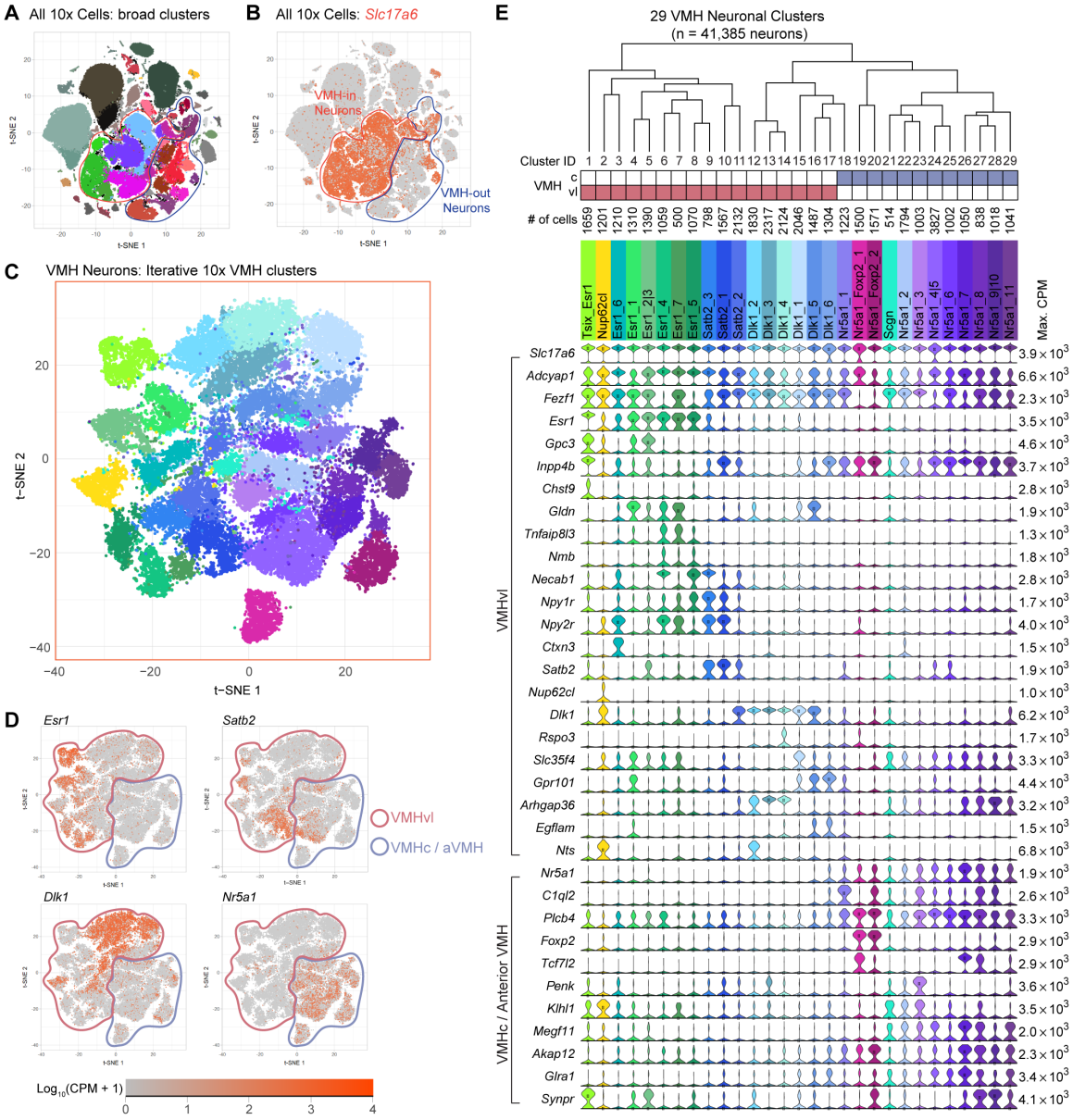

Figure 4. VMHvl Transcriptomic Cell Types Revealed by 10x scRNA-seq. (A and B) t-SNE plots showing clusters (A) and their *Vglut2* (*Slc17a6*) expressions (B) for all cells analyzed (N = 149,663). All neuronal clusters (n = 78,476) are outlined in either red (VMH-in) or blue (VMH-out), respectively. (C) t-SNE plot illustrating results of iterative clustering of VMH-in cells (n = 41,385; 29 clusters). (D) Expression levels of four main class-specific marker genes (*Esr1*, *Satb2*, *Dlk1*, and *Nr5a1*) are color coded (orange) on t-SNE plots; VMH subdivisions are outlined in different colors (VMHvl: light red; VMHc or anterior VMH: light blue). (E) Violin plots show differential

expressions of marker genes (VMHvl versus VMHc-enriched; indicated by left brackets) among 29 VMH clusters (C) with their spatial locations (top matrix). See also Figure S4.

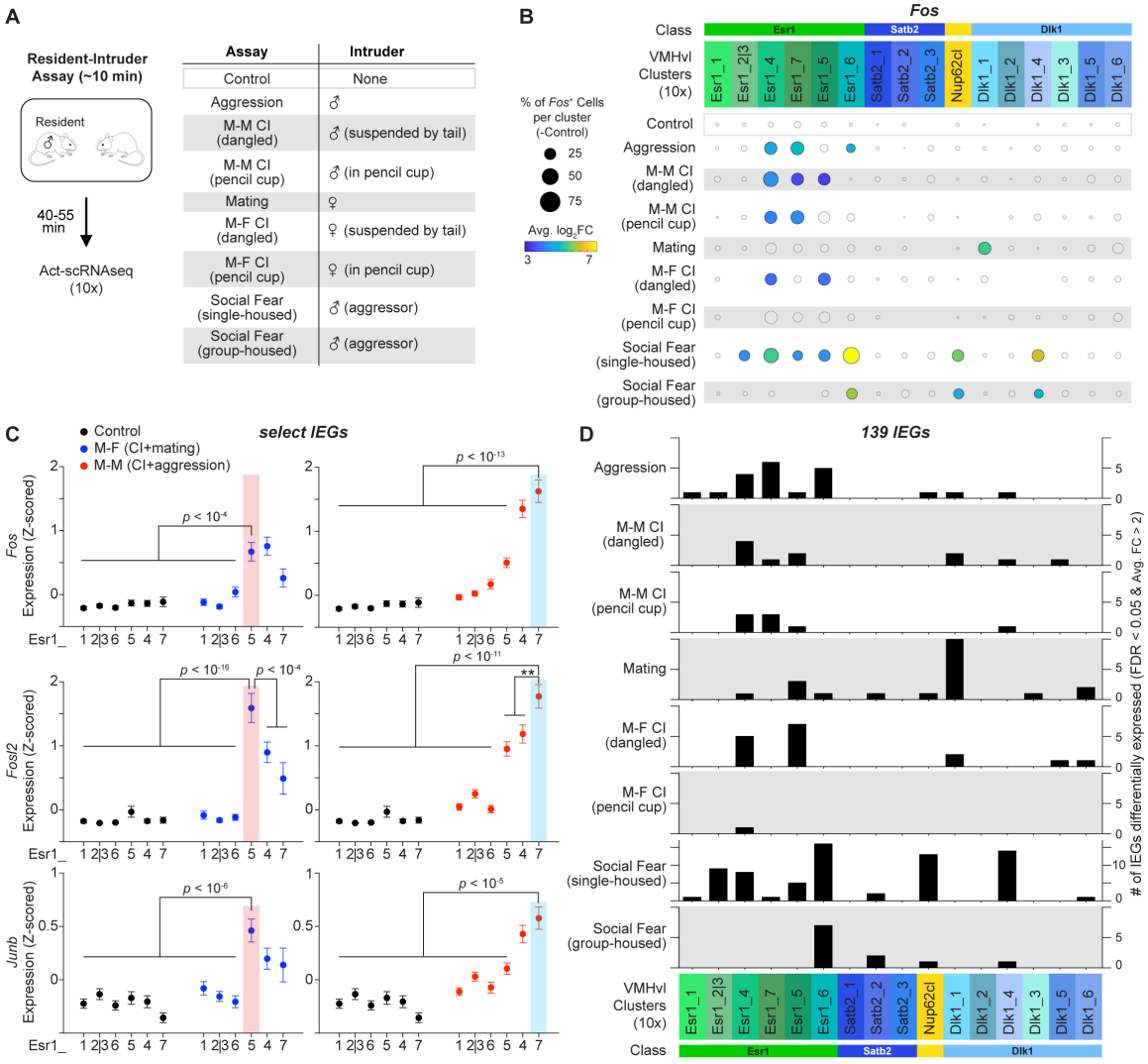

Figure 5. Activation of Neurons in VMHvl Clusters during Different Social Behaviors. (A) Schematic of Act-seq protocol and summary of behavioral assays (see Methods; Figure S4I). (B) Dot plots illustrating *Fos* induction in 10x VMHvl clusters. Colored and shaded gray dots indicate clusters with a significant ($p < 0.05$ after multiple comparison corrections across behaviors and clusters; three-way ANOVA and Bonferroni post hoc test), and non-significant induction ($p > 0.05$) versus control (or clusters from control animals), respectively. Dot size indicates proportion of *Fos*$^+$ cells per cluster after subtraction of control values; color indicates average $\log_2$fold-change (FC) versus control. (C) Expression levels of *Fos*, *Fosl2*, and *Junb* in 6 Esr1 clusters (male only). The clusters showing the highest induction of *Fos*, *Fosl2*, or *Junb* by either M-M (aggression plus M-M CI tests; red points) or M-F (mating plus M-F CI tests; blue points) versus control (black points) are highlighted by red (left;

Esr1_5) or blue (right; Esr1_7) shading, respectively (**p < 0.01, ***p < 0.001; three-way ANOVA and Bonferroni post hoc test; data are represented as mean ± SEM). (D) Bar graphs showing the number of different IEGs significantly induced during each social behavior in each cluster (see also Figure S5B) compared with control animals, where average FC > 2 and false discovery rate (FDR) <0.05. See also Figures S5 and S6.

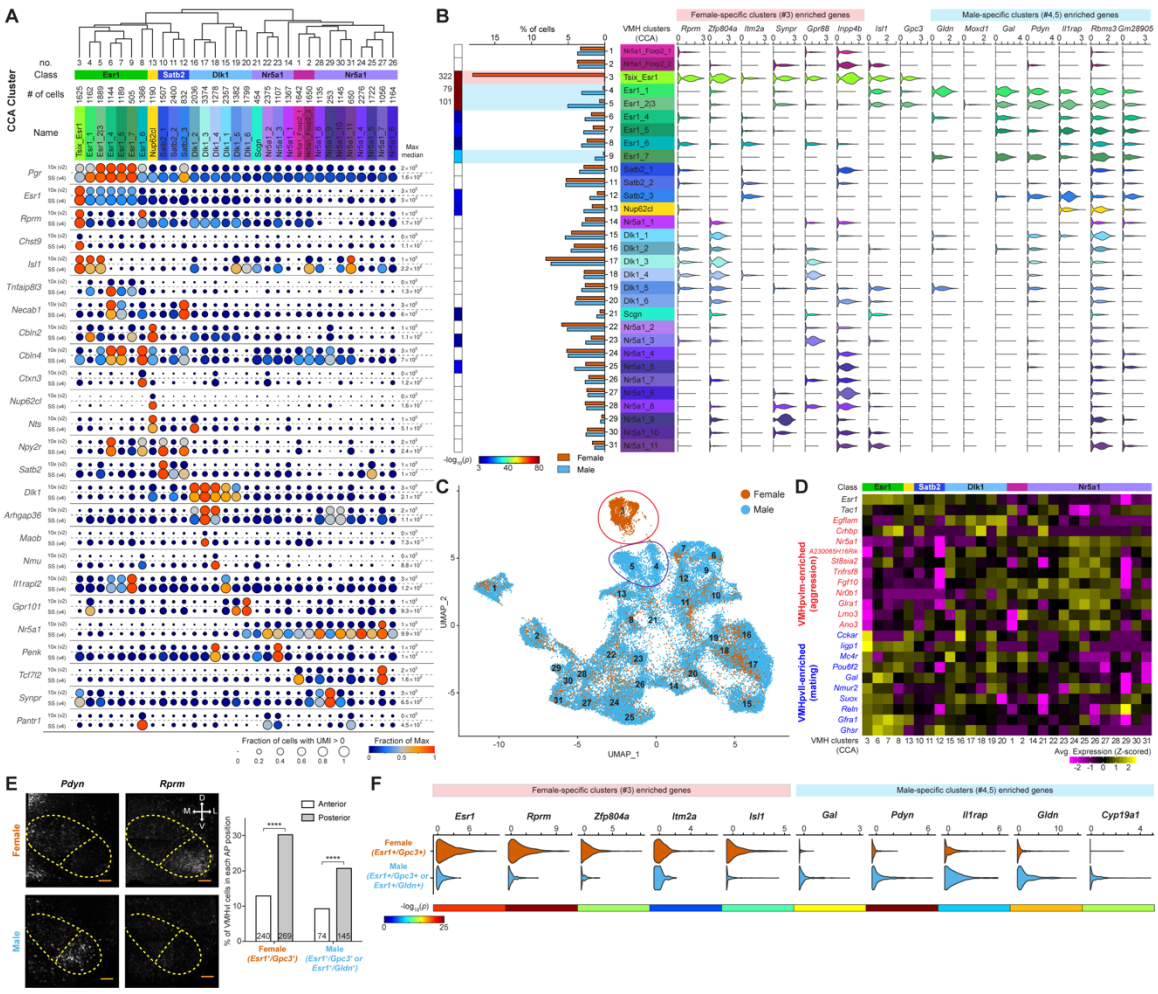

Figure 6. CCA Clusters and Sexually Dimorphic Cell Types in VMHvl. (A) Dot plots illustrating marker gene expressions in joint CCA clusters for SMART-seq (bottom) and 10x (top) datasets. Dot size and color indicate proportion of expressing cells and average expression level in each cluster, respectively. (B) Bar graph (left) represents the fractions of cells in clusters by sex (orange-red, female; sky blue, male). Far left, log-scaled p values (-log10p) are color coded (Fisher's exact test; white, $p > 0.001$). Violin plots show expression of genes specifically enriched in either female-specific (#3) or male-specific (#4, #5, #9) Esr1 clusters. (C) t-SNE plots of VMH cells from males (sky blue) and females (orange-red) illustrating distribution across joint CCA clusters. (D) Heatmap showing average expression in female VMH CCA clusters (n = 8,793) of genes previously identified by bulk RNA-seq (Hashikawa et al., 2017a) as localized to aggression-activated (red) or mating-activated (blue) VMHvl subdivisions in females (VMHpvlm versus VMHpvll). (E) Examples of seqFISH images (left; maximum intensity Z projections) for *Pdyn* and *Rprm*, major marker genes

differentially expressed between male and female VMHvl. Scale bars, 100 μm (inset). Bar graph (right) showing spatial distribution patterns (along A-P axis) of female *Esr1*$^{+}$ and *Gpc3*$^{+}$ or male *Esr1*$^{+}$ and *Gpc3*$^{+}$ or *Esr1*$^{+}$ and *Gldn*$^{+}$ cells in VMHvl (****$p < 0.0001$; Fisher's exact test). (F) Violin plots showing differential expressions of 10 genes (previously identified by scRNA-seq) between female- (*Esr1*$^{+}$/*Gpc3*$^{+}$) and male- (*Esr1*$^{+}$/*Gpc3*$^{+}$ or *Esr1*$^{+}$/*Gldn*$^{+}$) specific cell types in seqFISH. Log-scaled p values (-log10p) are color coded (below; unpaired t test). See also Figure S7.

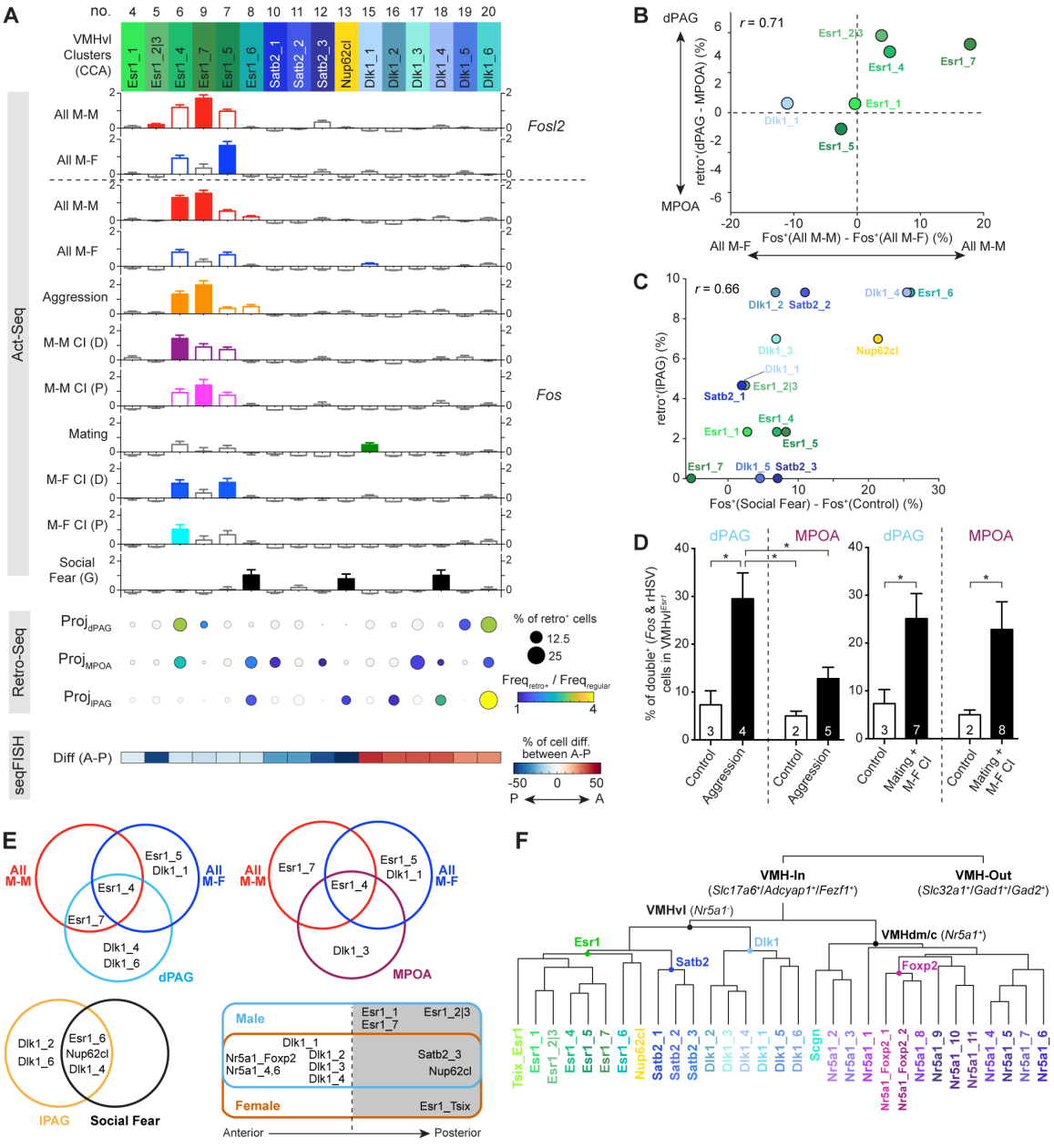

Figure 7. Relationship between Behavioral Activation, Projection Specificity, and Transcriptomic Identity in the CCA Framework. (A) Bar graphs, dot plots, and heatmap illustrating Act-seq (top; Figure 5B), Retro-seq (middle; Figures 3C and 3D), and seqFISH (bottom; Figure 2D) data projected onto VMHvl CCA clusters, respectively. Act-seq: bars with colored and gray outlines indicate significant ($p < 0.05$ after multiple-comparison corrections across behaviors and clusters; three-way ANOVA and Bonferroni post hoc test), and non-significant ($p > 0.05$) differences in IEG expression versus control, respectively. Filled and open bars indicate significant

and non-significant differences in IEG expression versus other clusters within a given behavior or behavioral category, respectively. All M-M and all M-F indicate combined data from all male-male or male-female social behaviors, respectively, for both *Fos* and *Fosl2* expression. Lower bar graphs illustrate data for individual behaviors (*Fos* only). Retro-seq: dot size indicates the percent of retrogradely labeled VMH cells from dPAG, lPAG, or MPOA (middle); dot colors indicate relative ratio (retrogradely labeled divided by total sequenced populations) of clusters. Shaded gray dots indicate clusters with ratio <1). seqFISH: ‘‘Diff (A-P)’’ indicates percent difference in cell number between anterior versus posterior VMHvl (color scale) for cells in corresponding seqFISH clusters. (B and C) Scatterplots showing correlation between preferential *Fos* activation during all M-M versus all M-F behaviors (x axis) and projection preference between dPAG versus MPOA (y axis) for major Act-seq clusters (B) and correlation between preferential *Fos* activation during social fear (using group-housed mice) versus control (x axis) and proportion of cells projecting to lPAG (y axis) for VMHvl CCA clusters (C).

2.6 Supplementary figures

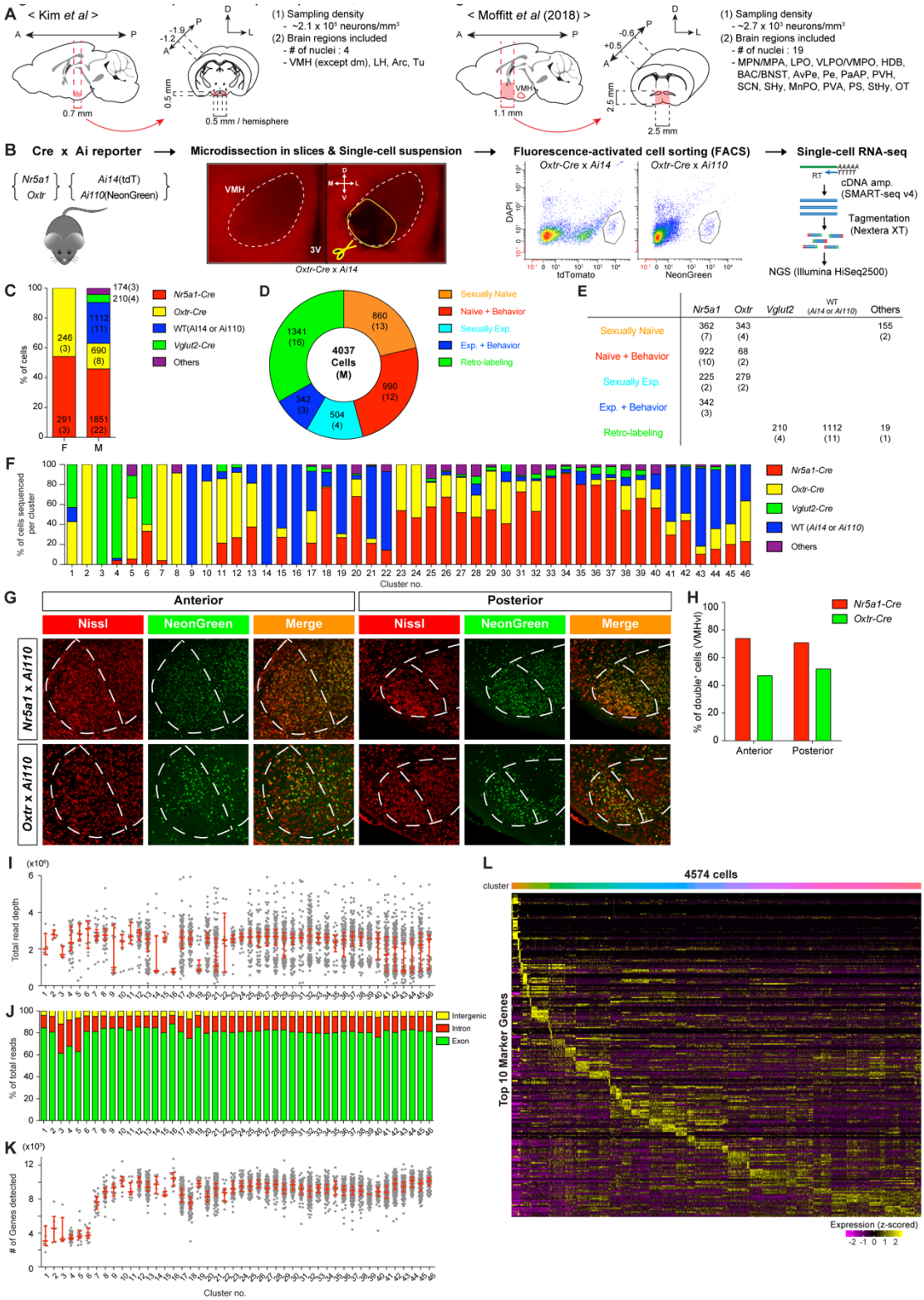

Figure S1. SMART-seq scRNA-seq Sample Info, Related to Figure 1 and Table S1. (A) Schematic diagrams illustrating relative size of region analyzed in this study versus POR (Moffitt et al., 2018) (see also Table S1). (B) Overview of sample collection protocol for SMART-seq. Two main Cre recombinase lines (*Nr5a1* and *Oxtr*) were crossed with Cre-dependent reporter mice (*Ai14* or *Ai110*) for cell labeling. A region including the lower half of VMH was manually micro-dissected from acute brain slices (yellow circle in the epifluorescence image). Microdissected tissue was treated with protease and triturated to generate a single-cell suspension. Single cells were isolated by FACS (fluorescence-activated cell sorting) with optimized gates (high reporter fluorescence and low DAPI). mRNAs in each cell were reverse-transcribed, and their full-length cDNAs were amplified using SMART-seq v4. Sequencing libraries were prepared by Nextera XT (Illumina), and sequenced on an Illumina HiSeq 2500. (C) Bar plot showing proportion of cells sequenced from each Cre line in each sex (n = 537 cells from six females; n = 4,037 cells from 48 males). (D and E) Distribution of sequenced cells from male mice according to animal pre-treatment (sexually naive versus experienced; + Behavior, Resident-Intruder test; or retrograde labeling) prior to sample collection (D). Table (E) summarizes Cre lines used for each pre-treatment. (F) Stacked bar plot showing proportion of cells sequenced from each Cre line in 46 SMART-seq clusters. (G) Images showing distribution of cells in anterior versus posterior VMHvl in two Cre lines, *Nr5a1*, and *Oxtr*, crossed with *Ai110*. (H) Quantifications of data in (G), relative to Nissl. (I–K) Sequencing depth (I), mean mapping percentages in each category of sequence (J), and number of genes detected per cell (K) indicated for each SMARTseq cluster. Data (I and K) are represented as median with interquartile range. (L) Heatmap showing relative expression of marker genes exhibiting differential expression across clusters.

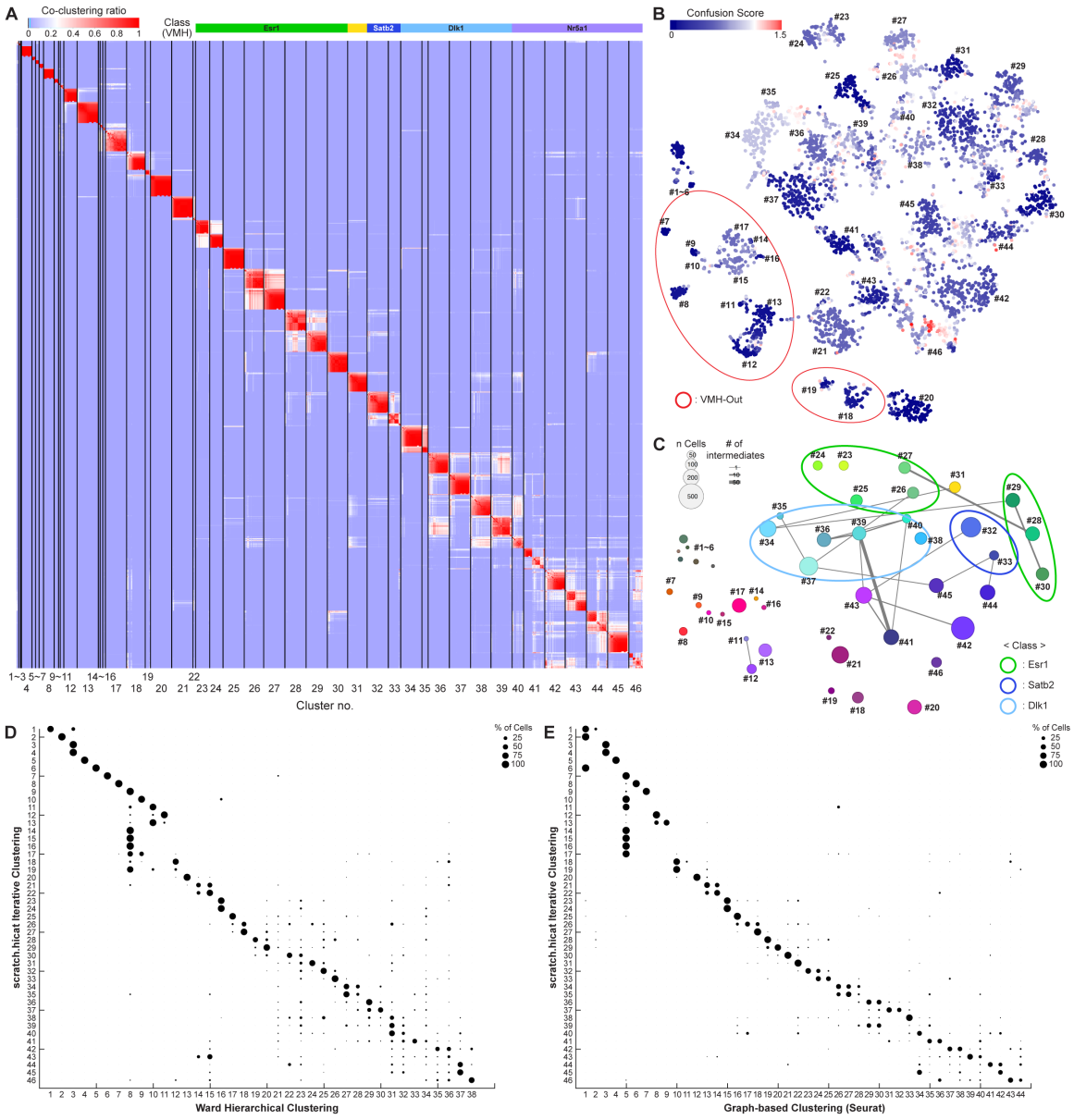

Figure S2. Robustness of SMART-seq Clusters, Related to Figure 1. (A–C) Co-clustering frequency matrix grouped by clusters (A), confusion scores color-coded (low-blue, high-red) onto t-SNE plot (B), and constellation diagram based on t-SNE plot (C) are shown. Confusion score is a metric showing how confidently each cell is assigned to its cluster (low score = high confidence). In (C), cells are considered as intermediates (lines joining two clusters; only displayed if they have more than 2 cells) if their confusion scores are above a threshold (0.9); number represented by line (n = 340 cells). Remaining cells are core cells in each cluster (disc size; n = 4,234 cells). (D and E) Dot plots

show comparisons of cluster membership based on scrattch.hicat iterative clustering (Figure 1C) with Ward's hierarchical clustering (D) or using Seurat (graph-based clustering; E).

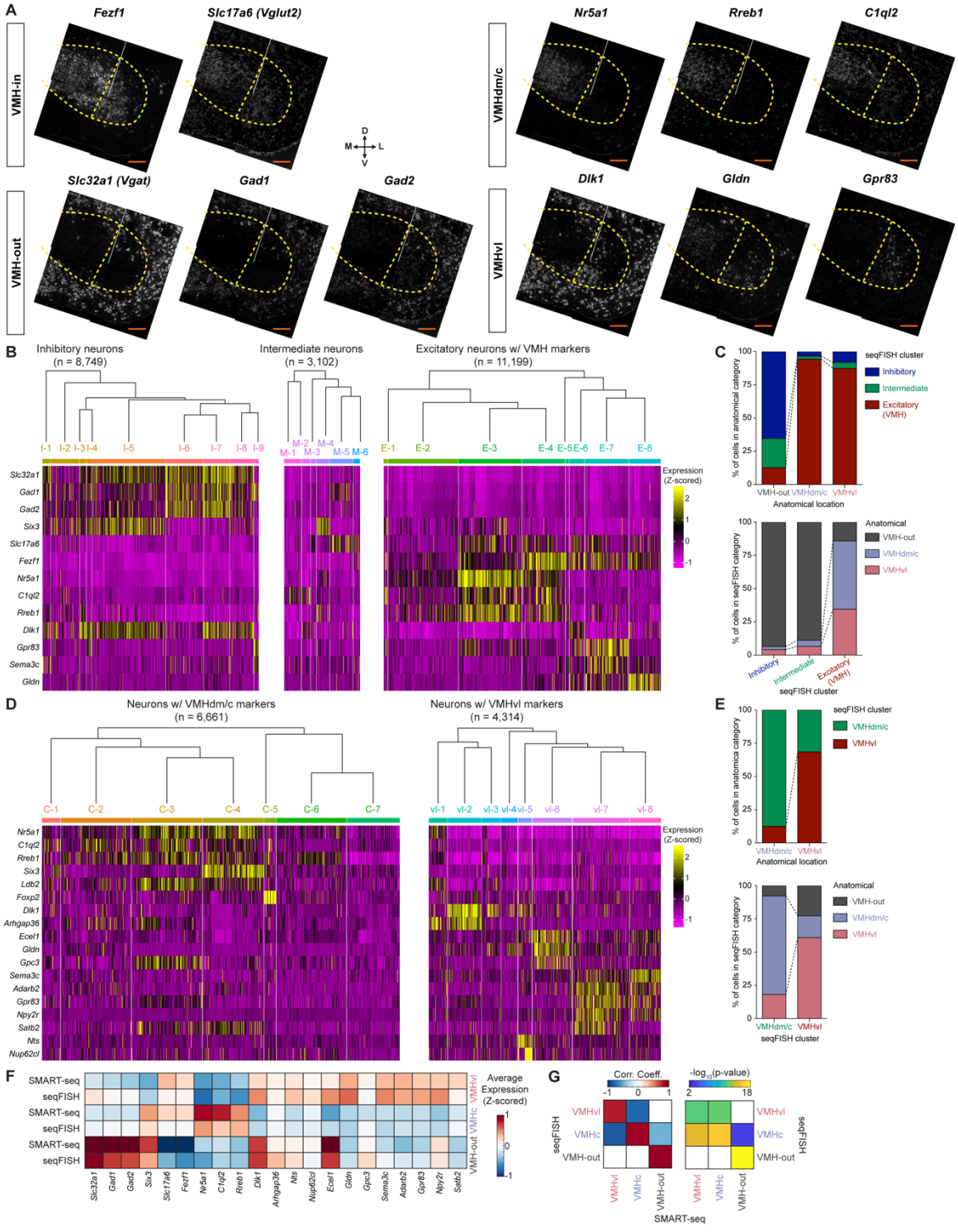

Figure S3. Iterative Clustering of Whole seqFISH Data with Region-specific Marker Genes, Related to Figure 2. (A) Examples of seqFISH images (maximum intensity Z-projections) for 11

region-specific marker genes. Scale bars, 100 μm (inset). (B–E) Iterative unbiased clustering of whole seqFISH data (including all segmented cells; n = 23,050). (B and D) Initial clustering was performed and clusters were assigned as inhibitory (left; nine clusters), intermediate (middle; six clusters), or excitatory with VMH markers (right; eight clusters) identities (B), based on the expression of region-specific marker genes (*Slc32a1*, *Gad1*, *Gad2*, *Six3*, *Slc17a6*, *Fezf1*, *Nr5a1*, *C1ql2*, *Rreb1*, *Dlk1*, *Gpr83*, *Sema3c*, and *Gldn*). A second round of clustering was performed on all excitatory neurons with VMH markers, and clusters were assigned as ‘‘Neurons w/ VMHdm/c markers’’ or ‘‘Neurons w/ VMHvl markers’’ (D), based on the expression of region-specific marker genes (*Nr5a1*, *C1ql2*, *Rreb1*, *Six3*, *Ldb2*, *Foxp2*, *Dlk1*, *Arhgap36*, *Ecel1*, *Gldn*, *Gpc3*, *Sema3c*, *Adarb2*, *Gpr83*, *Npy2r*, *Satb2*, *Nts*, and *Nup62cl*). (C and E) Stacked bar plots showing proportion of cells from each anatomical location (VMH-out, VMHdm/c, VMHvl) onto broad seqFISH cluster categories (inhibitory, intermediate, and VMH excitatory in C; VMHdm/c and VMHvl in E; upper), and vice versa (lower). (F) Heatmap showing average expression (Z scored) of marker genes in seqFISH or scRNA-seq (SMART-seq) data, in each region. (G) Left, overall correlation between SMART-seq and seqFISH marker expression, by region; right, significance of correlation (white, $p > 0.05$).

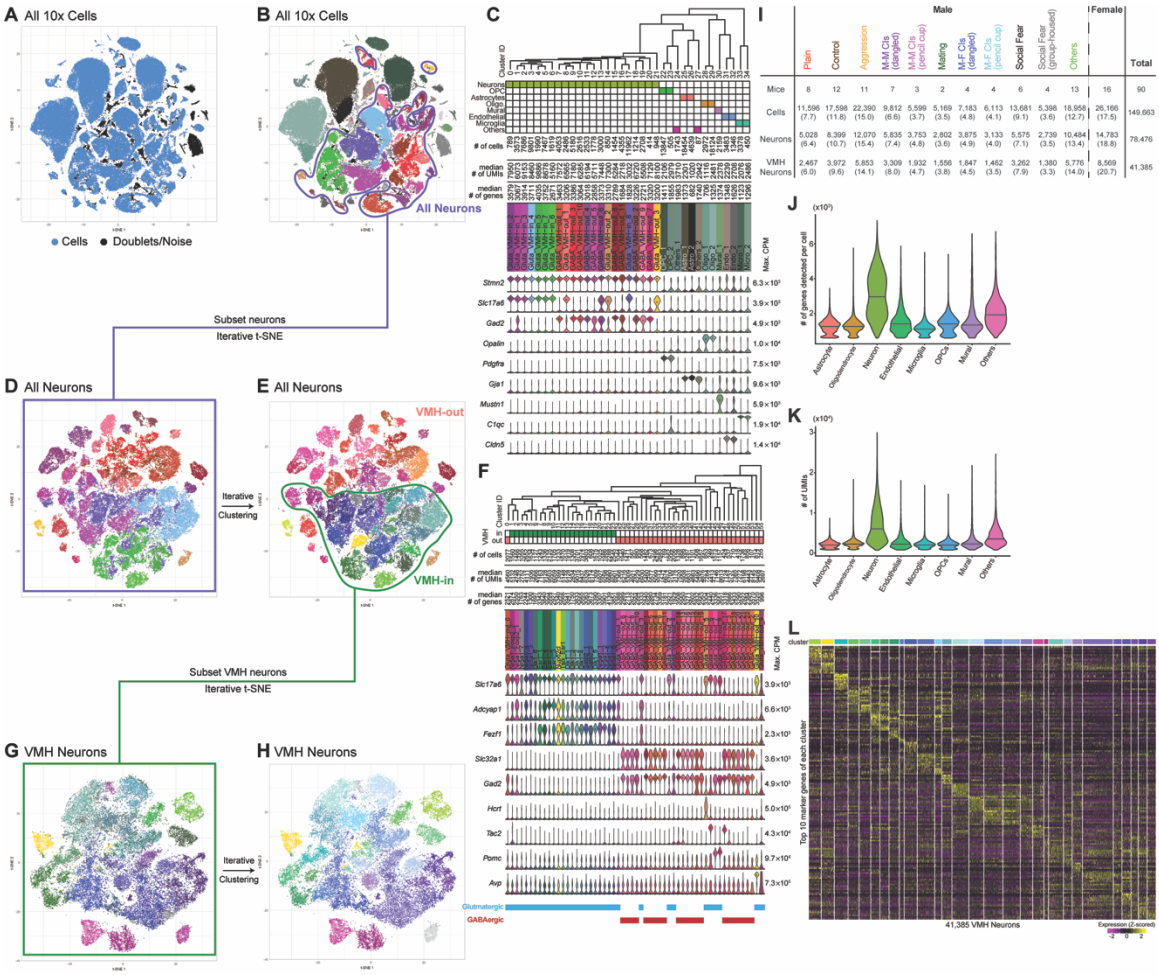

Figure S4. 10x scRNA-seq Sample Statistics, Related to Figure 4. (A–H) Iterative clustering. After filtering out doublets or noise cells (A), initial clustering was performed with remaining QC-qualified cells (B). (C) Clusters were assigned either neuronal (red, green, or blue) or non-neuronal (gray or brown) identities, based on expression of marker genes. (D and E) A second round of clustering was performed on all neuronal cells (enclosed by purple in B; color-coded according to initial clustering on t-SNE in D). (E) Clusters were assigned as VMH-in (green/blue) or VMH-out (red/brown) based on the expression of known markers (F; see also Figure 1B). Glutamatergic (blue bars) or GABAergic (red bars) clusters are indicated below violin plots in (F) based on the expression level of *Vglut2* (*Slc17a6*) and *Vgat* (*Slc32a1*), respectively. (G and H) In a third round of clustering, VMH-in neurons (enclosed by green in E; color-coded on t-SNE plot in G according to cluster identity in F) were further classified into 29 clusters (see Figure 4C). (I) Number of mice and cells sampled for 10x scRNA-seq analysis, and their percentages at each successive stage of iterative clustering (bracket

below). Mice are categorized by behavioral assays run for Act-seq just before sample collection. ‘‘Plain’’ indicates animals taken directly from home cage in housing room; ‘‘control’’ indicates animals taken after moving home cage from housing to experimental room. (J and K) Comparisons of number of genes (J) and unique molecular identifiers (UMIs) detected per cell (K), among broad cell type categories (solid black line indicating median). (L) Heatmap showing single-cell expression levels of the top 10 most differentially expressed genes in each VMH cluster.

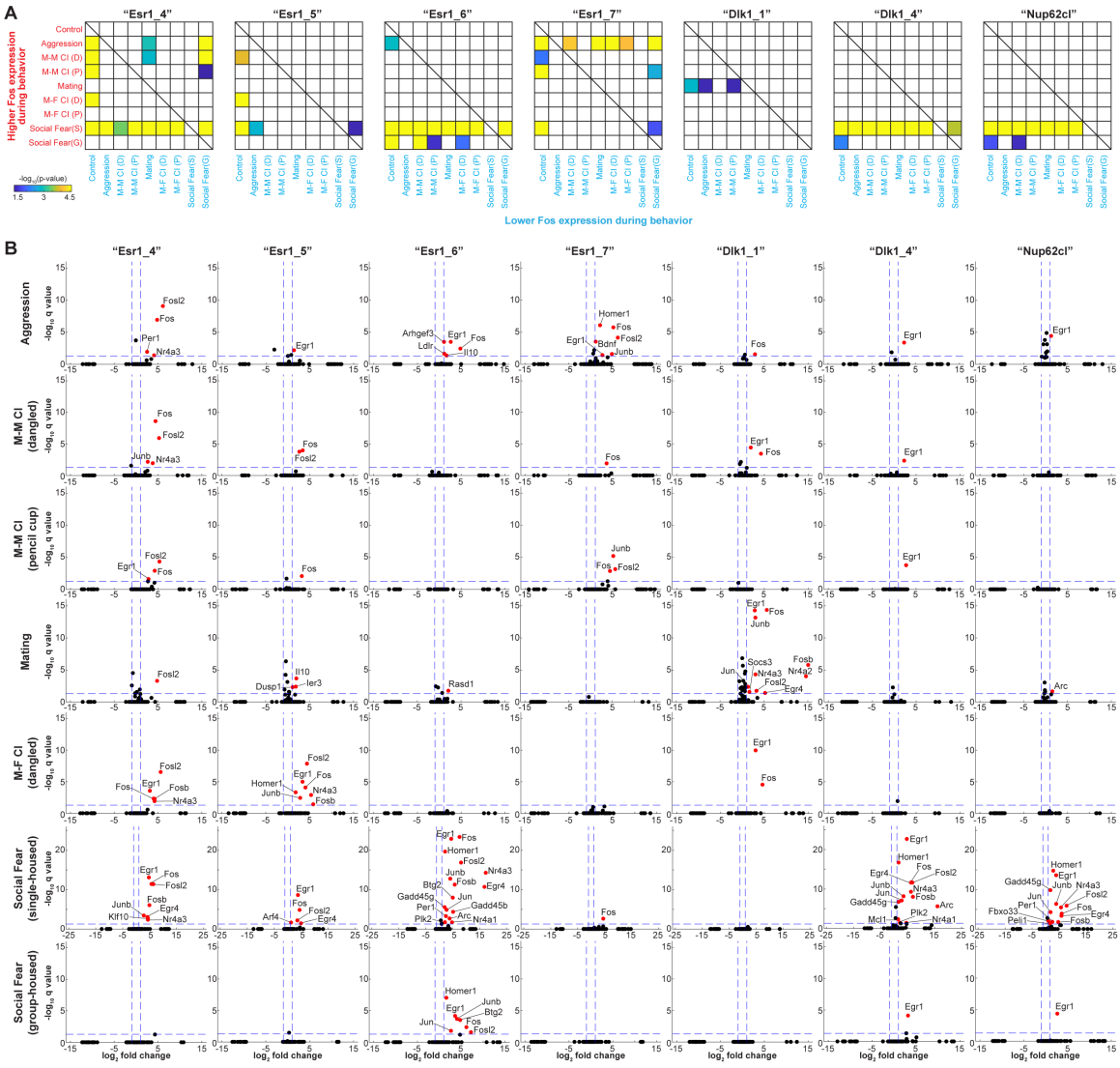

Figure S5. Expression Patterns of IEGs in Behaviorally Relevant Clusters, Related to Figure 5.

(A) Pairwise comparisons between behaviors exhibiting significant differences in *Fos* expression (colored squares), for seven major Act-seq clusters; y axis, higher levels of *Fos* (red) during behavior; x axis, lower levels (blue; control animals included). For example, cluster Dlk1_4 shows significantly higher *Fos* expression, compared to control, during two social fear assays but *Fos* levels during social fear from single-housed mice, social fear (S) are also significantly higher than during all other behaviors, including social fear from group-housed mice, Social Fear (G). (B) Scatterplots for the seven clusters in (A) showing expression level of different IEGs. Red dots indicate IEGs with fold change >2 (x axis cut-off) and FDR <0.05 (y axis cut-off; blue dashed lines), compared with control, for each indicated cluster and behavior assay (among 139 IEGs; see also Table S2).

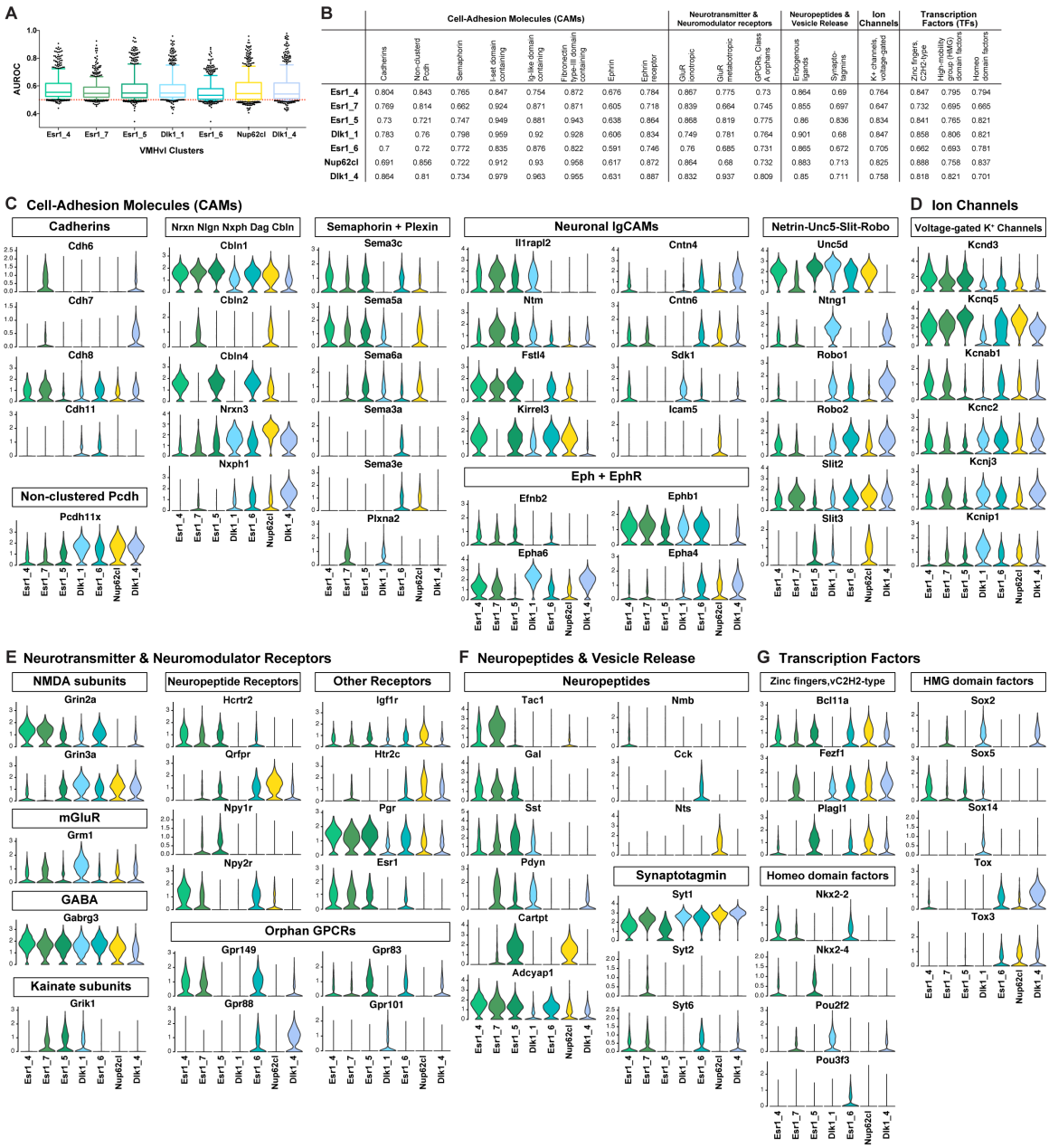

Figure S6. Comparisons of Differentially Expressed Genes between Behaviorally Relevant Clusters, Related to Figure 5 and Table S3. (A) Distribution of AUROC (area under receiver-operator characteristics curve) values for each of ~500 HGNC (HUGO Genome Nomenclature Committee) gene families, calculated using MetaNeighbor, for each major Act-seq cluster (see Figure 5). AUROC score reflects the probability of correctly identifying a given cluster (originally identified using entire gene set), using only members of that family of genes to perform clustering. Datapoints below red horizontal dashed line represent gene families who members cannot identify clusters better

than chance (AUROC = 0.5). Box-whisker plots indicate median, interquartile range, and 5^{th}–95^{th} percentiles of the distribution, respectively. Only outlier points are indicated. (B) Individual AUROC scores for specific gene families of interest. Note the high scores for the I-set domain-containing subset of Ig superfamily genes. (C–G) Violin plots show the relative expression levels of differentially expressed genes in gene families having high AUROC scores among the major Act-seq clusters: (C) cell-adhesion molecules, (D) ion channels, (E) neurotransmitter and neuromodulator receptors, (F) neuropeptides and vesicle release proteins, (G) transcription factors; see also Table S3).

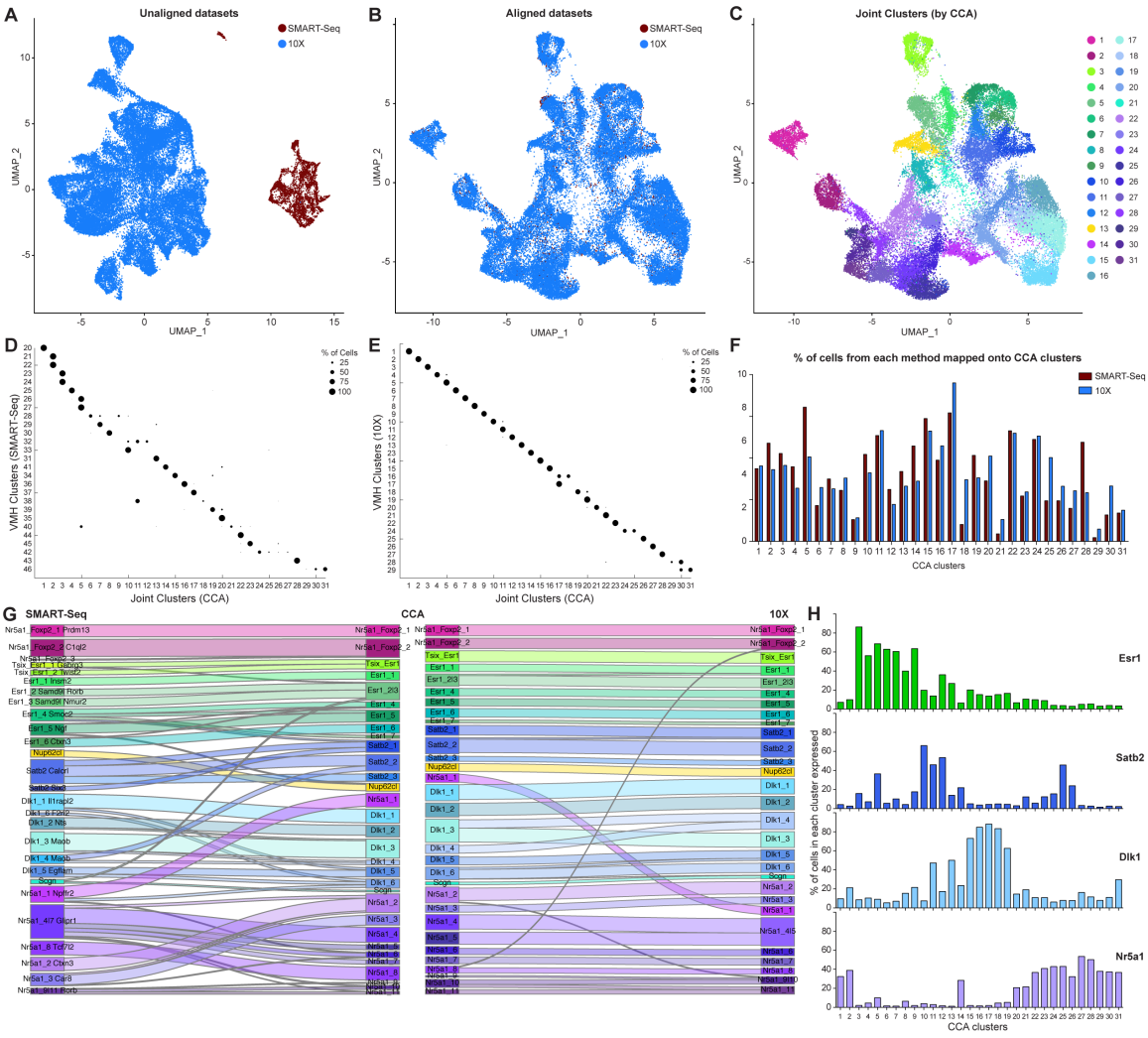

Figure S7. Comparison of SMART-seq versus 10x Clusters Using Canonical Correlation Analysis (CCA), Related to Figure 6. (A and B) t-SNE plots of all VMH cells (n = 45,005) combining SMART-seq (dark red; n = 3,824) and 10x (dark blue; n = 41,181) before (A) and after (B) CCA alignment. (C) Joint CCA clusters are labeled on t-SNE plots following alignment. (D and E) Dot plots comparing cluster membership in joint CCA clusters (x axis) with either SMART-seq (D) or 10x (E) clusters (y axis). (F) Bar graph showing distribution of cells from SMART-seq (non-Retro-seq cells only; n = 2,756) and 10x datasets across joint CCA clusters. (G) ‘‘River plots’’ illustrating how SMART-seq (left) and 10x (right) clusters are mapped onto joint CCA clusters (middle). (H) Bar graphs showing relative proportions of cells in each CCA cluster expressing marker genes for the four major VMH classes identified here.

2.7 Materials and methods

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All procedures were performed in accordance with NIH guidelines and approved by the Institutional Animal Care and Use Committee (IACUC) at the California Institute of Technology (Caltech) and Allen Institute for Brain Science (AIBS). Animals were housed and maintained on a reverse 12 h light-dark cycle with food and water ad libitum. For SMART-seq, the following transgenic and reporter lines were used: *Nr5a1*$^{Cre/+}$ (Dhillon et al., 2006), *Oxtr*$^{Cre/+}$ (Daigle et al., 2018), *Vglut2*$^{Cre/+}$ (Jackson Laboratory, stock no. 016963), *Esr1*$^{Cre/+}$ (Lee et al., 2014), *Ai14*$^{Tg/+}$ (Madisen et al., 2010), and *Ai110*$^{Tg/+}$ (Daigle et al., 2018), and maintained on the C57BL/6 background. For 10x Act-seq and seqFISH, wild-type (WT) C57BL/6N male mice (experimental), C57BL/6N female mice (for sexual experience), and BALB/c male mice (intruders) were obtained from Charles River (Burlington, MA). We used 54 mice to collect 4,574 cells for SMART-seq, 90 mice to collect 149,663 cells for 10x (Figures S1C–S1E and S4I), and 7 mice for seqFISH (3 mice with 58 probes and 4 mice with 24 probes, respectively). Animals were euthanized at P50–P110 in this study. Behavior was tested during the dark cycle.

METHOD DETAILS

Single-cell isolation, cDNA amplification and library construction for SMART-seq

We isolated fluorescently labeled single cells from the mouse brain as previously described (Tasic et al., 2016; Tasic et al., 2018). Briefly, adult (8–12-week-old) male mice were anaesthetized with isoflurane and perfused with cold artificial cerebrospinal fluid (ACSF) containing CaCl2 (0.5 mM), glucose (25 mM), HCl (96 mM), HEPES (20 mM), MgSO4 (10 mM), NaH2PO4 (1.25 mM), myoinositol (3 mM), N-acetylcysteine (12 mM), N-methyl-d-glucamine (NMDG; 96 mM), KCl (2.5 mM), NaHCO3 (25 mM), sodium l-ascorbate (5 mM), sodium pyruvate (3 mM), taurine (0.01 mM), thiourea (2 mM), bubbled with carbogen

gas (95% O_2 and 5% CO_2). The brain was sectioned at 350 μm using vibratome (VT1000S, Leica Microsystems) on ice, and the regions of interest (Figure S1B) were microdissected under a fluorescence dissecting microscope from two consecutive sections (-1.22 to -1.94 from Bregma (Franklin and Paxinos, 2008)). For enzymatic digestion, the dissected tissues were transferred to a microcentrifuge tube and treated with 1 mg/mL pronase (Sigma, P6911-1G) in carbogen-bubbled ACSF for ~70 min at room temperature without mixing in a closed tube. After incubation, the pronase solution was exchanged with cold ACSF containing 1% fetal bovine serum (FBS), The tissue pieces were dissociated into single cells by gentle trituration through Pasteur pipettes with polished tip openings of 600-μm, 300-μm and 150-μm diameter. Single cells were sorted by FACS into 8-well PCR strips containing lysis buffer from SMART-Seq v4 kit with RNase inhibitor (0.17 U/ml), immediately frozen on dry ice, and stored at -80˚C. We used the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara, 634894) to reverse transcribe poly(A) RNA and amplify full-length cDNA according to the manufacturer's instructions. We performed reverse transcription and cDNA amplification for 18 PCR cycles in 8-well strips, in sets of 12–24 strips at a time. At least 1 control strip was used per amplification set, which contained 4 wells without cells and 4 wells with 10 pg control RNA. Control RNA was either Mouse Whole Brain Total RNA (Zyagen, MR-201) or control RNA provided in the SMART-Seq v4 kit. All samples proceeded through Nextera XT DNA Library Preparation (Illumina, FC-131-1096) using Nextera XT Index Kit V2 Set A (Illumina, FC-131-2001). Nextera XT DNA Library prep was performed according to manufacturer's instructions except that the volumes of all reagents including cDNA input were decreased to 0.4 × or 0.5 × by volume.

SMART-seq sequencing data processing and quality control

Libraries were sequenced on an Illumina HiSeq2500 platform (paired-end with read lengths of 50 nt) and Illumina sequencing reads were aligned to GRCm38 (mm10) using a RefSeq annotation gff file retrieved from NCBI on 18 January 2016 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/all/). Sequence alignment was

performed using STAR v2.5.3 (Dobin et al., 2013) in twopass Mode. PCR duplicates were masked and removed using STAR option 'bamRemoveDuplicates'. Only uniquely aligned reads were used for gene quantification. Gene counts were computed using the R GenomicAlignments package (Lawrence et al., 2013) summarizeOverlaps function using 'IntersectionNotEmpty' mode for exonic and intronic regions separately. For the SMART-seq dataset, we only used exonic regions for gene quantification. Cells that met any one of the following criteria were removed: < 100,000 total reads, < 1,000 detected genes (CPM > 0), < 75% of reads aligned to genome, or CG dinucleotide odds ratio > 0.5.

Retro-seq sample preparation

Mice (8–12-week-old) were anaesthetized (with 2%–5% isoflurane for induction and 0.8%–2% for maintenance), and mounted on a stereotaxic frame (David Kopf Instruments) with heating pad placed underneath. We bilaterally injected a volume of 200 nL containing retrograde tracer viruses into dPAG, lPAG, or MPOA of using a pulled glass capillary (World Precision Instruments) by pressure injection at a flow rate of 100 nL/min (Micro4 controller, World Precision Instruments; Nanojector II, Drummond Scientific). The combination of retrogradely transported virus and mice used were as follows: 1) rAAV2-retro-EF1a-Cre (Tervo et al., 2016) in a Cre-reporter mouse line (*Ai14*$^{Tg/+}$ or *Ai110*$^{Tg/+}$) and 2) HSV1-LS1L-EYFP (rHSV; MGH Vector Core) in *Vglut2*$^{Cre/+}$ or *Esr1*$^{Cre/+}$ (Figure S1E). Stereotactic injection coordinates of dPAG, lPAG, and MPOA were obtained from the Paxinos and Franklin atlas (Franklin and Paxinos, 2008) (AP: _4.72, ML: ± 0.12, DV: -2.0 mm for dPAG; AP: -4.48, ML: ± 0.45, DV: -2.65 mm for lPAG; AP: 0.022, ML: ± 0.35, DV: -5.37 mm for MPOA). After a 2–3 week viral incubation, retrogradely labeled single cells were manually dissected from VMHvl, isolated by FACS, and subjected to scRNA-seq using SMART-seq as described above.

Single-cell isolation, cDNA amplification and library construction for 10x scRNA-seq and Act-seq

We isolated single cells from the mouse brain as previously described (Wu et al., 2017) with some modifications. Briefly, 40-55 min after the resident-intruder assay (see 'Resident intruder assay' section below), adult (8–12-week-old) resident mice (2–4 brains collected at a time) were anaesthetized with isoflurane and transcardially perfused with cold NMDG-ACSF (adjusted to pH 7.3–7.4) containing CaCl2 (0.5 mM), glucose (25 mM), HCl (92 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), MgSO4 (10 mM), NaHCO3 (30 mM), NaH2PO4 (1.2 mM), NMDG (92 mM), sodium L-ascorbate (5 mM), sodium pyruvate (3 mM), thiourea (2 mM), bubbled with carbogen gas. The brain was sectioned at 350 μm using vibratome (VT1000S, Leica Microsystems) on ice, and the regions of interest were microdissected under a fluorescence dissecting microscope from two consecutive sections (-1.22 to -1.94 from Bregma (Franklin and Paxinos, 2008)). The microdissected tissues were accumulated in a microcentrifuge tube containing NMDG-ACSF with 30 mM actinomycin D on ice to suppress further immediate early gene (IEG) activation (Wu et al., 2017). Thereafter the sections were transferred to a new microcentrifuge tube for papain digestion (60 U/mL, Sigma, P3125-250MG; pre-activated at 34˚C for 30 min) in Trehalose-HEPES-ACSF (adjusted to pH 7.3–7.4) containing actinomycin D (15mM), CaCl2 (2 mM), L-cysteine (2.5 mM), EDTA (0.5mM), glucose (25 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), MgSO4 (2 mM), NaCl (92 mM), NaHCO3 (30 mM), NaH2PO4 (1.2 mM), trehalose (2.5% w/v) and gently carbogenated. During incubation for enzymatic digestion (~55 min at room temperature), the solution was mixed by gently pipetting a few times every 5-10 min. After incubation, the solution was exchanged with cold Trehalose-HEPES-ACSF containing Egg White/BSA ovomucoid inhibitor (3 mg/ mL, Worthington, OI-BSA) and DNase I (25 U/mL, Thermo scientific, 90083). The tissue pieces were dissociated into single cells by gentle, successive trituration through Pasteur pipettes with polished tip openings of 600-μm, 300-μm, and 150-μm diameters. After trituration and filtering through a 40 μm cell strainer, single cells were pelleted at 300 g for 5 min at 4˚C, the supernatant was carefully removed, the cells resuspended with cold Trehalose-HEPES-

ACSF, and filtered through a 20 μm filter. Cells were pelleted again at 300 g for 5 min at 4˚C and resuspended with Resuspension-HEPES-ACSF containing BSA (0.05%), CaCl2 (2 mM), glucose (25 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), MgSO4 (1 mM), NaCl (117 mM), NaHCO3 (30 mM), NaH2PO4 (1.2 mM) (osmolarity verified to be within 10 mOsm of Trehalose-HEPES-ACSF). After manually determining the cell concentration using a hemocytometer, suspensions were further diluted to desired concentrations (300–1,000 cells/μL) if necessary. The appropriate volume of reverse transcription (RT) mix was added in order to target 6,000–10,000 cells recovered, and loaded into the chip. The Chromium Single Cell 30 Library & Gel Bead Kit v2 (PN- 120237), Chromium Single Cell 30 Chip kit v2 (PN-120236), and Chromium i7 Multiplex Kit (PN-120262) were used for all downstream RT, cDNA amplification (11 PCR cycles), and library preparation as instructed by the manufacturer (Chromium Single Cell 30 Reagents Kits v2 User Guide).

10x Sequencing data processing and quality control

Libraries were sequenced on an Illumina HiSeq4000 or NovaSeq6000 (paired-end with read lengths of 150 nt) and Illumina sequencing reads were aligned to the mouse pre-mRNA reference transcriptome (mm10) using the 10x Genomics CellRanger pipeline (version 3.0.2) with the default parameters. The median read depth per cell was 66,319 reads, yielding 2,998/1,309 genes (median), and 6,096/2,238 unique molecular identifier (UMI) counts (median) per neuronal/non-neuronal cell, respectively (Figure S4C, J, K). Cells that met any one of the following criteria were filtered out for downstream processing in each 10x run: < 600 detected genes (for UMI count > 0), > 30,000 UMI counts (potential multiplets), or the proportion of the UMI count attributable to mitochondrial genes was greater than 15%. Doublets were further removed by first classifying cells into broad cell classes (neuronal versus non-neuronal) based on the co-expression of any pair of their marker genes (*Stmn2* for neurons; *Cldn5* for endothelial cells; *C1qc* for microglia; *Opalin* for oligodendrocytes; *Gja1* for astrocytes; *Pdgfra* for OPCs; *Mustn1* for mural cells; see Figure S4A).

seqFISH data generation

seqFISH data in VMH and its surrounding area were generated using the seqFISH+ protocol with some modifications (Eng et al., 2019). In brief, primary probes of 30-nt sequences of each gene were extracted using the exons from within the CDS region and BLASTed against the mouse transcriptome to ensure specificity. A minimum of 15 to a maximum of 48 primary probes were designed for each targeted gene. The number of probes per gene was dependent on the length of the CDS region. The same 15-nt readout probes were used as in previous study (Shah et al., 2018). The reverse complements of these readout probe sequences were included as four binding sites in the primary probes to increase signals. The final primary probes were ordered as a complex oligopool from Twist Bioscience and were constructed as previously described (Eng et al., 2019; Shah et al., 2018).

Adult (8–12-week-old) C57BL/6N male and female mice were perfused with perfusion buffer (10U/mL heparin, 0.5% NaNO2 (w/v) in 1X PBS at 4˚C). Mice were then perfused with fresh4%paraformaldehyde (PFA; Thermo Scientific, 28908)/1X PBS buffer (Invitrogen, AM9624). The brain was dissected and immediately placed in a 4% PFA buffer for 2 h at room temperature. The brain was then immersed in 30% RNase-free Sucrose (Amresco, 0335) in 1X PBS for 48 h at 4˚C until the brain sank to the bottom of the tube. The brain was then embedded in OCT and cryo-sectioned (15 μm thick), and 4–5 sections including VMHvl were collected per mouse (consecutive sections were ~100 μm apart along A-P axis; 20 sections in total). The sections were stored at -80˚C until use.

For seqFISH data acquisition, sections were prepared as described with modifications (Eng et al., 2019). In brief, sections were permeabilized in 70% ethanol for 18 h at 4˚C, and further permeabilized by the addition of RNase-free 8% SDS (Ambion, AM9822) in 1X PBS for 15 min at room temperature. Tissue slices were then rinsed multiple times with 70% ethanol to remove SDS. Primary probe hybridization buffer composed of 40% Formamide (Invitrogen, AM9344), 2X SSC (Invitrogen, 15557-044) 10% (w/v) Dextran Sulfate (Sigma, D8906), and ~5nM/primary probes were hybridized to the tissue slices by spreading out the hybridization buffer solution with a coverslip. The hybridization was allowed to proceed for

~36 h at 37˚C incubator in a humid chamber. After primary probe hybridization, the tissue slices were washed with 40% Wash Buffer (40% WB: 2x SSC, 40% Formamide (v/v), 0.1% Triton X-100 (Sigma, D8906) at 37˚C for 30 min. After rinsing with 2X SSC 3 times and 1X PBS once, the sample was subjected to 0.1mg/mL Acryoloyl-X SE (Thermo Fisher, A20770) in 1X PBS treatment for 30 min at room temperature. Next, the tissue slices were incubated with 4% acrylamide (1:19 crosslinking) hydrogel solution in 2X SSC for 30 min at room temperature. Then the hydrogel solution was aspirated and 30 μL of 4% hydrogel solution containing 0.05% APS and 0.05% TEMED in 2x SSC was dropped onto the tissue slice and sandwiched by Gel-Slick functionalized coverslip. The samples were transferred to 4˚C in a homemade nitrogen gas chamber for 30 min before transferring to 37˚C for 2.5 h to complete polymerization. After polymerization, the hydrogel embedded tissue slices were cleared with digestion buffer as previously described (Eng et al., 2019). After digestion, the tissue slices were rinsed with 2X SSC multiple times and subjected to 0.1mg/mL Label-X modification for 45 min at 37˚C. The tissue slices were reembedded in hydrogel solution as in the previous step, and were gelated at 37˚C incubator for 2 h for stronger stabilization before subsequent multiple rounds of imaging. Imaging platform and automated fluidics delivery system were similar to those previously described (Eng et al., 2019).

Each round of imaging contained the 405 channel, which included the DAPI stain of the cell. The DAPI images from all the rounds of hybridization were aligned to the first image using a 3D phase correlation algorithm. Tissue background and auto-fluorescence were then removed by dividing the initial background with the fluorescence images. To correct for the non-uniform background, a flat field correction was applied by dividing the normalized background illumination with each of the fluorescence images while preserving of the intensity profile of the fluorescent points. The background signal was then subtracted using the ImageJ rolling ball background subtraction algorithm with a radius of 3 pixels. For semi-automatic cell segmentation, images of sections stained with Nissl and DAPI were trained with Ilastik (Sommer et al., 2011), an interactive supervised machine learning toolkit, to output probability maps, which were then used in the Multicut (Beier et al., 2017) tool to produce volumetric labeled cells. The labeled cells were filtered by size and region in

MATLAB. For image analysis, potential mRNA transcript signals were located by finding the local maxima in the processed image above a predetermined pixel threshold. The transcript spots were assigned to the corresponding labeled cells according to location. For more details regarding the seqFISH method, see (Shah et al., 2016).

Resident intruder assay

The resident intruder assay was performed as previously described (Hong et al., 2014; Lee et al., 2014). Briefly, all experimental male mice (‘‘residents’’) were had been individually housed for two weeks and received sexually experience (for 1–3 days), except residents for ‘‘Social Fear’’ assays using group-housed (sexually naive) males. Before testing, the resident mice were pre-screened for baseline aggression and sexual behavior using resident-intruder testing sessions (~3 times). On the experimental day, residents were transported in their home cage to a novel behavioral testing room (under infrared light) where they acclimated for 5–15 min. An unfamiliar group-housed BALB/c mouse (‘‘intruders’’; types of intruders listed in Figure 5A) was then placed in the home cage of the resident, and residents were allowed to freely interact with it for ~10 min. For ‘‘M-M CI’’ (male-male close investigation only) and ‘‘M-F CI’’ (male-female close investigation only) tests, intruders were either dangled by the tail or placed inside a wire-mesh barrier (an inverted pencil cup) when introduced into the resident’s home cage, so that resident behavior was restricted to investigation and not allowed to proceed to the consummatory phases (fighting or copulation). ‘‘Control’’ indicates home-caged animals moved to the behavior testing room, but not otherwise manipulated. ‘‘Plain’’ indicated animals taken directly from their home cage in the housing room. In the cases of ‘‘Aggression’’ and ‘‘Mating,’’ only resident mice that were actively engaging in either aggression or sexual behaviors throughout the session were used for subsequent 10x scRNA-seq experiments. For ‘‘Social Fear,’’ experimental mice were introduced to the home cage of singly housed male aggressors for ~10 min to allow them to be attacked (more than 10 times), and then returned into their home cage. For female mice, sexually naive and group-housed mice were used as residents for the ‘‘Plain’’ category (n =

5,935 cells from 4 mice), ‘‘Control’’ (n = 7,964 cells from 4 mice), or social interaction tests with male intruders (n = 12,267 cells from 8 mice; most of them were unreceptive), respectively.

Retrograde labeling with *c-fos* immunohistochemistry

For retrograde labeling, 200 nL of HSV-LS1L-mCherry or HSV-LS1L-EYFP (MGH Vector Core) was injected into either dPAG or MPOA of adult *Esr1*$^{cre/+}$ mice (8–12 weeks old). ~3 weeks after injections, mice were tested in resident-intruder assays (3 main categories: male-male aggression, male-female mounting, and male-female close investigation only). 80–90 min after the resident-intruder assay, mice were euthanized via transcardiac perfusion with 4% PFA, and serial sections generated using a cryostat. To visualize *c-fos*, sections were stained with goat anti-*c-fos* antibody (Santa Cruz Biotechnology, sc-52-g, 1:500) overnight at 4˚C, followed by Alexa-647 donkey anti-goat (Invitrogen, A-21447, 1:1000) at room temperature for 3 h. For manual quantification of fluorescent cells, images were analyzed using Fiji/ImageJ as follows: 1) a manually set threshold was applied (to the point that background autofluorescence was not visible), 2) watershed-based segmentation and particle analysis (‘‘analyze particles’’ function) was used for the identification of individual cells (with size and circularity limited to an appropriate range), and 3) automatic particle counting was visually inspected for accuracy and manually adjusted.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were processed and analyzed using MATLAB, R, Prism 6, ImageJ, and Excel. The following statistical analyses were conducted: ANOVAs followed by Bonferroni post hoc tests to compare 1) the expression of *Fos* induced by each behavioral paradigm with control across clusters (Figures 5B and 7A), and 2) the proportion of double^{+} (*Fos* & rHSV) cells between aggression and control injected in dPAG and MPOA (Figure 7D); Fisher’s exact

tests to see 1) which retrogradely labeled SMART-seq classes/clusters from a given projection were significantly different from the proportions of retrogradely labeled cells from other projection (Figures 3B and 3C), 2) what kinds of joint CCA clusters were significantly different from the expected proportions of cells in a given sex (Figure 6B), and 3) how differently cells from sex-specific clusters were located between anterior versus posterior VMHvl (Figure 6E); a likelihood ratio test (McDavid et al., 2013) was used to calculate adjusted p-values (FindMarkers function in Seurat; after Bonferroni correction using all genes in the dataset) for determining the significance values of differential IEG expression compared to controls (Figure S5B). Unpaired t tests were used to compute p-values for 1) the differential expressions of genes in seqFISH (Figure 6F), and 2) different proportion of double^{+} (*Fos* & rHSV) cells between M-F social interaction tests (mating + M-F CI tests) and control both in dPAG and MPOA injections (Figure 7D). The n value and statistically significant effects are reported in each figure/figure legend. The significance threshold was held at $\alpha = 0.05$ ($p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

SMART-seq data analysis and clustering

Data analysis for the SMART-seq dataset (including clustering, dendrogram construction and differential gene expression) was performed using a pre-release version of R package scrattch.hicat (https://github.com/AllenInstitute/scrattch.hicat), as described previously (Tasic et al., 2018). In brief, all QC qualified cells were clustered using high variance gene selection, dimensionality reduction (using default PCA mode), dimension filtering, and Jaccard-Louvain or hierarchical (Ward) clustering. This process was repeated within each resulting cluster until no more child clusters met differential gene expression or cluster size termination criteria. The entire clustering procedure was repeated 100 times using 80% of all cells sampled at random, and the frequency with which cells co-cluster was used to generate a final set of clusters (Figure S2A), again subject to differential gene expression and cluster size termination criteria. Pairwise differential gene expression was performed using the de_score function in the scrattch.hicat package based on $\log_2(CPM+1)$ values. The cell type

tree was built using the build_dend function in scrattch.hicat based on cluster median expression profiles of marker genes. Neighboring clusters in the cell type taxonomy with shared key markers are defined as types within the subclass, and clusters are named for key markers with a suffix for further discrimination.

To evaluate the clustering robustness, we compared the original SMART-seq cluster membership (Figure 1C; based on scratch.hicat iterative clustering) with different data processing and clustering algorithms (either Ward's hierarchical clustering (Figure S2D) or graph-based clustering by Seurat (Figure S2E)). For Ward's hierarchical clustering (Ward, 1963), gene counts across 45,768 exons in single cells were log-normalized. Because dimensionality reduction on this dataset was spatially structured by cell sequencing depth, we normalized for sequencing depth per cell by regressing out the log of the sum of the exon counts per cell. After normalization, PCA followed by t-SNE no longer arranged cells by sequencing depth. scipy.cluster.hierarchy.linkage was then used to perform hierarchical clustering using the Ward variance minimization metric on the normalized counts, and the hierarchical clustering was cut at 38 clusters. For graph-based clustering, SMART-seq gene counts were loaded into Seurat. After log-normalization and data scaling (using default parameters), the top 2,000 highly variable genes were used (FindVariableGenes function; selection.method = 'vst') to identify 44 clusters (FindClusters function; using 40 PCs given the highly variable genes and a resolution of 4.5).

All retrogradely labeled SMART-seq cells were subjected to the same experimental and data processing and clustering with all other quality control-passed single-cells. Clustering was performed blinded to the experimental source of retrogradely labeled cells.

10x Data analysis and clustering

All downstream analysis of Act-seq 10x scRNA-seq data was performed with R package Seurat (v3.0.3) (Butler et al., 2018; Satija et al., 2015; Stuart et al., 2019), using default

parameters unless specified (https://satijalab.org/seurat/). Briefly, after initial cell filtering, genes expressed in fewer than 0.1% of total cells in each experiment were also removed, and all 10x gene expression matrices from each experiment were loaded and merged into Seurat. Gene expressions of each cell were normalized by total number of molecules, multiplied by a scale factor (10,000) per cell, and log-transformed (NormalizeData function). Then the expressions of each gene were scaled (its mean/variance across cells is 0 and 1, respectively) and the proportion of mitochondrial UMIs was regressed out (ScaleData function). Next, highly variable genes were identified (FindVariableGenes function; top 2,000 genes with the highest standardized variance selected by selection.method = 'vst') and used as input for dimensionality reduction via principal component analysis (PCA) after removing sex-specific genes (*Ddx3y*, *Eif2s3y*, *Uty*, *Kdm5d*, *Xist*, *Tsix*), immediate early genes (e.g., *Fos*, *Fosl2*, *Junb*, *Egr1*, *Arc*, *Homer1*; 139 genes in total from (Wu et al., 2017)), 30 retro-virus-induced genes (e.g., *B2m*, *Bst2*, *Oasl2*, *Ifit1*), and 1,000 noise-sensitive genes (high abundance genes sensitive to technical noise; see also Table S2). The resulting PCs were then used as input for clustering analysis (FindClusters function). Specifically, we performed iterative rounds of clustering and cell selection (Figures S4A–S4H). For the initial 149,663 cells, we used 40 PCs to classify 35 unique clusters (a resolution of 0.6) to separate neuronal cells from non-neuronal cells. We then selected the putative neuronal clusters (#0–21) for the 2nd round of clustering analysis. 56 neuronal clusters were identified using 40 PCs (a resolution of 2.0) from 78,476 neurons to separate glutamatergic VMH-in (inside VMH) neurons from GABAergic VMH-out (outside VMH) neurons. The final round of clustering analysis was performed given 41,385 VMH neurons (#1–23), and 29 VMH clusters were identified using 40 PCs (a resolution of 1.6; Figures 4C–4E). Clustering robustness for VMH clusters was evaluated by varying the parameters (number of PCs: 30–60; resolution: 1–2), and cluster identification was robust across the range of PCs and resolutions. For VMH clusters, top 10 differentially expressed genes were computed by FindAllMarkers function (min.pct = 0.25, logfc.threshold = 0.25; Figure S4L).

To compare the expression level of *Fos* in each resident-intruder assay with control in whole VMHvl clusters (Figure 5B; 14,763 cells from 10x VMH cluster #2–17; only cells from male

mice were included), we calculated p-values after multiple comparison correction across behaviors and clusters (using 3-way ANOVA and Bonferroni post hoc test). We also performed the same analysis using only cells from M-M social interaction tests (‘‘M-M’’; aggression and two M-M CI tests) or M-F social interaction tests (‘‘M-F’’; mating and two M-F CI tests) samples in the 6 Esr1 clusters (#3-8; Figure 5C; 3,972, 11,094, and 4,865 cells from control, M-M, and M-F, respectively). For 7 selected VMHvl clusters (Esr1_4, Esr1_5, Esr1_6, Esr1_7, Nup62cl, Dlk1_1, and Dlk1_4; showing strong *Fos* inductions in Figure 5B), 1) pairwise comparisons between each social behavior exhibiting significant differences in Fos expression (Figure S5A; after multiple comparison correction), and 2) differential IEG expression analysis (for 139 IEGs; see Table S2) using Seurat (a likelihood ratio test (McDavid et al., 2013) by FindMarkers function) to identify significantly upregulated (average fold change > 2 and FDR < 0.05) IEGs in each behavior compared to control (Figure S5B) were performed. For each VMHvl cluster, the number of IEGs significantly induced (average fold change > 2 and FDR < 0.05) during each social behavior compared to control was also computed (Figure 5E).

seqFISH Data analysis and clustering

The smFISH counts for each gene were normalized by Z-scoring within each brain section, and for imaging ROIs that covered the same portion of VMHvl, their spatial coordinates were integrated into a common coordinate system. Based on Nissl-stained images, VMH and its subdivisions (VMHvl or VMHc) were identified, and cells were assigned to one of the major anatomic regions (VMHvl, VMHc and VMH-out). For 4,686 cells in VMHvl, their gene expression matrix was loaded into Seurat, PCA and clustering was performed (using 40 PCs and a resolution of 3.4 in FindClusters). Among initial 29 clusters, we filtered out 2 clusters that showed very low expression levels of all genes and are mostly located nearby the boundary of the ROIs. The dendrogram for final 27 clusters (4,497 cells) was constructed based on the average distance between clusters in PCA space (BuildClusterTree function; Figure 2B).

To establish correspondence between scRNA-seq and seqFISH clusters, we computed the pairwise Pearson's correlation coefficient between the expression patterns of 58 seqFISH genes in SMART-seq (3,824 cells) and seqFISH (4,497 cells) VMHvl dataset (using corrcoef function in MATLAB). We then identified the seqFISH clusters that were significantly correlated with SMART-seq clusters ($p < 0.05$), and whose correlation coefficients were greater than the median value of bootstrap replicates ($r > 0.35$). Among these, the seqFISH clusters most strongly correlated with each scRNA-seq cluster are marked by red squares (Figure 2C). 23 out of 27 seqFISH clusters corresponded to one or multiple SMART-seq clusters, and the pairwise correlation between these seqFISH clusters and scRNA-seq subclasses showed a clear diagonal relationship (Figure 2C).

MetaNeighbor analysis

To identify the set of genes that highly predict each of the 7 major Act-seq clusters identified in VMHvl (Esr1_4, Esr1_5, Esr1_6, Esr1_7, Nup62cl, Dlk1_1, and Dlk1_4), we calculated AUROC (Area Under Receiver-Operator Characteristics Curve) scores using MetaNeighbor as described previously (Crow et al., 2018; Paul et al., 2017). Briefly, we inputted ~500 HGNC gene families (Paul et al., 2017) and a gene expression matrix for each of the 7 clusters (7,632 cells only from male mice), with their cell type identities (and a single experiment labeling only) to MetaNeighbor function. Output AUROC scores reflect the probability of a correct assignment of cell types based on a given set of genes, when making a single binary choice, using randomized labels (AUROC ~0.5) as control (see Table S3).

Canonical correlation analysis (CCA) with SMART-seq and 10x datasets

To integrate the two different scRNA-seq datasets, their shared sources of variation were identified by Seurat CCA alignment (Butler et al., 2018; Stuart et al., 2019). Because the 10x data contained a large number of contaminating cells outside of VMH, in order to make a

fair comparison, only cells from VMH clusters were considered from each dataset (3,824 cells from SMART-seq clusters #20–46; 41,385 cells from 10x). Specifically, after data preprocessing (log-normalization, data scaling for each gene, and detecting variable genes) on SMART-seq data in Seurat, we selected the top 2,000 genes with the highest dispersion (variance to mean ratio) from each dataset. We took the union of these two resulting gene sets (2,338 in total) as input to identify the pairwise correspondences between single cells across two datasets, called anchors (FindIntegrationAnchors function). We then used these anchors to integrate the two datasets together (IntegrateData function), performed a joint clustering on these aligned embeddings (FindClusters function; resolution of 1.9), yielding 31 clusters (referred to as ‘‘Joint Clusters (CCA)’’; Figures 6A and S7C). Lastly, differentially expressed genes that were conserved across datasets were identified (FindConservedMarkers function), and a dendrogram was built using build_dend function in scrattch.hicat based on cluster median expression profiles of marker genes (Figure 6I).

DATA AND CODE AVAILABILITY

Data analysis and visualization software

The accession number for the scRNA-seq data reported in this paper is https://doi.org/10.17632/ypx3sw2f7c.1 (Mendeley Data). Analysis and visualization of transcriptomic data were performed using MATLAB (version 2017a) and R v.3.3.0 and greater, assisted by the Rstudio IDE (Integrated Development Environment for R v.1.1.442; https://rstudio.com/) as well as the following R packages: cowplot v.0.9.2 (https://rdrr.io/cran/cowplot/), dendextend v.1.5.269, dplyr v.0.7.4 (https://dplyr.tidyverse.org/), feather v0.3.1 (https://rdrr.io/cran/feather/), ggbeeswarm v.0.6.0 (https://cran.r-project.org/web/packages/ggbeeswarm/index.html), ggExtra v.0.8 (https://rdrr.io/cran/ggExtra/), ggplot2 v.2.2.170, ggrepel v.0.7.0 (https://cran.r-project.org/web/packages/ggrepel/vignettes/ggrepel.html), googlesheets v.0.2.2 (https://cran.r-project.org/web/packages/googlesheets/vignettes/basic-usage.html),

gridExtra v.2.3 (https://cran.r-project.org/web/packages/gridExtra/index.html), Hmisc v.4.1-1 (https://cran.r-project.org/web/packages/Hmisc/index.html), igraph v.1.2.1 (https://www.rdocumentation.org/packages/igraph/versions/1.2.1), limma v.3.30.1366,71, Matrix v.1.2-12 (https://rdrr.io/rforge/Matrix/), matrixStats v.0.53.1 (https://cran.rstudio.com/web/packages/matrixStats/index.html), pals v.1.5 (https://rdrr.io/cran/pals/), purrr v.0.2.4 (https://purrr.tidyverse.org/), pvclust v.2.0-0 (http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/), reshape2 v.1.4.2 (https://www.statmethods.net/management/reshape.html), Rphenograph v.0.99.1 (https://rdrr.io/github/JinmiaoChenLab/Rphenograph/), Rtsne v.0.14. (https://cran.r-project.org/web/packages/Rtsne/citation.html), Seurat v.3.0.3, viridis v.0.5.0 (https://rdrr.io/cran/viridisLite/man/viridis.html), and xlsx v.0.5.7 (https://cran.r-project.org/web/packages/xlsx/index.html).

Code availability

An R package for iterative clustering (scrattch.hicat) is available on GitHub at https://github.com/AllenInstitute/scrattch.hicat. All MATLAB and R scripts used in this manuscript are available on reasonable request.

2.8 Additional information

ACKNOWLEDGMENTS

We thank A. Jones and C. Koch for support at the Allen Institute for Brain Sciences during the writing of this manuscript, J.-S. Chang for cell counting, Y. Huang for genotyping, G. Mancuso for administrative assistance, C. Chiu for lab management, S. Diamond for assistance with FACS, the Single Cell Profiling and Engineering Center (SPEC) in the Beckman Institute at Caltech for initial help for 10x scRNAseq experiments, S. Pease for assistance with transgenic mouse strains, J. Costanza for mouse colony management, members of the Anderson laboratory for helpful comments on this project, and an anonymous reviewer for suggesting social fear testing in group-housed mice. This work was supported by NIH BRAIN Initiative Grants U01MH105982 and U19MH114830 to H.Z. and D.J.A., and NIH grants MH070053 and TR01 OD024686 to D.J.A. and L.C., respectively. D.-W.K. was supported by an HHMI International Student Research Fellowship. D.J.A. is an Investigator of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

D.-W.K., B.T., H.Z., and D.J.A. contributed to the study design. D.-W.K. performed most of the experiments. T.K.K. and K.A.S. prepared sequencing libraries for SMART-seq scRNA-seq. L.T.G and O.F. contributed data visualization. D.-W.K. and T.N.N performed Retro-seq experiments. D.-W.K., Z.Y, L.T.G, L.Y., and L.P. analyzed the scRNA-seq data. D.-W.K. and N.K. performed seqFISH experiments. D.-W.K., N.P., S.S., and L.C. analyzed the seqFISH data. D.-W.K. and L.L. performed retrograde labeling with *c-fos* immunohistochemistry experiments. A.-H.P. and Y.O. developed the tissue preparation protocols for 10x Act-seq. D.J.A. supervised the project. D.-W.K. and D.J.A. wrote the manuscript with contributions from B.T. and H.Z. All authors discussed and commented on the manuscript.

2.9 References

Beier, T., Pape, C., Rahaman, N., Prange, T., Berg, S., Bock, D.D., Cardona, A., Knott, G.W., Plaza, S.M., Scheffer, L.K., et al. (2017). Multicut brings automated neurite segmentation closer to human performance. Nat. Methods 14, 101–102.

Bravo, R. (1990). Growth factor-responsive genes in fibroblasts. Cell Growth Differ. 1, 305–309.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420.

Cachero, S., Ostrovsky, A.D., Yu, J.Y., Dickson, B.J., and Jefferis, G.S.X.E. (2010). Sexual dimorphism in the fly brain. Curr. Biol. 20, 1589–1601.

Campbell, J.N., Macosko, E.Z., Fenselau, H., Pers, T.H., Lyubetskaya, A., Tenen, D., Goldman, M., Verstegen, A.M.J., Resch, J.M., McCarroll, S.A., et al. (2017). A molecular census of arcuate hypothalamus and median eminence cell types. Nat. Neurosci. 20, 484–496.

Canteras, N.S., Simerly, R.B., and Swanson, L.W. (1994). Organization of projections from the ventromedial nucleus of the hypothalamus: a Phaseolus vulgaris-leucoagglutinin study in the rat. J. Comp. Neurol. 348, 41–79.

Chen, P., and Hong, W. (2018). Neural Circuit Mechanisms of Social Behavior. Neuron 98, 16–30.

Chen, R., Wu, X., Jiang, L., and Zhang, Y. (2017). Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. Cell Rep. 18, 3227–3241.

Cheung, C.C., Kurrasch, D.M., Liang, J.K., and Ingraham, H.A. (2013). Genetic labeling of steroidogenic factor-1 (SF-1) neurons in mice reveals ventromedial nucleus of the

hypothalamus (VMH) circuitry beginning at neurogenesis and development of a separate non-SF-1 neuronal cluster in the ventrolateral VMH. J. Comp. Neurol. 521, 1268–1288.

Choi, G.B., Dong, H.W., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., Swanson, L.W., and Anderson, D.J. (2005). Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. Neuron 46, 647–660.

Correa, S.M., Newstrom, D.W., Warne, J.P., Flandin, P., Cheung, C.C., Lin- Moore, A.T., Pierce, A.A., Xu, A.W., Rubenstein, J.L., and Ingraham, H.A. (2015). An estrogen-responsive module in the ventromedial hypothalamus selectively drives sex-specific activity in females. Cell Rep. 10, 62–74.

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

Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell 174, 465–480 e422.

Dhillon, H., Zigman, J.M., Ye, C., Lee, C.E., McGovern, R.A., Tang, V., Kenny, C.D., Christiansen, L.M., White, R.D., Edelstein, E.A., et al. (2006). Leptin directly activates SF1 neurons in the VMH, and this action by leptin is required for normal body-weight homeostasis. Neuron 49, 191–203.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84.

Eng, C.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., and Cai, L. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature 568, 235–239.

Falkner, A.L., Dollar, P., Perona, P., Anderson, D.J., and Lin, D. (2014). Decoding ventromedial hypothalamic neural activity during male mouse aggression. J. Neurosci. 34, 5971–5984.

Franklin, K.B.J., and Paxinos, G. (2008). The mouse brain in stereotaxic coordinates (Academic Press).

Greenberg, M.E., and Ziff, E.B. (1984). Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature 311, 433–438.

Hahn, J.D., Sporns, O., Watts, A.G., and Swanson, L.W. (2019). Macroscale intrinsic network architecture of the hypothalamus. Proc. Natl. Acad. Sci. USA 116, 8018–8027.

Hashikawa, K., Hashikawa, Y., Tremblay, R., Zhang, J., Feng, J.E., Sabol, A., Piper, W.T., Lee, H., Rudy, B., and Lin, D. (2017a). Esr1^{+} cells in the ventromedial hypothalamus control female aggression. Nat. Neurosci. 20, 1580–1590.

Hashikawa, Y., Hashikawa, K., Falkner, A.L., and Lin, D. (2017b). Ventromedial Hypothalamus and the Generation of Aggression. Front. Syst. Neurosci. 11, 94.

Hong, W., Kim, D.W., and Anderson, D.J. (2014). Antagonistic control of social versus repetitive self-grooming behaviors by separable amygdala neuronal subsets. Cell 158, 1348–1361.

Hrvatin, S., Hochbaum, D.R., Nagy, M.A., Cicconet, M., Robertson, K., Cheadle, L., Zilionis, R., Ratner, A., Borges-Monroy, R., Klein, A.M., et al. (2018). Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. Nat. Neurosci. 21, 120–129.

Ishii, K.K., Osakada, T., Mori, H., Miyasaka, N., Yoshihara, Y., Miyamichi, K., and Touhara, K. (2017). A Labeled-Line Neural Circuit for Pheromone-Mediated Sexual Behaviors in Mice. Neuron 95, 123–137 e128.

Jorgenson, L.A., Newsome, W.T., Anderson, D.J., Bargmann, C.I., Brown, E.N., Deisseroth, K., Donoghue, J.P., Hudson, K.L., Ling, G.S., MacLeish, P.R., et al. (2015). The BRAIN Initiative: developing technology to catalyse neuroscience discovery. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20140164.

Kennedy, A., Asahina, K., Hoopfer, E., Inagaki, H., Jung, Y., Lee, H., Remedios, R., and Anderson, D.J. (2014). Internal States and Behavioral Decision-Making: Toward an Integration of Emotion and Cognition. Cold Spring Harb. Symp. Quant. Biol. 79, 199–210.

Kohl, J., Babayan, B.M., Rubinstein, N.D., Autry, A.E., Marin-Rodriguez, B., Kapoor, V., Miyamishi, K., Zweifel, L.S., Luo, L., Uchida, N., and Dulac, C. (2018). Functional circuit architecture underlying parental behaviour. Nature 556, 326–331.

Krause, W.C., and Ingraham, H.A. (2017). Origins and Functions of the Ventrolateral VMH: A Complex Neuronal Cluster Orchestrating Sex Differences in Metabolism and Behavior. Adv. Exp. Med. Biol. 1043, 199–213.

Kunwar, P.S., Zelikowsky, M., Remedios, R., Cai, H., Yilmaz, M., Meister, M., and Anderson, D.J. (2015). Ventromedial hypothalamic neurons control a defensive emotion state. eLife 4, e06633.

Kurrasch, D.M., Cheung, C.C., Lee, F.Y., Tran, P.V., Hata, K., and Ingraham, H.A. (2007). The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. J. Neurosci. 27, 13624–13634.

Lawrence, M., Huber, W., Page` s, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. 9, e1003118.

Lee, H., Kim, D.W., Remedios, R., Anthony, T.E., Chang, A., Madisen, L., Zeng, H., and Anderson, D.J. (2014). Scalable control of mounting and attack by $Esr1^{+}$ neurons in the ventromedial hypothalamus. Nature 509, 627–632.

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., Luginbuhl, D.J., Quake, S.R., and Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. Cell 171, 1206–1220 e1222.

Lin, D., Boyle, M.P., Dollar, P., Lee, H., Lein, E.S., Perona, P., and Anderson, D.J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. Nature 470, 221–226.

Liu, K.S., and Sternberg, P.W. (1995). Sensory regulation of male mating behavior in Caenorhabditis elegans. Neuron 14, 79–89.

Lo, L., Yao, S., Kim, D.W., Cetin, A., Harris, J., Zeng, H., Anderson, D.J., and Weissbourd, B. (2019). Connectional architecture of a mouse hypothalamic circuit node controlling social behavior. Proc. Natl. Acad. Sci. USA 116, 7503–7512.

Lovett-Barron, M., Andalman, A.S., Allen, W.E., Vesuna, S., Kauvar, I., Burns, V.M., and Deisseroth, K. (2017). Ancestral Circuits for the Coordinated Modulation of Brain State. Cell 171, 1411–1423 e1417.

Luiten, P.G.M., ter Horst, G.J., and Steffens, A.B. (1987). The hypothalamus, intrinsic connections and outflow pathways to the endocrine system in relation to the control of feeding and metabolism. Prog. Neurobiol. 28, 1–54.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140.

McDavid, A., Finak, G., Chattopadyay, P.K., Dominguez, M., Lamoreaux, L., Ma, S.S., Roederer, M., and Gottardo, R. (2013). Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. Bioinformatics 29, 461–467.

McHenry, J.A., Otis, J.M., Rossi, M.A., Robinson, J.E., Kosyk, O., Miller, N.W., McElligott, Z.A., Budygin, E.A., Rubinow, D.R., and Stuber, G.D. (2017). Hormonal gain control of a medial preoptic area social reward circuit. Nat. Neurosci. 20, 449–458.

Mickelsen, L.E., Bolisetty, M., Chimileski, B.R., Fujita, A., Beltrami, E.J., Costanzo, J.T., Naparstek, J.R., Robson, P., and Jackson, A.C. (2019). Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. Nat. Neurosci. 22, 642–656.

Millhouse, O.E. (1973a). Certain ventromedial hypothalamic afferents. Brain Res. 55, 89–105.

Millhouse, O.E. (1973b). The organization of the ventromedial hypothalamic nucleus. Brain Res. 55, 71–87.

Moffitt, J.R., Hao, J., Wang, G., Chen, K.H., Babcock, H.P., and Zhuang, X. (2016). High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. Proc. Natl. Acad. Sci. USA 113, 11046–11051.

Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science 362, eaau5324.

Moon, Y.S., Smas, C.M., Lee, K., Villena, J.A., Kim, K.H., Yun, E.J., and Sul, H.S. (2002). Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. Mol. Cell. Biol. 22, 5585–5592.

Morgan, J.I., Cohen, D.R., Hempstead, J.L., and Curran, T. (1987). Mapping patterns of c-fos expression in the central nervous system after seizure. Science 237, 192–197.

Morrell, J.I., and Pfaff, D.W. (1982). Characterization of estrogen-concentrating hypothalamic neurons by their axonal projections. Science 217, 1273–1276.

Musatov, S., Chen, W., Pfaff, D.W., Mobbs, C.V., Yang, X.J., Clegg, D.J., Kaplitt, M.G., and Ogawa, S. (2007). Silencing of estrogen receptor alpha in the ventromedial nucleus of hypothalamus leads to metabolic syndrome. Proc. Natl. Acad. Sci. USA 104, 2501–2506.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature 508, 207–214.

Paul, A., Crow, M., Raudales, R., He, M., Gillis, J., and Huang, Z.J. (2017). Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. Cell 171, 522–539 e520.

Persson-Augner, D., Lee, Y.W., Tovar, S., Dieguez, C., and Meister, B. (2014). Delta-like 1 homologue (DLK1) protein in neurons of the arcuate nucleus that control weight homeostasis and effect of fasting on hypothalamic DLK1 mRNA. Neuroendocrinology 100, 209–220.

Pfaff, D.W. (2017). How the Vertebrate Brain Regulates Behavior: Direct from the Lab (Harvard University Press).

Pfaff, D.W., and Sakuma, Y. (1979a). Deficit in the lordosis reflex of female rats caused by lesions in the ventromedial nucleus of the hypothalamus. J. Physiol. 288, 203–210.

Pfaff, D.W., and Sakuma, Y. (1979b). Facilitation of the lordosis reflex of female rats from the ventromedial nucleus of the hypothalamus. J. Physiol. 288, 189–202.

Picelli, S., Björklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10, 1096–1098.

Remedios, R., Kennedy, A., Zelikowsky, M., Grewe, B.F., Schnitzer, M.J., and Anderson, D.J. (2017). Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex. Nature 550, 388–392.

Roeling, T.A., Veening, J.G., Kruk, M.R., Peters, J.P., Vermelis, M.E., and Nieuwenhuys, R. (1994). Efferent connections of the hypothalamic ‘‘aggression area’’ in the rat. Neuroscience 59, 1001–1024.

Romanov, R.A., Zeisel, A., Bakker, J., Girach, F., Hellysaz, A., Tomer, R., Alpár, A., Mulder, J., Clotman, F., Keimpema, E., et al. (2017). Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. Nat. Neurosci. 20, 176–188.

Sakurai, K., Zhao, S., Takatoh, J., Rodriguez, E., Lu, J., Leavitt, A.D., Fu, M., Han, B.-X., and Wang, F. (2016). Capturing and Manipulating Activated Neuronal Ensembles with CANE Delineates a Hypothalamic Social-Fear Circuit. Neuron 92, 739–753.

Sano, K., Tsuda, M.C., Musatov, S., Sakamoto, T., and Ogawa, S. (2013). Differential effects of site-specific knockdown of estrogen receptor a in the medial amygdala, medial pre-optic area, and ventromedial nucleus of the hypothalamus on sexual and aggressive behavior of male mice. Eur. J. Neurosci. 37, 1308–1319.

Saper, C.B., and Lowell, B.B. (2014). The hypothalamus. Curr. Biol. 24, R1111–R1116.

Saper, C.B., Swanson, L.W., and Cowan, W.M. (1976). The efferent connections of the ventromedial nucleus of the hypothalamus of the rat. J. Comp. Neurol. 169, 409–442.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502.

Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., et al. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. Cell 174, 1015–1030 e1016.

Seung, H.S., and Sümbül, U. (2014). Neuronal cell types and connectivity: lessons from the retina. Neuron 83, 1262–1272.

Shah, S., Lubeck, E., Zhou, W., and Cai, L. (2016). In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. Neuron 92, 342–357.

Shah, S., Takei, Y., Zhou, W., Lubeck, E., Yun, J., Eng, C.L., Koulena, N., Cronin, C., Karp, C., Liaw, E.J., et al. (2018). Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. Cell 174, 363–376 e316.

Silva, B.A., Mattucci, C., Krzywkowski, P., Murana, E., Illarionova, A., Grinevich, V., Canteras, N.S., Ragozzino, D., and Gross, C.T. (2013). Independent hypothalamic circuits for social and predator fear. Nat. Neurosci. 16, 1731–1733.

Simerly, R.B. (2002). Wired for reproduction: organization and development of sexually dimorphic circuits in the mammalian forebrain. Annu. Rev. Neurosci. 25, 507–536.

Sommer, C., Straehle, C., Kothe, U., and Hamprecht, F.A. (2011). Ilastik: Interactive learning and segmentation toolkit. In 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 230–233.

Sternson, S.M. (2013). Hypothalamic survival circuits: blueprints for purposive behaviors. Neuron 77, 810–824.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 e1821.

Tasic, B. (2018). Single cell transcriptomics in neuroscience: cell classification and beyond. Curr. Opin. Neurobiol. 50, 242–249.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. 19, 335–346.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature 563, 72–78.

Tervo, D.G., Hwang, B.Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K.D., Lindo, S., Michael, S., Kuleshova, E., Ojala, D., et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron 92, 372–382.

Tovote, P., Esposito, M.S., Botta, P., Chaudun, F., Fadok, J.P., Markovic, M., Wolff, S.B., Ramakrishnan, C., Fenno, L., Deisseroth, K., et al. (2016). Midbrain circuits for defensive behaviour. Nature 534, 206–212.

Tsuneoka, Y., Tsukahara, S., Yoshida, S., Takase, K., Oda, S., Kuroda, M., and Funato, H. (2017). Moxd1 Is a Marker for Sexual Dimorphismin the Medial Preoptic Area, Bed Nucleus of the Stria Terminalis and Medial Amygdala. Front. Neuroanat. 11, 26.

Wang, L., Chen, I.Z., and Lin, D. (2015). Collateral pathways from the ventromedial hypothalamus mediate defensive behaviors. Neuron 85, 1344–1358.

Wang, L., Talwar, V., Osakada, T., Kuang, A., Guo, Z., Yamaguchi, T., and Lin, D. (2019). Hypothalamic Control of Conspecific Self-Defense. Cell Rep 26, 1747–1758 e1745.

Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58, 236–244.

Wei, Y.C., Wang, S.R., Jiao, Z.L., Zhang, W., Lin, J.K., Li, X.Y., Li, S.S., Zhang, X., and Xu, X.H. (2018). Medial preoptic area in mice is capable of mediating sexually dimorphic behaviors regardless of gender. Nat. Commun. 9, 279.

Welch, J.D., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E.Z. (2019). Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. Cell 177, 1873–1887 e1817.

Wermter, A.K., Scherag, A., Meyre, D., Reichwald, K., Durand, E., Nguyen, T.T., Koberwitz, K., Lichtner, P., Meitinger, T., Scha¨ fer, H., et al. (2008). Preferential reciprocal transfer of paternal/maternal DLK1 alleles to obese children: first evidence of polar overdominance in humans. Eur. J. Hum. Genet. 16, 1126–1134.

Wu, Z., Autry, A.E., Bergan, J.F., Watabe-Uchida, M., and Dulac, C.G. (2014). Galanin neurons in the medial preoptic area govern parental behaviour. Nature 509, 325–330.

Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. Neuron 96, 313–329 e316.

Xu, Y., Nedungadi, T.P., Zhu, L., Sobhani, N., Irani, B.G., Davis, K.E., Zhang, X., Zou, F., Gent, L.M., Hahner, L.D., et al. (2011). Distinct hypothalamic neurons mediate estrogenic effects on energy homeostasis and reproduction. Cell Metab. 14, 453–465.

Xu, X., Coats, J.K., Yang, C.F., Wang, A., Ahmed, O.M., Alvarado, M., Izumi, T., and Shah, N.M. (2012). Modular genetic control of sexually dimorphic behaviors. Cell 148, 596–607.

Yang, C.F., and Shah, N.M. (2014). Representing sex in the brain, one module at a time. Neuron 82, 261–278.

Yang, C.F., Chiang, M.C., Gray, D.C., Prabhakaran, M., Alvarado, M., Juntti, S.A., Unger, E.K., Wells, J.A., and Shah, N.M. (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. Cell 153, 896–909.

Yang, T., Yang, C.F., Chizari, M.D., Maheswaranathan, N., Burke, K.J., Jr., Borius, M., Inoue, S., Chiang, M.C., Bender, K.J., Ganguli, S., et al. (2017). Social Control of Hypothalamus-Mediated Male Aggression. Neuron 95, 955–970 e954.

Yu, J.Y., Kanai, M.I., Demir, E., Jefferis, G.S.X.E., and Dickson, B.J. (2010). Cellular organization of the neural circuit that drives Drosophila courtship behavior. Curr. Biol. 20, 1602–1614.

Zeisel, A., Hochgerner, H., Lonnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular Architecture of the Mouse Nervous System. Cell 174, 999–1014 e1022.

Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nat. Rev. Neurosci. 18, 530–546.

Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049.

Chapter III

FUTURE DIRECTIONS

3.1 Linking electrophysiological and morphological properties to transcriptomic cell types in VMHvl

In previous chapter, I have been mainly focused on mapping behavioral and projectional information onto T-types in VMHvl. However, given the facts that the traditional cell-type characterizations primarily relied on morpho-electric properties and obviously they also contain crucial information/insights for reflecting functional phenotypes of cells (reviewed in Migliore and Shepherd, 2005), the first next step we can take is naturally trying to profile the electrophysiological and morphological cell types (E-type and M-type, respectively) in VMHvl and their relationships with our benchmark transcriptomic data.

Previously, I conducted whole-cell patch clamp recordings in acute brain slices to characterize E-types of VMHvl neurons (random-patching) and VMHvl^{Esr1} (using ChR2-assisted opto-tagging in *Esr1-Cre* mice; see also Figure 1A). At least 4 distinct E-types have been both identified as follows (Figure 1B): 1) "high-frequency firing" (showing significantly higher firing rates than others as increasing injected currents; Figure 1D), 2) "rebound firing" (low-threshold rebound spikes following hyperpolarization), 3) "burst firing" (initial strong bursts followed by inter-spike-interval adaptations), and 4) "regular firing" (regularly-spaced spikes). Interestingly, both VMHvl neurons and VMHvl^{Esr1} showed the similar proportions of each E-type each (Figure 1C), in spite of its relatively small sample size (n = 106).

Recently, a single-cell characterization pipeline, using standardized patch-clamp recordings in acute brain slices, biocytin-based morphological reconstructions, and transcriptomic profiling, was successfully established in Allen Institute for Brain Science (AIBS), and demonstrated in the primary visual cortex (Gouwens et al., 2019). Therefore, based on the information of VMHvl E-types and using the experimental/computational pipelines in AIBS,

we are going to perform this Patch-seq based tri-modal data acquisition in VMHvl, and will see how it would be mapped onto the previous 17 VMHvl T-types. Especially because each E-types in VMHvl has a clear link between its firing pattern and underlying ionic basis (e.g. low voltage activated, T-type calcium channels for Type 2; Ca^{2+}-dependent K^{+} channels for Type 3), the relatively separate mapping of each E-type onto T-types would be expected.

3.2 Neuromodulations in VMHvl and their functional implications

Earlier studies indicated that many neuromodulators (e.g. sex steroids and neuropeptide) could regulate neural activities at the circuit level in multiple brain regions involved in controlling social behaviors (reviewed in Pfaff et al., 2008) such as aggression and mating (reviewed in Nelson and Trainor, 2007). However, there have been very little studied about neuromodulatory influences on VMHvl neurons and their functional implications in innate social behaviors. To this end, I previously characterized neuromodulatory responses of VMHvl neurons, by performing two-photon (2P) Ca^{2+} imaging in VMHvl slices with a genetically encoded calcium indicator GCaMP6s to monitor the population activity responding to a bath application of multiple neuromodulators in VMHvl. Eventually, I have discovered two neuromodulators, arginine vasopressin (AVP) / oxytocin (OXT) and nitric oxide (NO), which exerted opposing effects on VMHvl neurons.

Firstly, it has been extensively reported in various vertebrate species that the neuropeptides AVP and OXT are well-known for modulating social cognition and complex social behaviors including aggression (reviewed in Donaldson and Young, 2008; Stoop, 2012; Stoop, 2014). However, the current knowledge about their roles in modulating aggressive behavior is far from complete, given the fact that previously reported results are highly varied as a result of genetic variations, prior social experience, technical limitations (mostly lack of specificity due to using simple pharmacological or genetic approaches), or unknown other experimental parameters (reviewed in Albers, 2012; De Jong and Neumann, 2017). In detail, some studies suggest that AVP can inhibit as well as promote aggression, depending on which brain areas were manipulated (Ferris and Potegal, 1988; Deville et al. 1996; Ferris et al, 1997; Veenema et al., 2010). Similarly, although the prevailing view is that the increase of OXT level in the brain produces anti-aggressive effects (Calcagnoli et al., 2013; Calcagnoli et al., 2015; also reviewed in De Jong and Neumann, 2017), some studies showed that OXT can promote aggression in a context-dependent manner (Pagani et al., 2015; also reviewed in Beery, 2015), even though some of them appeared to be a developmental effect in part (Takayanagi et al., 2005; Dhakar et al., 2012). Overall, a study using more systematic approaches to

investigate the role of AVP/OXT in aggression should be necessary, and more importantly, when and where endogenous AVP/OXT are released, and how they exert on distributed neural network to modulate natural aggression, are not identified yet.

Previously, I found that AVP/OXT can produce strong evoked responses in VMHvl upon their bath applications in slices (Figure 2A, C), in a dose-dependent manner (Figure 3). This AVP/OXT-evoked activity was also observed in slice electrophysiological recordings (Figure 2B), and mediated by both AVP receptor 1a (Avpr1a) and OXT receptor (Oxtr) (bottom in Figure 2C; a single blocker specific to one of receptors could not block the evoked response; data not shown). The involvement of both receptors seems to be reasonable because both Avpr1a and Oxtr showed relatively high affinity to both AVP and OXT (Koshimizu et al., 2012). More interestingly, there are two different subpopulations of neurons exhibiting two distinct temporal response profiles of AVP/OXT: time-locked vs. persistent (Figure 4). In order to figure out their underlying mechanisms, populational Ca^{2+} responses to AVP/OXT were monitored in the presence of various cation channel blockers. First, synaptic blockers could significantly but slightly decrease the level of persistent activities (Figure 5A, B), indicating that this persistent activity is more likely cell-autonomous (than based on recurrent network properties). Next, non-specific transient receptor potential canonical (Trpc) channel blockers, 2-ABP and flufenamic acid (FFA) could completely block the AVP/OXT-evoked responses whereas non-specific transient receptor potential vanilloid (Trpv) channel blocker, SKF96365 didn't show any effect (Figure 5C). Interestingly, the low concentration of FFA could only block persistent activities while time-locked responses were almost intact. Because Trpc4 and Trpc5 were specifically expressed in our previous VMHvl T-types among all 7 Trpc channels, it would be of particular interest in the future to repeat the same experiment with more Trpc4/5-specific blockers if available. Nevertheless, because persistent neural activity is one of potential neural substrates encoding emotional states (reviewed in Anderson, 2016), it would be extremely interesting to investigate if this persistently active AVP/OXT-responsive subpopulations in VMHvl are relevant to controlling persistent internal states of aggressiveness in vivo.

Secondly, Nitric oxide (NO) is an important gasotransmitter molecule which can act as as a "volume messenger", and associated with signaling across many physiological processes (reviewed in Calabrese et al., 2007, Garthwaite, 2008, and Steinert et al., 2010). Also, it was interestingly reported that male mice with a disruption of neuronal nitric oxide synthase (nNOS) by gene knockout (KO; Nelson et al., 1995) or pharmacological inhibitor (Demas et al., 1997) displayed a hyper-aggressiveness in resident-intruder assay. Although the follow-up study suggested that it would be due to the dysfunction of s serotonergic signaling in several brain regions (Chiavegatto et al., 2001), there is other possibility that the effects of NO can directly manipulate the activity of relevant neuronal populations specific to aggression, including VMHvl. Therefore, we tested this hypothesis by using the same approach (2P Ca^{2+} imaging in VMHvl slices with the bath application of neuromodulators), and surprisingly our preliminary data showed that NO could substantially inhibit baseline activity of VMHvl neurons in slices as opposed to AVP/OXT case, and conversely their baseline activities gradually increased after perfusing nNOS inhibitor (Figure 6). Based on the recent observations on NO as an activity-dependent regulator of synaptic transmission in the postsynaptic neurons (Steinert, et al., 2011), these results suggested that NO might play a role in regulating VMHvl activities in an activity-dependent manner and its disruption can cause neuronal hyperactivities in VMHvl, resulting in abnormal aggressive state as shown in $nNOS^{-/-}$ mice.

The specific aims of the future study are three-fold. First, by sequentially performing 2P Ca^{2+} imaging system and Patch-seq on the same slice preparation, we can get transcriptomic information of AVP/OXT- (especially separate between time-locked vs. persistent), or NO-responsive populations, respectively, and relate them to previous T-types in VMHvl. Next, based on the genetic handles identified from Patch-seq, we can directly manipulate and/or monitor these responsive populations in various ways (e.g. CRISPR-Cas9 knockout, shRNAi, optogenetics) to determine their potential behavioral significances. Lastly, the source of these neuromodulators to VMHvl during aggressive behaviors (or social interactions) should be identified. In case of AVP/OXT, the first place to be tested is AVP/OXT-expressing cells located along the ventral edge of the hypothalamus adjacent to

VMHvl (data not shown) because AVP/OXT can be released from both their axons and dendrites (Ludwig, 1998). For NO, it is very tempting to try to delete a nNOS gene locally in VMHvl using genetic tools including CRISPR-Cas9 knockout.

3.3 Figures

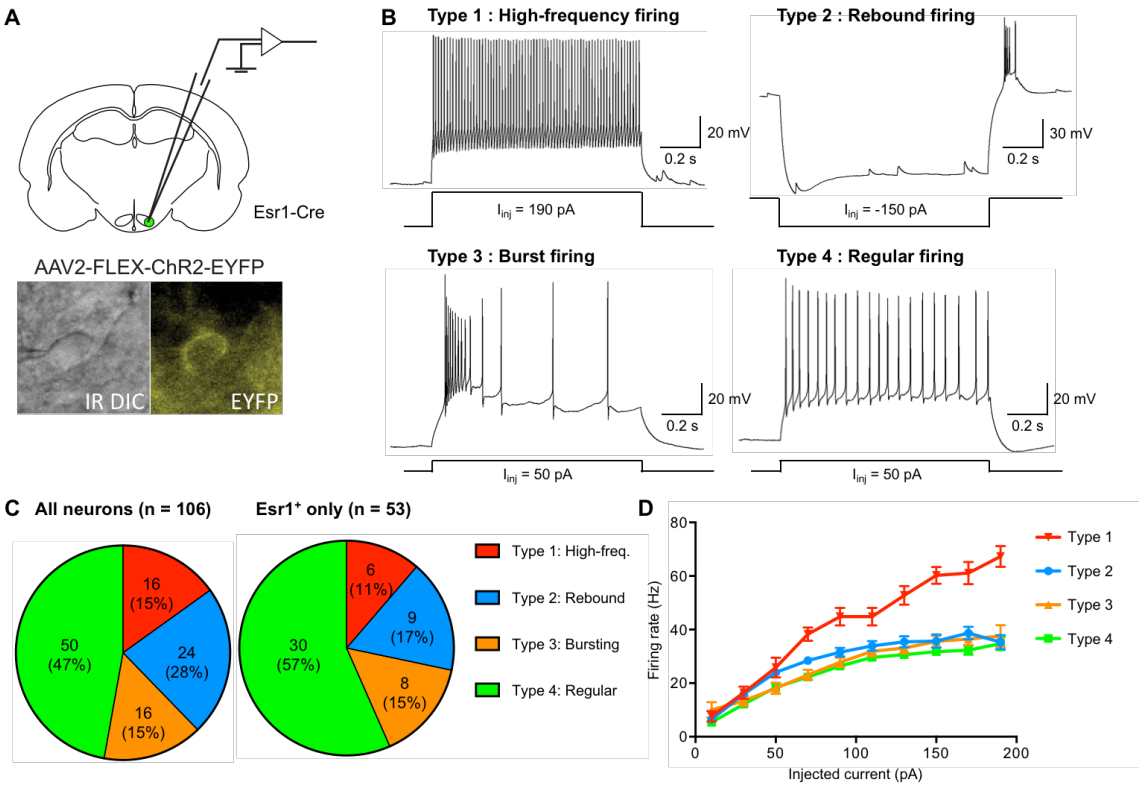

Figure 1. Four distinct electrophysiological types of VMHvl^{Esr1}. (A) Schematic diagram of whole-cell patch clamp recording in VMHvl. In case of VMHvl^{Esr1} recording, cells were optogenetically tagged by injecting cre-dependent adeno-associated virus (AAV) encoding ChR2-EYFP in Esr1-Cre mice. (B-D) Based on firing patterns, 4 distinct eletrophysiological types (E-types) have been identified in VMHvl^{Esr1} (B), and their proportions (C) and firing rates given 1s injected currents (D) were shown.

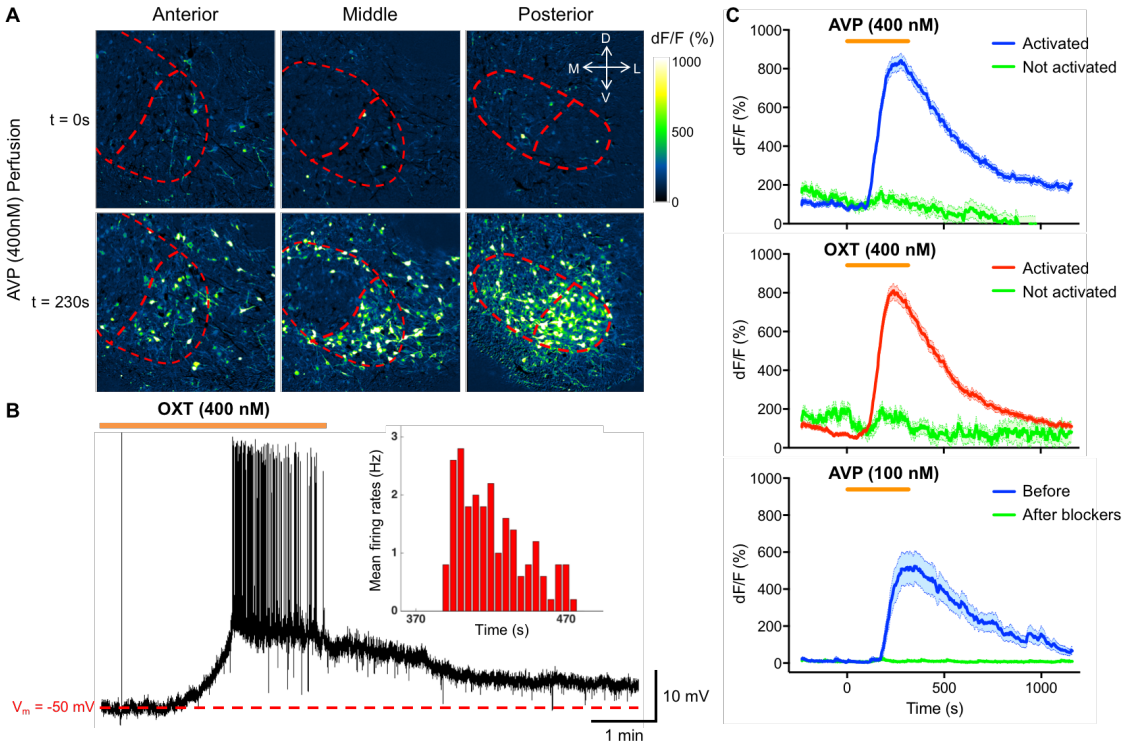

Figure 2. Strong persistent activations of VMHvl neurons induced by vasopressin/oxytocin. (A-B) The acute bath applications vasopressin (AVP) and oxytocin (OXT) were able to trigger the strong neuronal activations in VMHvl slices observed by both two-photon Ca^{2+} imaging with cre-independent GCaMP6s (A), and whole-cell patch clamp recording (B), respectively. (C) Average Ca^{2+} signals (upper: 342 out of 413 neurons were active; middle: 356 out of 433 neurons were active; from 4 mice each) showed that AVP/OXT strongly induced prolong persistent activities, and these responses were completely abolished by the presence of Avpr1a and Oxtr blockers (1 μM SR49059 and 1 μM OTVA; bottom: 19 neurons in total).

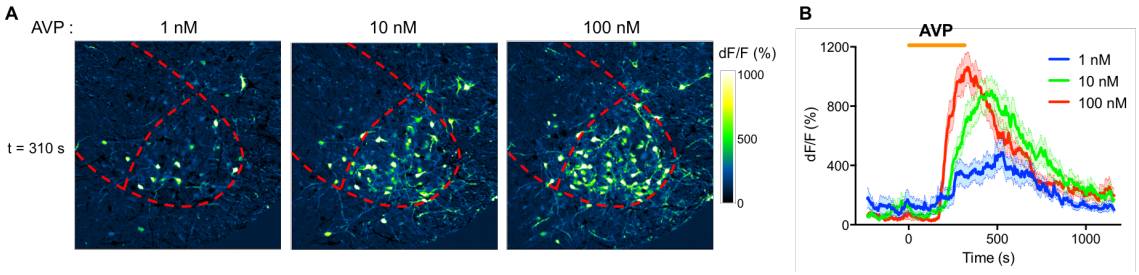

Figure 3. Dose-dependent responses of AVP in VMHvl. (A) Representative fluorescence images varying the concentrations (1, 10, and 100 nM) of AVP (bath-applied at t = 0) at the time point (t = 310 s). Red dashed lines outline VMH and VMHvl. (B) Average dF/F Ca^{2+} signals in $GCaMP6s^{+}$ VMHvl cells (n = 46) at the indicated AVP concentration.

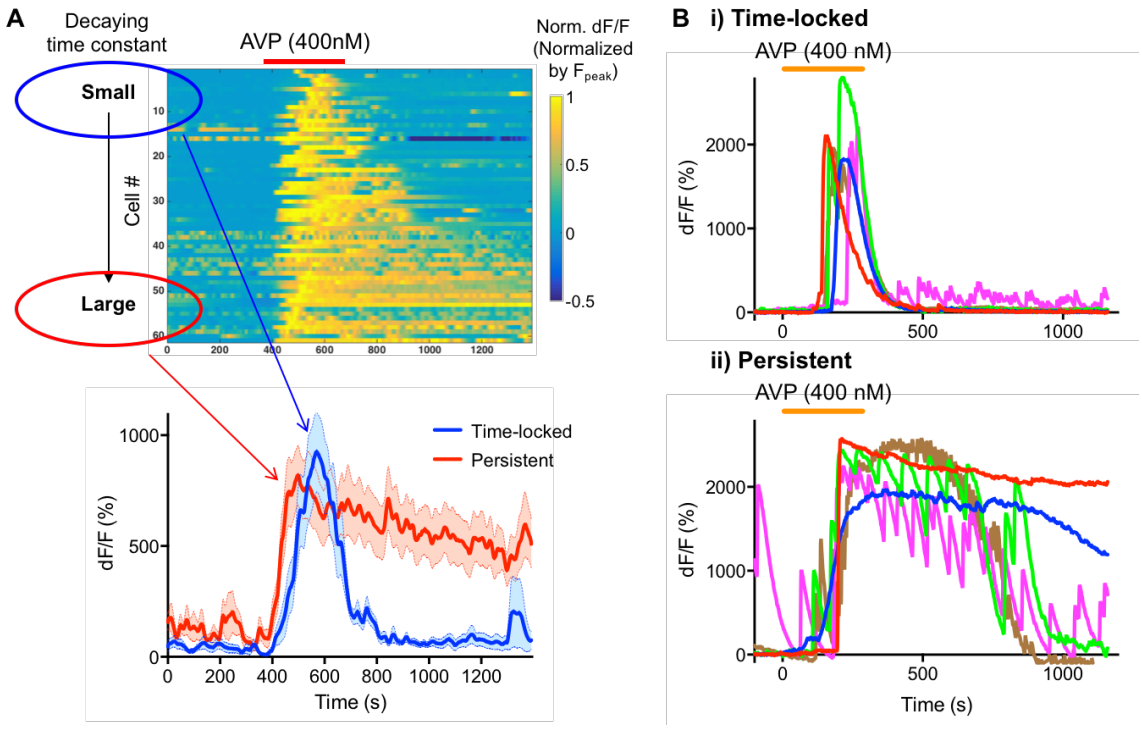

Figure 4. Two types of AVP/OXT responses in VMHvl: Time-locked vs. Persistent. (A) Heatmap visualizes normalized dF/F (by F_{peak}) traces of AVP-responding cells in VMHvl (n = 61). The order of cells was sorted by decaying time constant from F_{peak} (top). Average Ca^{2+} traces of cells with small (blue) and large (red) decaying time constants (bottom; top 15 cells were selected each) illustrated two types of VMHvl cells responding to AVP/OXT: time-locked vs. persistent. (B) Five representative Ca^{2+} traces from both time-locked and persistent responses.

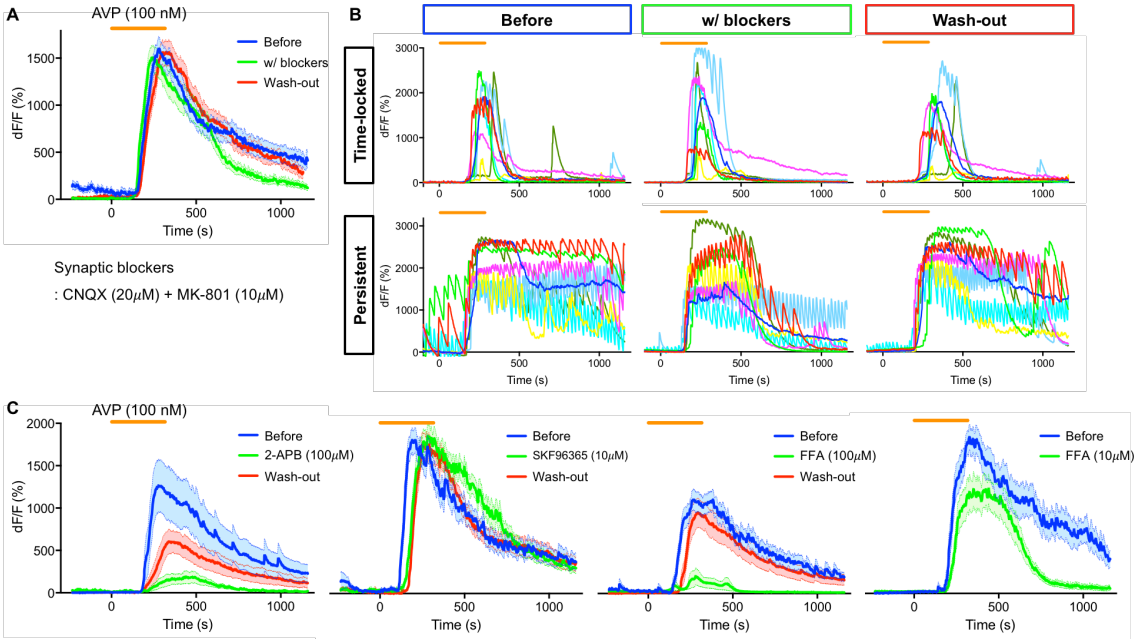

Figure 5. Underlying mechanisms eliciting persistent neural activities in VMHvl by AVP/OXT. Averaged (A, C) and representative individual (B) Ca^{2+} traces for the bath applications of AVP (100 nM) with or without pharmacological blockers of different ionic channels were shown. (A-B) In spite of presence of major synaptic blockers (20 μM CNQX for AMPA and 10 μM MK-801 for NMDA receptors, respectively), AVP response profiles were pretty much similar to those before/after the blockers except mild but significant reductions of persistent activities (A; n = 45), which were also observed in the individual traces (middle in B). (C) Multiple non-specific transient receptor potential (Trp) channel blockers were tested to see which trp channels were involved in the persistent VMHvl activities by AVP/OXT. 2-APB (100 μM; n = 14), SKF96365 (10 μM; n = 20), and FFA (100 and 10 μM; n = 30 and 17).

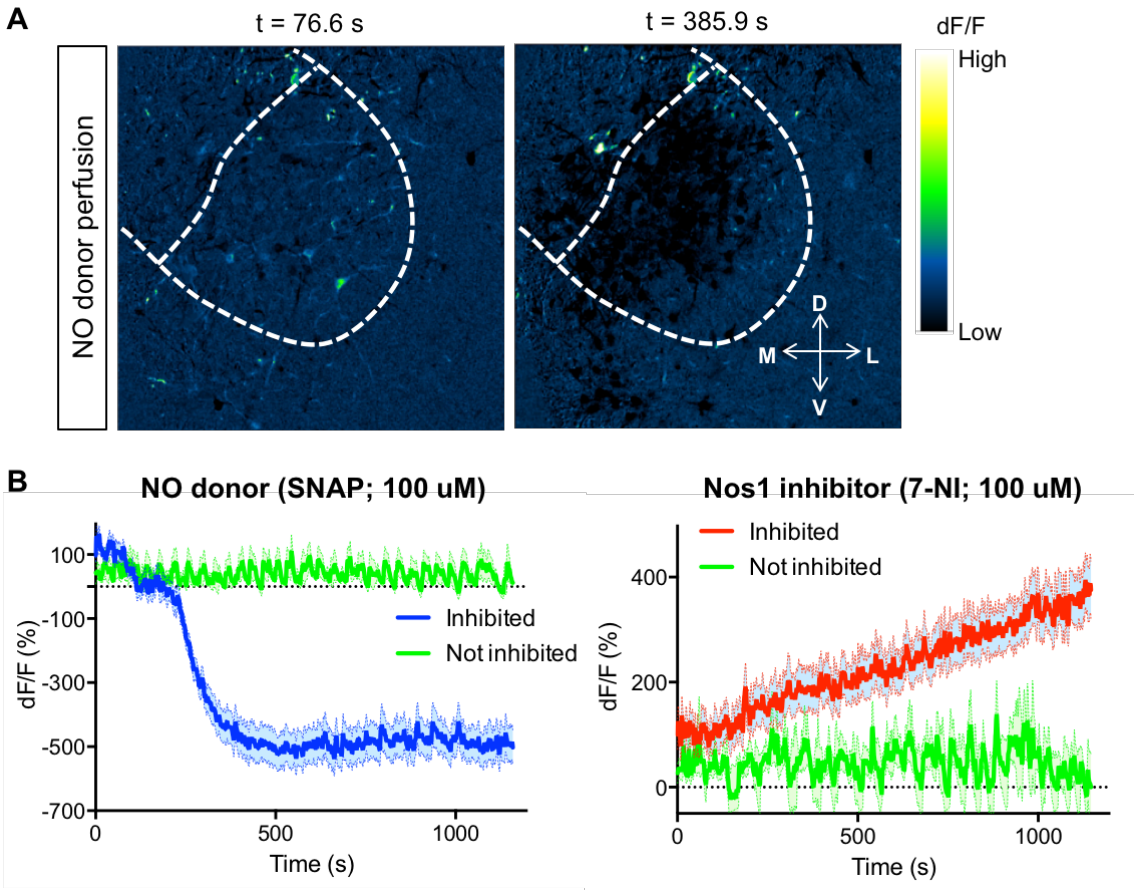

Figure 6. Neuromodulatory effects of nitric oxide (NO) on VMHvl. (A) Representative fluorescence images at the indicated time points after NO donor (100 μM SNAP) was applied in bath (t = 0). White dashed lines outline VMH and VMHvl. (B) Average dF/F Ca^{2+} signals in $GCaMP6s^{+}$ VMHvl cells in the presence of either No donor (left) or Nos1 inhibitor (right; 100 μM 7-NI), respectively (60 out of 86 VMHvl cells were inhibited by NO donors; n = 2 mice).

3.4 References

Albers, H.E. (2012). The regulation of social recognition, social communication and aggression: vasopressin in the social behavior neural network. *Horm. Behav.* 61, 283-292.

Beery, A.K. (2015). Antisocial oxytocin: complex effects on social behavior. *Curr Opin Behav. Sci.* 6, 174-182.

Calabrese, V., Mancuso, C., Calvani, M., Rizzarelli, E., Butterfield, D.A., and Stella, A.M. (2007). Nitric oxide in the central nervous system: neuroprotection versus neurotoxicity. *Nat. Rev. Neurosci.* 8, 766-775.

Calcagnoli, F., Stubbendorff, C., Meyer, N., de Boer, S.F., Althaus, M., and Koolhaas, J.M. (2015). Oxytocin microinjected into the central amygdaloid nuclei exerts anti-aggressive effects in male rats. *Neuropharmacology.* 90, 74-81.

Chiavegatto, S., Dawson, V.L., Mamounas, L.A., Koliatsos, V.E., Dawson, T.M., and Nelson, R.J. (2001). Brain serotonin dysfunction accounts for aggression in male mice lacking neuronal nitric oxide synthase. *Proc. Natl. Acad. Sci. USA* 98, 1277-1281.

De Jong, T.R., Neumann, I.D. (2017). Oxytocin and Aggression. In: Hurlemann R., Grinevich V. (eds) Behavioral Pharmacology of Neuropeptides: Oxytocin. *Curr. Top. Behav. Neurosci.* 35, Springer, Cham

Delville, Y., Mansour, K.M., and Ferris, C.F. (1996). Testosterone facilitates ag- gression by modulating vasopressin receptors in the hypothalamus. *Physiol. Behav.* 60, 25–29.

Demas, G.E., Eliasson, M.J., Dawson, T.M., Dawson, V.L., Kriegsfeld, L.J., Nelson, R.J., and Snyder, S.H. (1997). Inhibition of neuronal nitric oxide synthase increases aggressive behavior in mice. *Mol. Med.* 3, 610-616.

Dhakar, M.B., Rich, M.E., Reno, E.L., Lee, H.J., and Caldwell, H.K. (2012). Heightened aggressive behavior in mice with lifelong versus postweaning knockout of the oxytocin receptor. *Horm. Behav.* 62, 86-92.

Donaldson, Z.R., and Young, L.J. (2008). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* 322, 900-904.

Ferris, C.F., Melloni, R.H., Koppel, G., Perry, K.W., Fuller, R.W., Delville, Y. (1997). Vasopressin/serotonin interactions in the anterior hypothala- mus control aggressive behavior in golden hamsters. *J. Neurosci.* 17, 4331– 4340.

Ferris, C.F., Potegal, M. (1988). Vasopressin receptor blockade in the anterior hypothalamus suppresses aggression in hamsters. *Physiol. Behav.* 44, 235–239.

Garthwaite, J. (2008). Concepts of neural nitric oxide-mediated transmission. *Eur. J. Neurosci.* 27, 2783-2802.

Gouwens, N.W., Sorensen, S.A., Berg, J., Lee, C., Jarsky, T., Ting, J., Sunkin, S.M., Feng, D., Anastassiou, C.A., Barkan, E., et al. (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat. Neurosci.* 22, 1182-1195.

Koshimizu, T.A., Nakamura, K., Egashira, N., Hiroyama, M., Nonoguchi, H., and Tanoue, A. (2012). Vasopressin V1a and V1b receptors: from molecules to physiological systems. *Physiol Rev* 92, 1813-1864.

Ludwig, M. (1998). Dendritic release of vasopressin and oxytocin. *J. Neuroendocrinol.* 10, 881–895.

Migliore, M., and Shepherd, G.M. (2005). An integrated approach to classifying neuronal phenotypes. *Nat. Rev. Neurosci*. 6, 810–818.

Nelson, R.J., Demas, G.E., Huang, P.L., Fishman, M.C., Dawson, V.L., Dawson, T.M., and Snyder, S.H. (1995). Behavioural abnormalities in male mice lacking neuronal nitric oxide synthase. *Nature* 378, 383-386.

Nelson, R.J., and Trainor, B.C. (2007). Neural mechanisms of aggression. *Nat. Rev. Neurosci.* 8, 536-546.

Pagani, J.H., Williams Avram, S.K., Cui, Z., Song, J., Mezey, E., Senerth, J.M., Baumann, M.H., and Young, W.S. (2015). Raphe serotonin neuron-specific oxytocin receptor knockout reduces aggression without affecting anxiety-like behavior in male mice only. *Genes Brain Behav.* 14, 167-176.

Pfaff, D., et al. (2008). Hormones and Social Behavior (Springer).

Steinert, J.R., Chernova, T., and Forsythe, I.D. (2010). Nitric oxide signaling in brain function, dysfunction, and dementia. *Neuroscientist* 16, 435-452.

Steinert, J.R., Robinson, S.W., Tong, H., Haustein, M.D., Kopp-Scheinpflug, C., and Forsythe, I.D. (2011). Nitric oxide is an activity-dependent regulator of target neuron intrinsic excitability. *Neuron* 71, 291-305.

Steinman, M.Q., Duque-Wilckens, N., and Trainor, B.C. (2019). Complementary Neural Circuits for Divergent Effects of Oxytocin: Social Approach Versus Social Anxiety. *Biol. Psychiatry* 85, 792-801.

Stoop, R. (2012). Neuromodulation by oxytocin and vasopressin. *Neuron* 76, 142-159.

Stoop, R. (2014). Neuromodulation by oxytocin and vasopressin in the central nervous system as a basis for their rapid behavioral effects. *Curr. Opin. Neurobiol.* 29, 187-193.

Takayanagi, Y., Yoshida, M., Bielsky, I.F., Ross, H.E., Kawamata, M., Onaka, T., Yanagisawa, T., Kimura, T., Matzuk, M.M., Young, L.J., et al. (2005). Pervasive social deficits, but normal parturition, in oxytocin receptor-deficient mice. *Proc Natl Acad Sci USA* 102, 16096-16101.

Veenema, A.H., Beiderbeck, D.I., Lukas, M., and Neumann, I.D. (2010). Distinct correlations of vasopressin release within the lateral septum and the bed nucleus of the stria terminalis with the display of intermale aggression. *Horm. Behav.* 58, 273-281.