

NEUROSCIENCE

Sensory coding mechanisms revealed by optical tagging of physiologically defined neuronal types

Donghoon Lee, Maiko Kume, Timothy E. Holy*

Neural circuit analysis relies on having molecular markers for specific cell types. However, for a cell type identified only by its circuit function, the process of identifying markers remains laborious. We developed physiological optical tagging sequencing (PhOTseq), a technique for tagging and expression profiling of cells on the basis of their functional properties. PhOTseq was capable of selecting rare cell types and enriching them by nearly 100-fold. We applied PhOTseq to the challenge of mapping receptor-ligand pairings among pheromone-sensing neurons in mice. Together with *in vivo* ectopic expression of vomeronasal chemoreceptors, PhOTseq identified the complete combinatorial receptor code for a specific set of ligands.

Molecular markers are a powerful tool for labeling and analyzing neuronal cell types. However, in many cases, a single marker is insufficient to define a distinct cell type and may label a few, or a few hundred, physiologically distinguishable cell types (1–3). In such cases, one wishes to select specific physiological populations and discover their molecular identities. However, tools for proceeding from function to molecular markers are not fully mature. Markers such as *c-fos* label active neurons promiscuously across many cell types (4). Cells can be chosen for sequencing based on patch recording (5–8); however, this approach faces obstacles for rare cell types or when one needs many cells of the same type to profile low-abundance transcripts.

We reasoned that the key bottlenecks of this pipeline could be replaced by an all-optical approach. One prototype is CaMPARI, which uses the simultaneous presence of calcium and excitation by short-wavelength light under the control of the experimenter (9). Although CaMPARI allows one to control the labeling of excited neurons temporally, as with *c-fos*, any active neuron will be labeled, so this may comprise dozens or more cell types. There is considerable need for tools that provide the specificity of patch-seq with the convenience and high throughput of optical methods.

Here, we report a method called physiological optical tagging sequencing (PhOTseq). PhOTseq allows specific cell types to be defined intersectionally through their physiological activity under diverse conditions. We exploited two coexpressed fluorescent proteins, GCaMP and photoactivatable mCherry (PAmCherry) (10–12); GCaMP permits recording neuronal activity by large-scale calcium imaging, and PAmCherry enables long-term labeling of selected neurons by photoactivation

(PA). Using these reporters, we performed calcium imaging with a diverse set of stimuli or experiences and then selected a specific collection of cells for tagging by PA. Last, the tagged neurons were harvested and profiled for their mRNA expression (Fig. 1A).

We created several plasmids and viruses, all expressing variants of GCaMP-2A-PAmCherry. The self-cleaving 2A peptide (13) allowed both proteins to be expressed from a single mRNA; the tight coupling facilitated ratiometric analysis to compensate for variability in expression. Initial *in vitro* and *in vivo* experiments demonstrated the feasibility of performing both calcium imaging and phototagging by using this reporter (figs. S1 and S2). To test the performance of the complete pipeline (Fig. 1A), we turned to mouse vomeronasal sensory neurons (VSNs), whose primary function is pheromone sensing and social communication. VSNs exhibit dense cell packing (14) (thus challenging phototagging accuracy), extreme functional heterogeneity, and an unambiguous relationship between molecular identity and physiological function (15–18). We generated a transgenic tetO-GCaMP5g-2A-PAmCherry mouse and drove expression in all VSNs through olfactory marker protein (OMP)-internal ribosomal entry site (IRES)-tetracyclin transactivator (tTA) (fig. S3A). We observed calcium responses from the VSNs, and two-photon PA resulted in phototagging at single-cell resolution (figs. S3 and S4).

We evoked combinatorial VSN responses using two ligands: 5-androsten-3 β , 17 β -diol disulfate (A7864) and 1,3,5(10)-estratrien-3, 17 β -diol disulfate (E1050). A custom online algorithm automatically segmented responsive neurons; out of 430 neurons responding to at least one ligand, 192 responded to both ligands. Voxels corresponding to these dually responsive neurons were chosen to create a mask so that only these voxels were illuminated during PA (Fig. 1, B to D). Photoactivated neurons were readily distinguished from the background, and PA tagged an av-

erage of 1.32 cells per region (Fig. 1E and fig. S5). The photoactivated vomeronasal epithelia (VNEs) were subsequently subjected to dissociation and fluorescence-activated cell sorting (FACS) (Fig. 1, F and G). In the FACS analysis, four to five abundant clusters were observed, of which only one (P1) was almost entirely dependent on PA. This population represents the one selected later for sequencing.

We used PhOTseq to begin unraveling the logic of chemosensation. Fewer than a dozen vomeronasal receptors (VRs) have ligands of known structure (5, 19, 20), and these are scattered widely across the gene family. We reasoned that a saturation analysis of VRs responsible for encoding a focused region of “chemical space” will provide insights into the molecular logic of chemosensation. First, we examined functional types of VSNs broadly by light-sheet calcium imaging while presenting 15 different sulfated steroids, of which six were also delivered over a range of concentrations, for a total of 27 different chemosensory stimuli. Approximately 15,000 neurons, ~10% of the total VSN population per hemisphere, were visualized in each imaging volume (3). From three imaging volumes, more than 5000 neurons responded to at least one of the ligands, and they fell naturally into 20 different response types (Fig. 2, A and B). From this reference clustering, we chose to phototag the most abundant cell type, cluster 1, and three others showing similar chemoreceptive fields (clusters 2, 3, and 6) (fig. S6), among which cluster 3 comprised fewer than 2% of the responsive neurons and presumably ~0.2% of the total VSN population (Fig. 2B). As a control group, we selected one of the most dissimilar cell types, cluster 19, which strongly responded to sulfated pregnanolones.

For cell type selection under two-photon microscopy, on the basis of Fig. 2B, we used a subset of stimuli that distinguish cells among these five selected clusters (figs. S7 and S8 and movie S1). After online two-photon PA, we collected photoactivated or nonphotoactivated experimental control cells by FACS and performed single-cell RNA sequencing. We obtained a total of 622 qualified cells (fig. S9), among which 61, 44, 49, 40, and 65 photoactivated cells were from experiments aimed at functionally defined cluster 1, cluster 2, cluster 3, cluster 6, and cluster 19 cells, respectively (Fig. 2C).

Physiological responses of chemosensory neurons are thought to be largely determined by the VR genes they express (21). Therefore, our investigation was focused on the expression of VR genes. Among photoactivated cells, any given cluster exhibited substantial enrichment of at least one VR gene, with different VR genes being enriched in different clusters (Fig. 2C). By contrast, among control cells, VR gene expression was sporadic (fig. S10). Because our targeting accuracy was less than 100% (fig. S5C), we

Department of Neuroscience, Washington University School of Medicine in St. Louis, St. Louis, MO, USA.

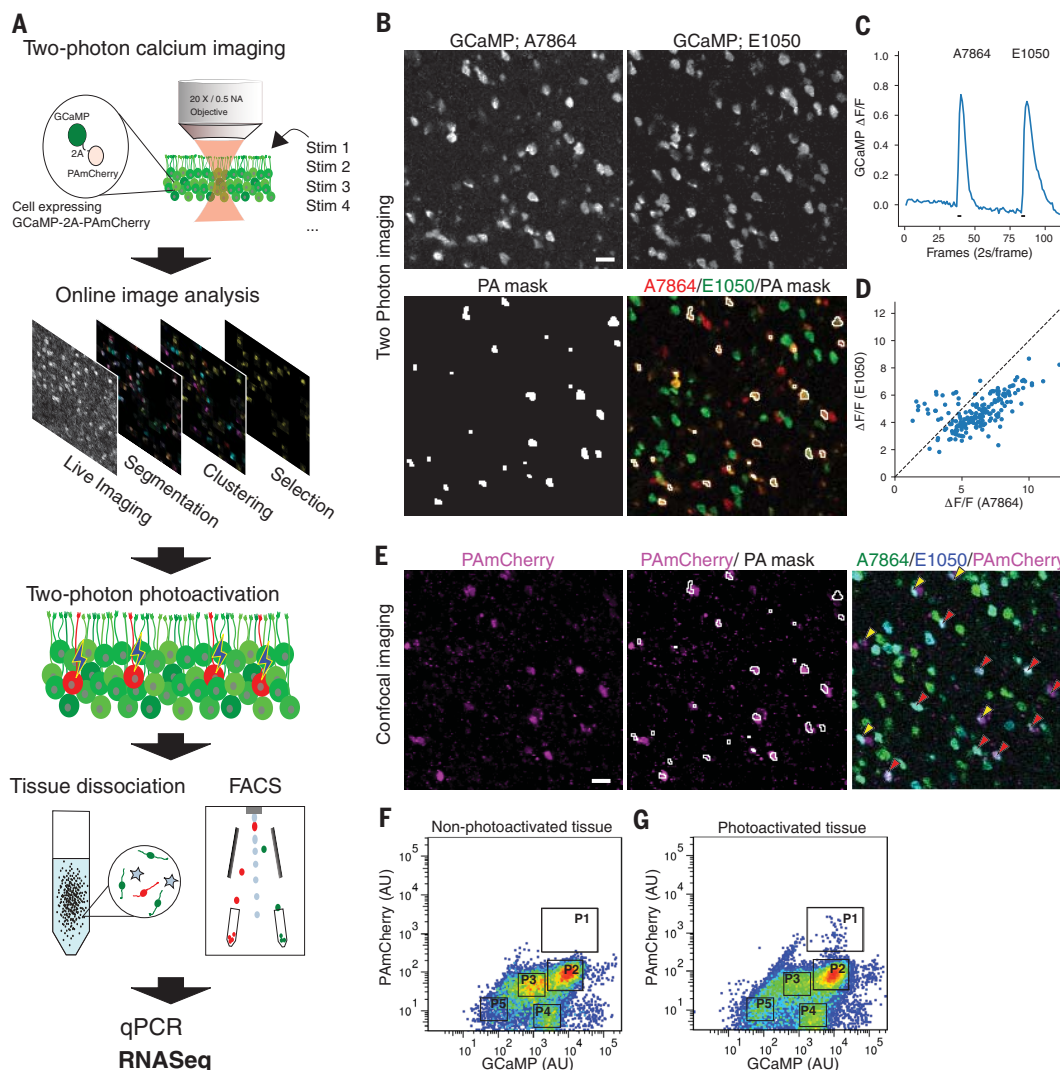
*Corresponding author. Email: holy@wustl.edu

Fig. 1. Two-photon PA can tag neurons chosen by activity pattern.

(A) Workflow of PhOTseq. NA, numerical aperture; qPCR, quantitative polymerase chain reaction; RNAseq, RNA sequencing; Stim, stimulus.

(B to D) Two-photon calcium imaging of the VNE explanted from a GCaMP5g-p2a-PAmCherry transgenic mouse. (B) Representative images of calcium chemosensory response evoked by either 10 μ M A7864 (top left) or 10 μ M E1050 (top right). PA mask (bottom) selects cells responsive to both stimuli. Scale bar, 20 μ m. (C) The average GCaMP intensity obtained from masked cells. Black bars: delivery time course of stimuli. (D) Individual responses ($\Delta F/F$) among masked cells to 10 μ M A7864 and 10 μ M E1050. (E) Confocal imaging after online two-photon PA. PAmCherry signals (left), PAmCherry signals with two-photon PA mask (middle), and PAmCherry signals with calcium responses to either 10 μ M A7864 (cyan) or 10 μ M E1050 (green) (right). Red and yellow arrow heads: the examples of PA mask regions associated with single and double PAmCherry-positive

neurons, respectively. Scale bar, 20 μ m. (F and G) FACS analysis of cells from nonphotoactivated tissue (F) or photoactivated tissue (G). Clusters were marked as P1, P2, P3, P4, and P5. P1 and P2 represent a photoactivated and an experimental control group, respectively. Of the total population, P1 and P2 accounted for 0.013% and 33.9% (F) or 0.218% and 37.3% (G), respectively. AU, arbitrary unit.



also expected some sporadic expression among cluster-selected photoactivated neurons (Fig. 2C). When the type of a VSN was defined by its maximally expressed VR gene, only one abundant type was found among cluster-selected photoactivated neurons (Fig. 2D) [cluster 1: *VmnIr89* type (25%); cluster 2: *VmnIr86* type (30%); cluster 3: *VmnIr78* type (27%); cluster 6: *VmnIr237* type (50%); cluster 19: *VmnIr58* type (58%)]. These five VR genes were the only significantly enriched receptor gene in each photoactivated group compared with all the other cells (Fig. 2E). During these analyses, we updated the *VmnIr237* gene model because the read coverage of *VmnIr237* and cloning revealed 3' untranslated regions missing from existing gene models (fig. S11 and data file S1). We also

unexpectedly observed coexpression of *VmnIr85* in *VmnIr86* neurons (Fig. 2F and fig. S12, A to C). Including data from our control and sporadic cells, coexpression was also observed for several other VR gene pairs, in each case consisting of a genomically adjacent pair (fig. S12, D to F). Taken together, we identified six putative VR genes mediating a focused set of responses.

To test whether PhOTseq identified true receptor-ligand pairs, we performed a gain-of-function study. Because of abnormal localization of VR proteins (22), in vitro heterologous expression systems have had little success, particularly for the V1R family (23). We reasoned that a VSN cell's endogenous machinery would allow functional expression of a VR gene; seeking a more efficient route than making a trans-

genic mouse (5), we explored a virus-mediated approach. Among the viral vectors we tested, intravenous injection of rAAV2/8-CAG-green fluorescent protein (GFP) was able to induce expression of GFP in a subset of VSNs (fig. S13). We delivered rAAV2/8-CAG-GCaMP-2A-VR, in which VR was one of the genes identified by PhOTseq, and for each obtained multiple neurons that showed statistically significant responses to at least one ligand (Fig. 3, A and B, and movie S2). In each case, the dominant response pattern matched the expected PhOTseq response pattern (Fig. 3, C and D). We also tested *VmnIr85* because this gene was found to be coexpressed with *VmnIr86* (Fig. 2C). The response pattern matched that of cluster 4, which was another group that was similar

Fig. 2. PhOTseq-identified VR genes of overlapping cell types. (A) Example calcium traces (normalized ΔF) from two cells. Amplitude was normalized by the maximum amplitude of all cells recorded.

Ligands are listed at the top. OCPI, objective coupled planar illumination. (B) Neuronal responses to sulfated steroids. Cells are on columns, and stimuli are on rows. If not indicated, the ligand concentration is 10 μM . The color bar indicates normalized response.

Cluster identities are reported at the top. Asterisks indicate PhOTseq target cell types. Asterisks in parentheses indicate functional type whose receptor identity was discovered during analysis by ectopic expression in Fig. 3.

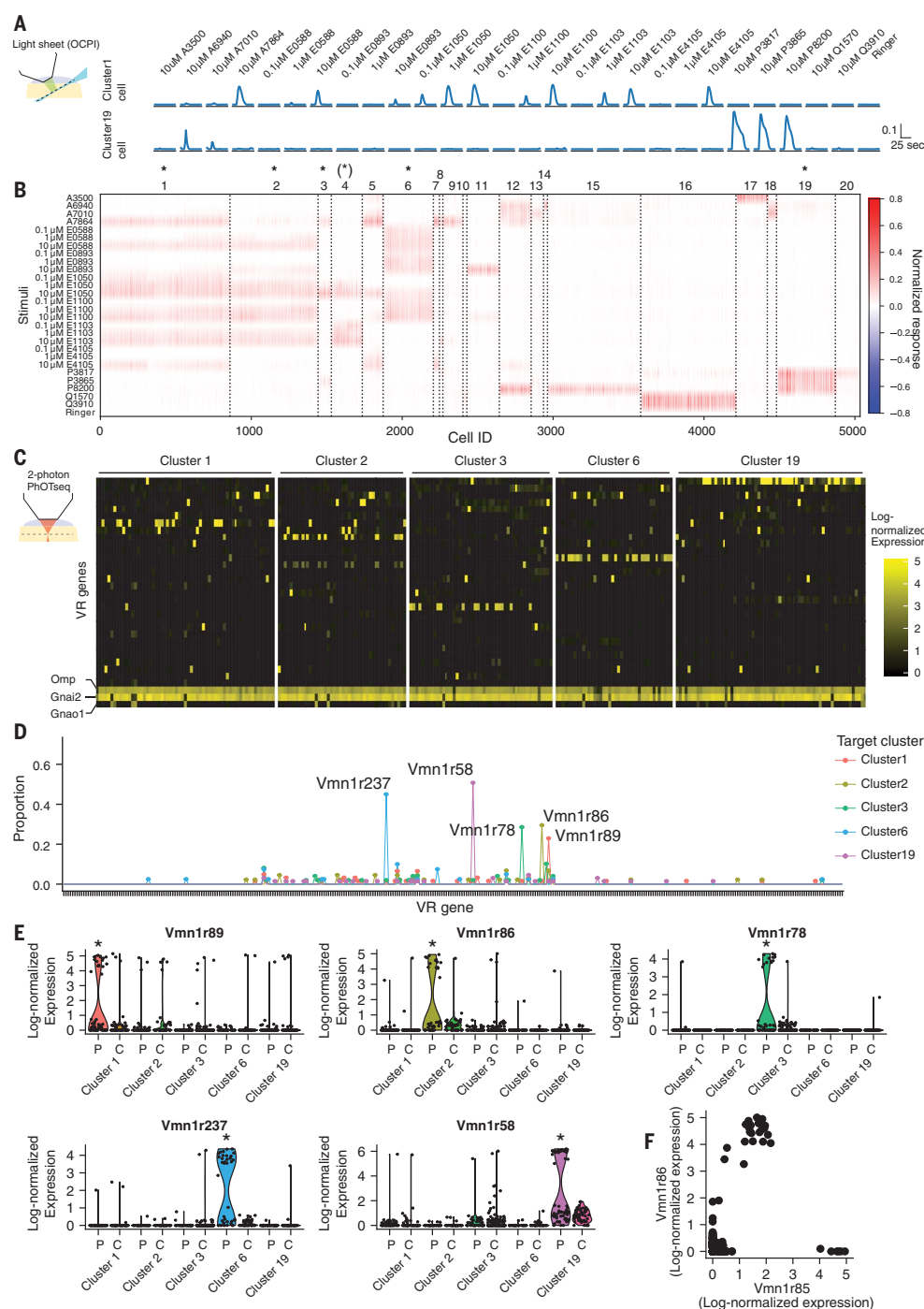
(C) Expression of the 30 most highly expressed VR genes when averaged across all sequenced cells; also shown are three marker genes (*Omp*, *Gnai2*, and *Gnao1*).

Cells are on columns, and genes are on rows. PhOTseq-targeted functional types are shown at the top; cells belonging to these functional clusters were specifically photoactivated and sequenced. The color bar indicates the log-normalized expression level. ID, identification.

(D) Proportion of a VSN type in each group shown in (C). Each tick on the x axis represents a different VR gene. Each functional type exhibited only one common VSN type.

(E) Expression of five VR genes across different experiment groups. "P" indicates photoactivated cells, and "C" indicates nonphotoactivated control cells. $*P_{\text{adj}} < 0.01$ (Wilcoxon rank sum test; $P_{\text{adj}} < 0.01$) (Wilcoxon rank sum test; $P_{\text{adj}} = 2.7 \times 10^{-11}$, 2.4×10^{-10} , 1.1×10^{-20} , 1.3×10^{-40} , and 2.7×10^{-30} ; average fold difference: 6.2, 23.8, 80.7, 75.7, and 45.9 for *Vmn1r89*, *Vmn1r86*, *Vmn1r78*, *Vmn1r237*, and *Vmn1r58*, respectively).

(F) Single-cell expression of *Vmn1r85* and *Vmn1r86* is nonexclusive.



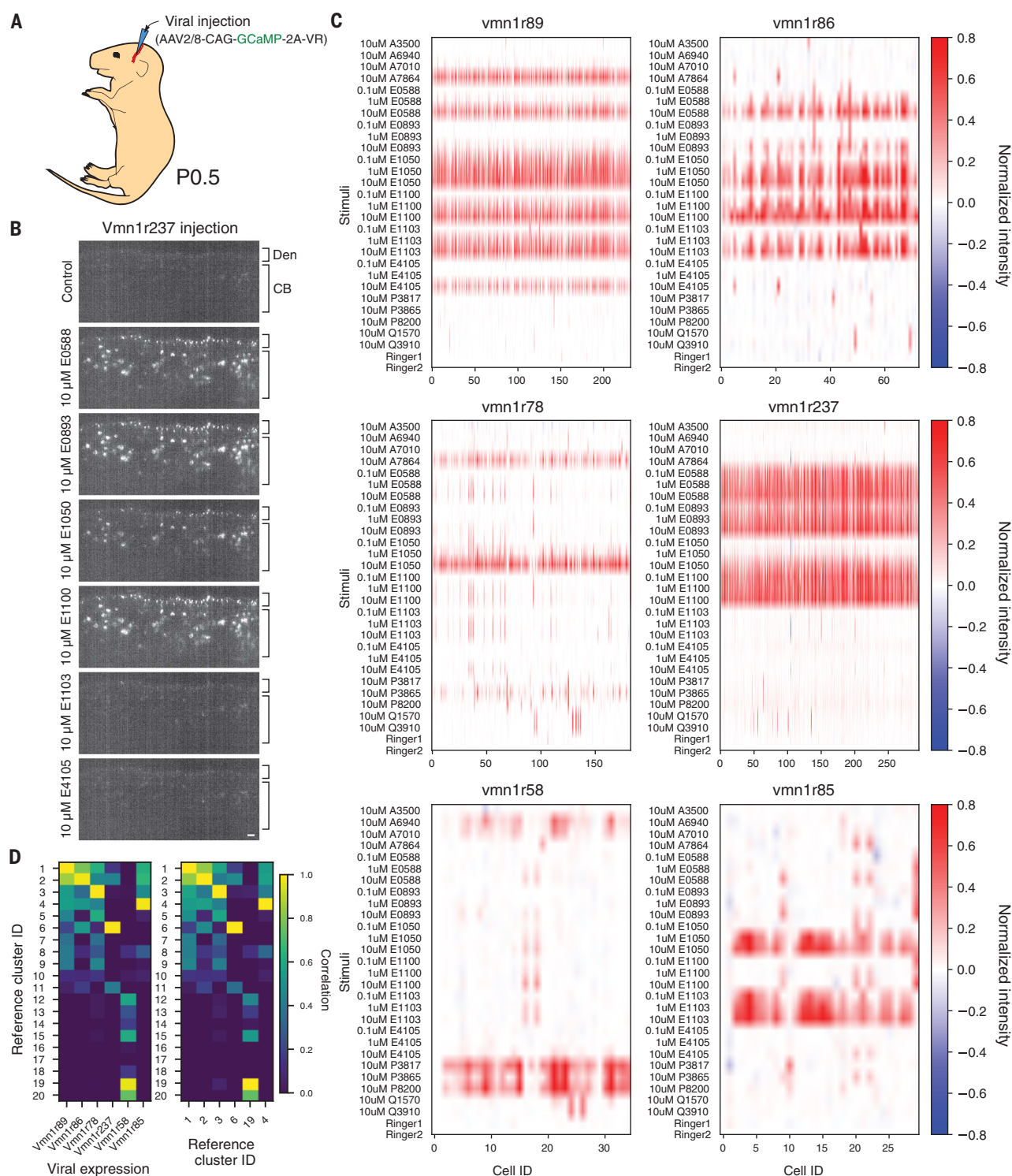


Fig. 3. Ectopic expression-enabled functional analysis of VRs. (A) Expression by means of AAV injection in the temporal vein of newborn pups. AAV, adeno associated virus; P0.5, postnatal day 0.5. (B) Optical section from light-sheet calcium imaging after the ectopic expression of GCaMP-2A-Vmn1r237 in response to different ligands. The apical layer is occupied by dendritic tips (Den)

with cell bodies (CB) below. Scale bar, 20 μ m. (C) Calcium response after ectopic expression of GCaMP-2A-Vmn1r89, -Vmn1r86, -Vmn1r78, -Vmn1r237, -Vmn1r58, or -Vmn1r85. Ligands are identical to those in Fig. 2B. (D) Pairwise correlation between the reference clustering and the ectopic responses shown in (C) (left) or autocorrelogram of the reference clustering (right).

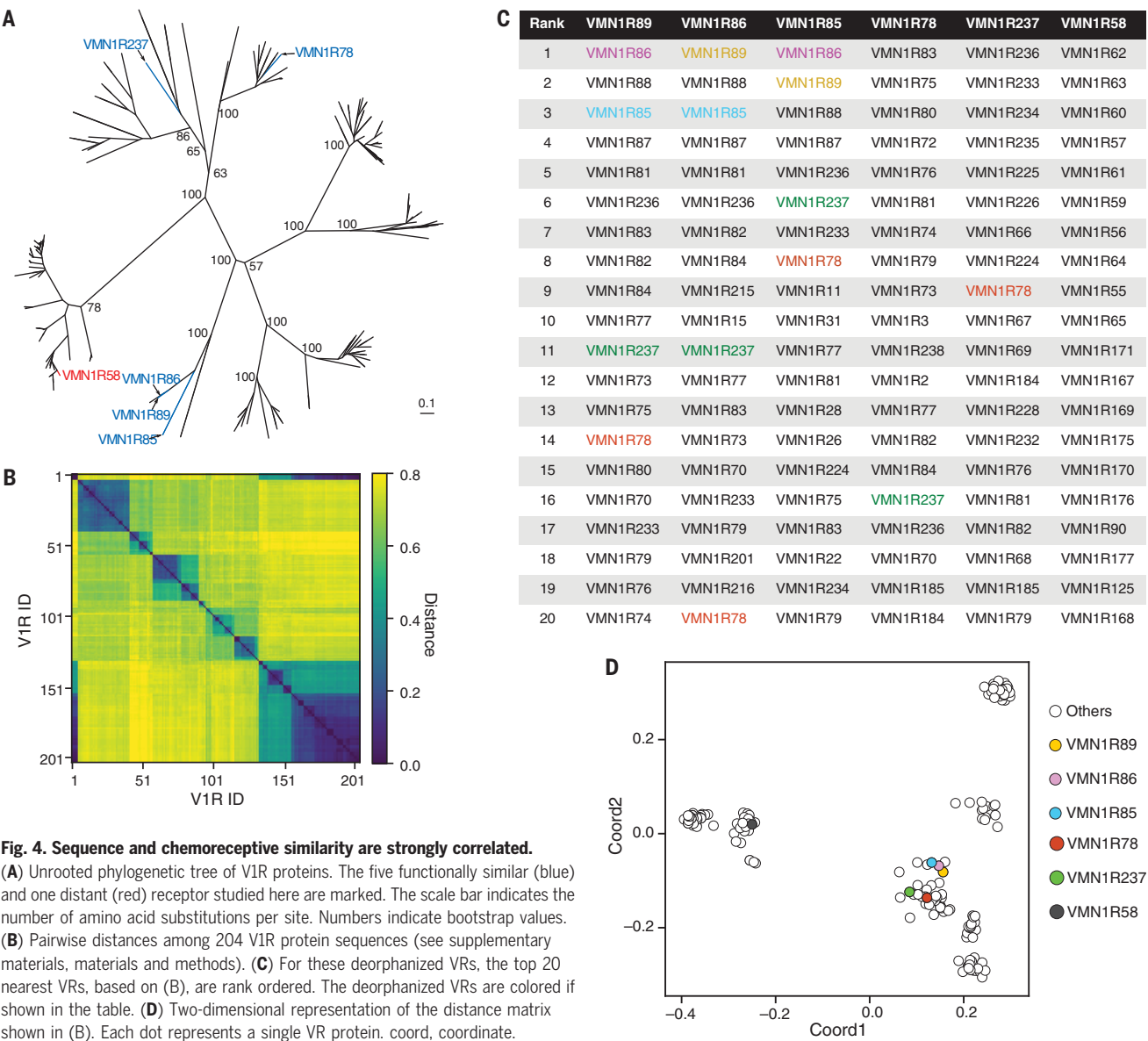


Fig. 4. Sequence and chemoreceptive similarity are strongly correlated. (A) Unrooted phylogenetic tree of V1R proteins. The five functionally similar (blue) and one distant (red) receptor studied here are marked. The scale bar indicates the number of amino acid substitutions per site. Numbers indicate bootstrap values. (B) Pairwise distances among 204 V1R protein sequences (see supplementary materials, materials and methods). (C) For these deorphanized VRs, the top 20 nearest VRs, based on (B), are rank ordered. The deorphanized VRs are colored if shown in the table. (D) Two-dimensional representation of the distance matrix shown in (B). Each dot represents a single VR protein. coord, coordinate.

to cluster 1 but was not chosen for analysis by PhOTseq (Fig. 3, C and D). Because sequencing revealed control neurons that expressed only *Vmn1r85* (Fig. 2F), functional cluster 4 cells likely expressed *Vmn1r85*.

Only a few studies have examined the relationship between VR sequence and chemosensory function (5, 19) and never from the vantage point of complete knowledge of the VRs that underlie a set of nearest neighbors in terms of function. To investigate this relationship, we performed multiple sequence alignment (MSA) of vomeronasal type 1 receptor (V1R) protein sequences followed by phylogenetic tree analysis (Fig. 4A). Among V1Rs expressed in functionally similar types, VMN1R89, VMN1R86, and VMN1R85 were

close to each other in the sense of putative evolutionary distance. Two of the functionally similar cell types, VMN1R78 and VMN1R237, belonged to different branches. In this representation, VMN1R58, which was the chosen outlier in terms of function, was not notably more divergent from VMN1R89, VMN1R86, and VMN1R85 than the more functionally similar VMN1R78 and VMN1R237. Consequently, this tree structure representation did not suggest a particularly close correspondence between function and primary sequence.

However, when the pairwise distances acquired from the MSA analysis were examined carefully, we noticed that, despite their apparent evolutionary distance, VMN1R78

and VMN1R237 were consistently among the VR genes most closely related to VMN1R89, VMN1R86, and VMN1R85 (Fig. 4, B and C). Among 204 V1R sequences examined, the median neighbor rank among these VRs was 11, implying that these VRs are typically among the top-few-percent-closest matches to one another. We therefore considered the possibility that phylogeny, in attempting to construct plausible ancestry, inaccurately models relatedness between sequences that are not nearest neighbors. To more comprehensively evaluate and visualize all pairwise distances, we performed a classical multidimensional scaling analysis to project the high-dimensional sequence-based distance relationships into two dimensions (Fig. 4D). This resulted in a

different picture of the gene family, whose dominant feature was the presence of seven to nine apparent clusters. The functionally similar receptors, including VMN1R78 and VMN1R237, were near neighbors, whereas the functionally dissimilar receptor VMN1R58 was positioned in a distant cluster. We conclude that relationships among the primary sequences of these VRs are strongly correlated with their degree of functional similarity.

Our results demonstrate the utility of PhOTseq, an all-optical solution to the problem of cell type selection and labeling. It provides an opportunity to study the relationships between genes and physiological function even in extremely rare cell types that can be defined only through extensive functional characterization. PhOTseq will be applicable when individual neurons respond combinatorially to various conditions, including sensory stimulation, emotional status, and behavior. We further anticipate that in vivo ectopic expression will be widely applicable for characterizing VR-ligand pairings. In addition, our data show unexpected receptor coexpression, a specific and systematic departure from the “one neuron, one receptor rule” (21), whose potential roles will need to be evaluated in

many aspects of olfactory function. Lastly, our study provides biological insight on chemosensation by comprehensively mapping receptor-ligand pairings for chosen subsets of the vomeronasal sensory population and suggests that saturation analyses can reveal sequence-function coupling unanticipated by single-ligand studies (24, 25).

REFERENCES AND NOTES

1. H. Markram et al., *Nat. Rev. Neurosci.* **5**, 793–807 (2004).
2. M. Nassar et al., *Front. Neural Circuits* **9**, 20 (2015).
3. P. S. Xu, D. Lee, T. E. Holy, *Neuron* **91**, 878–892 (2016).
4. H. Okuno, *Neurosci. Res.* **69**, 175–186 (2011).
5. S. Haga-Yamanaka et al., *eLife* **3**, e03025 (2014).
6. J. Fuzik et al., *Nat. Biotechnol.* **34**, 175–183 (2016).
7. S. Qiu et al., *Front. Genet.* **3**, 124 (2012).
8. C. R. Cadwell et al., *Nat. Biotechnol.* **34**, 199–203 (2016).
9. B. F. Fosque et al., *Science* **347**, 755–760 (2015).
10. J. Akerboom et al., *J. Neurosci.* **32**, 13819–13840 (2012).
11. T.-W. Chen et al., *Nature* **499**, 295–300 (2013).
12. F. V. Subach et al., *Nat. Methods* **6**, 153–159 (2009).
13. M. D. Ryan, A. M. Q. King, G. P. Thomas, *J. Gen. Virol.* **72**, 2727–2732 (1991).
14. K. C. P. Wilson, G. Raisman, *Brain Res.* **185**, 103–113 (1980).
15. T. Bozza, P. Feinstein, C. Zheng, P. Mombaerts, *J. Neurosci.* **22**, 3033–3043 (2002).
16. P. Feinstein, T. Bozza, I. Rodriguez, A. Vassalli, P. Mombaerts, *Cell* **117**, 833–846 (2004).
17. E. A. Hallem, M. G. Ho, J. R. Carlson, *Cell* **117**, 965–979 (2004).
18. X. Grosmaître, A. Vassalli, P. Mombaerts, G. M. Shepherd, M. Ma, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1970–1975 (2006).
19. Y. Isogai et al., *Nature* **478**, 241–245 (2011).
20. W. M. Wong et al., *bioRxiv* 766592 [Preprint]. 12 September 2019. <https://doi.org/10.1101/766592>.
21. P. Mombaerts, *Nat. Rev. Neurosci.* **5**, 263–278 (2004).
22. J. Loconto et al., *Cell* **112**, 607–618 (2003).
23. B. Stein, M. T. Alonso, F. Zufall, T. Leinders-Zufall, P. Chamero, *PLOS ONE* **11**, e0156092 (2016).
24. Y. Jiang et al., *Nat. Neurosci.* **18**, 1446–1454 (2015).
25. B. von der Weid et al., *Nat. Neurosci.* **18**, 1455–1463 (2015).

ACKNOWLEDGMENTS

We thank R. Yu for sharing the TetO vector and OMP-IRES-tTA mouse, P. H. Taghert for the HEK cell line, H. Schoknecht for mouse care, and D. W. Kim for developing imaging software. We also thank T. Barnes, J. Chen, J. Corbo, X. Fu, C. Greer, A. Livi, R. Roberts, and M. Zhang for suggestions and comments. This work was supported by the Hope Center Viral Vectors Core and the Mouse Genetics Core at the Washington University School of Medicine in St. Louis. **Funding:** This work was funded by NIH/NIDCD R01 DC005964 and DC010381. **Author contributions:** M.K. performed in vivo imaging and PA. D.L. performed all other experiments and analysis. T.E.H. supervised the study and analysis. **Competing interests:** The authors have no competing interests. **Data and materials availability:** Accession numbers are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6471/1384/suppl/DC1
Materials and Methods
Figs. S1 to S13
References (26–52)
Movies S1 and S2
Data File S1

[View/request a protocol for this paper from Bio-protocol.](http://www.biorxiv.org/content/10.1101/000000)

25 April 2019; accepted 25 October 2019
10.1126/science.aax8055

Sensory coding mechanisms revealed by optical tagging of physiologically defined neuronal types

Donghoon Lee, Maiko Kume and Timothy E. Holy

Science **366** (6471), 1384-1389.
DOI: 10.1126/science.aax8055

Cell-type identification in neural circuits

In many cases, a single molecular marker is insufficient to define a specific cell type and may label a few, or a few hundred, physiologically distinguishable cell types. Lee *et al.* developed a high-throughput technique, called physiological optical tagging sequencing (PhOTseq), for identifying the expression profile of cells that exhibit a particular physiological profile (see the Perspective by Renninger). They used PhOTseq to identify genes encoding vomeronasal receptors in mice, which detect pheromones and subserve social communication.

Science, this issue p. 1384; see also p. 1311

ARTICLE TOOLS	http://science.sciencemag.org/content/366/6471/1384
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2019/12/11/366.6471.1384.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/366/6471/1311.full
REFERENCES	This article cites 46 articles, 9 of which you can access for free http://science.sciencemag.org/content/366/6471/1384#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2019 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works