

# Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex

Ryan Remedios<sup>1\*</sup>, Ann Kennedy<sup>1\*</sup>, Moriel Zelikowsky<sup>1</sup>, Benjamin F. Grewe<sup>2,3,†</sup>, Mark J. Schnitzer<sup>2,3,4</sup> & David J. Anderson<sup>1,4</sup>

**All animals possess a repertoire of innate (or instinctive<sup>1,2</sup>) behaviours, which can be performed without training. Whether such behaviours are mediated by anatomically distinct and/or genetically specified neural pathways remains unknown<sup>3–5</sup>. Here we report that neural representations within the mouse hypothalamus, that underlie innate social behaviours, are shaped by social experience. Oestrogen receptor 1-expressing (*Esr1*<sup>+</sup>) neurons in the ventrolateral subdivision of the ventromedial hypothalamus (VMHvl) control mating and fighting in rodents<sup>6–8</sup>. We used microendoscopy<sup>9</sup> to image *Esr1*<sup>+</sup> neuronal activity in the VMHvl of male mice engaged in these social behaviours. In sexually and socially experienced adult males, divergent and characteristic neural ensembles represented male versus female conspecifics. However, in inexperienced adult males, male and female intruders activated overlapping neuronal populations. Sex-specific neuronal ensembles gradually separated as the mice acquired social and sexual experience. In mice permitted to investigate but not to mount or attack conspecifics, ensemble divergence did not occur. However, 30 minutes of sexual experience with a female was sufficient to promote the separation of male and female ensembles and to induce an attack response 24 h later. These observations uncover an unexpected social experience-dependent component to the formation of hypothalamic neural assemblies controlling innate social behaviours. More generally, they reveal plasticity and dynamic coding in an evolutionarily ancient deep subcortical structure that is traditionally viewed as a ‘hard-wired’ system.**

We performed microendoscopic calcium imaging<sup>9,10</sup> (Inscopix, Inc.; Fig. 1a) of VMHvl *Esr1*<sup>+</sup> neurons expressing GCaMP6s<sup>11</sup> (Fig. 1b, c; see Methods) during resident–intruder assays<sup>12</sup> that were conducted in the home cage of the resident. On a given day, the implanted resident mouse engaged in five trials with (different) female intruders, shuffled with five trials with (different) male intruders (Fig. 1d). We observed no impairment in social behaviour in animals habituated to the miniature microscope (Fig. 1e, f). Behavioural data from synchronized videos were manually scored at 30 Hz. We recorded between 175 and 300 *Esr1*<sup>+</sup> cells per day per mouse ( $n = 7$ ; approximately 3,400 cells were imaged; see Methods).

In socially and sexually experienced residents, VMHvl *Esr1*<sup>+</sup> neural activity was low on baseline trials but increased by conspecifics of either sex (Fig. 1e–g and Extended Data Fig. 1a, b). Cellular response kinetics following introduction of an intruder varied continuously, from transient to persistent (Extended Data Fig. 1d–f). The number of cells activated ( $>2\sigma$  above baseline; Extended Data Fig. 1a) by conspecifics (approximately 20–30%) was significantly higher than by a toy mouse (less than 5%; Extended Data Fig. 1b). Repeated trials with different intruders of the same sex activated similar neural ensembles (Fig. 1g), reflected by a high Pearson’s correlation coefficient (PCC) between ensemble responses to the same sex (Fig. 1h). The population activity

vector during a social encounter, in the principal component space (Fig. 1k and Extended Data Fig. 2b, c), or in the partial least squares regression<sup>13</sup> space (Extended Data Fig. 2a), was remarkably similar across trials with different intruders of the same sex.

By contrast, males versus females activated different ensembles (Fig. 1g) and the PCC between sexes was close to zero (Fig. 1h), independent of inter-animal distance (Extended Data Fig. 3c). Male-versus female-preferring cells (defined as activated  $>2\sigma$  by one but not the other sex during social interactions) were spatially intermingled (Fig. 1i and Extended Data Fig. 1g), consistent with prior Fos-catFISH (cellular compartment analysis of temporal activity by fluorescent *in situ* hybridization of Fos) studies<sup>14</sup>. The number of male-preferring cells was around 50% higher than the number of female-preferring cells<sup>7,14</sup> (Fig. 1j and Extended Data Fig. 1b). Approximately 20% of cells activated by one sex were preferentially inhibited by the opposite sex (Fig. 1j and Extended Data Fig. 1c, g, h).

Intruder sex accounted for  $56 \pm 5.2\%$  of the observed variance (Extended Data Fig. 2a), indicating that population activity is dominated by the representation of conspecific sex. Reflecting this, a linear support vector machine (SVM) decoder could correctly classify an intruder’s sex identity with almost 100% accuracy, within 3 s after its introduction (Fig. 1l). SVM decoders trained on data from residents separated from intruders by a mesh barrier also performed accurately, suggesting that sex-specific ensembles represent sensory cues, not sex-specific behaviours (see below).

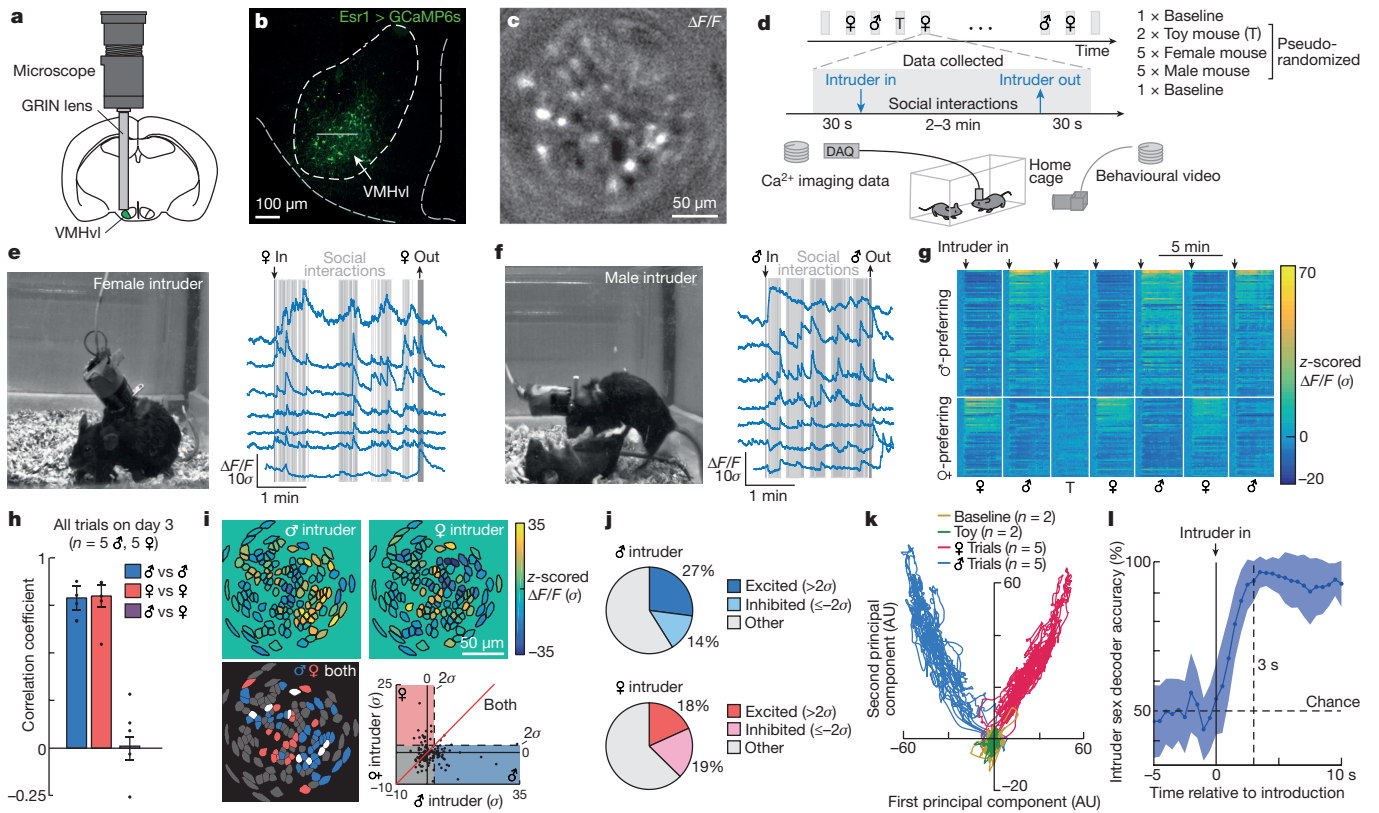
Optogenetic manipulation of *Esr1*<sup>+</sup> neurons in the VMHvl evoked specific social behaviours, such as mounting or attack<sup>7</sup>, but behaviour-triggered average analysis of calcium transients revealed few time-locked responses during such actions<sup>14–16</sup>. We therefore investigated whether VMHvl ensemble responses could at least predict epochs when these behaviours occurred (Fig. 2a). Linear SVM decoders showed significantly higher than chance performance in predicting episodes of attack, sniffing or mounting (Fig. 2b, c), with errors typically affecting their precise timing. Such errors may reflect inconsistencies in the manual scoring of behaviour onset/offset or biological factors.

Additionally, we examined whether VMHvl neurons might be probabilistically tuned to particular behaviours, by calculating the ‘choice probability’<sup>17</sup> for each cell, defined as the chances of correctly predicting which of two alternative behaviours is occurring based on that neuron’s activity (Extended Data Fig. 4a). Approximately half of the cells (52–53%) were tuned for periods when social behaviour occurred (Fig. 2h, dark grey). Within that category, there were cells tuned to sniff versus attack for males or to sniff versus mount for females (Fig. 2d–h and Extended Data Fig. 4).

Surprisingly, the behavioural tuning of many cells was independent of their sex-preference (Fig. 1g): for example, ‘sniff male’-tuned units were equally represented among male- versus female-preferring cells (Fig. 2i). Therefore, individual *Esr1*<sup>+</sup> neurons can participate in ensembles

<sup>1</sup>Division of Biology and Biological Engineering 156-29, Tianqiao and Chrissy Chen Institute for Neuroscience, California Institute of Technology, Pasadena, California 91125, USA. <sup>2</sup>James H. Clark Center for Biomedical Engineering & Sciences, Stanford University, Stanford, California 94305, USA. <sup>3</sup>CNC Program, Stanford University, Stanford, California 94305, USA. <sup>4</sup>Howard Hughes Medical Institute, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. <sup>†</sup>Present address: Institute of Neuroinformatics, ETH Zurich, CH-8057 Zurich, Switzerland.

\*These authors contributed equally to this work.



**Figure 1 | VMHvl *Esr1*<sup>+</sup> neurons represent intruder sex.** **a**, Schematic of preparation. Redrawn from Allen Mouse Brain Atlas, version 1 (2008). **b**, Approximate imaging plane (dashed line). **c**, Sample frame showing active neurons. **d**, Experimental design. DAQ, data acquisition box. **e**, **f**, Sample video frames and  $\Delta F/F$  traces. **g**, Raster of responses of male-prefering vs. female-prefering cells ( $n = 135$ ) ranked by response strength. **h**, Average PCC between cell responses to males and females, each datapoint is a single mouse. **i**, Spatial maps of averaged neuronal

responses (top); intruder sex-prefering cells ( $>2\sigma$  above baseline; bottom left); scatterplot of response intensities (bottom right). **j**, Fraction of cells activated or inhibited ( $\pm 2\sigma$  from baseline) by conspecifics for 10 trials ( $n = 5$  mice; 2,379 cells imaged). **k**, Population activity vectors in principal component space. AU, arbitrary units. **l**, Decoder performance during the introduction of an intruder ( $n = 5$  mice that mounted and fought, 10 trials each). **e–g**, **i**, **k**, Panels are from one example mouse. Data are mean  $\pm$  s.e.m in all figures and extended data figures.

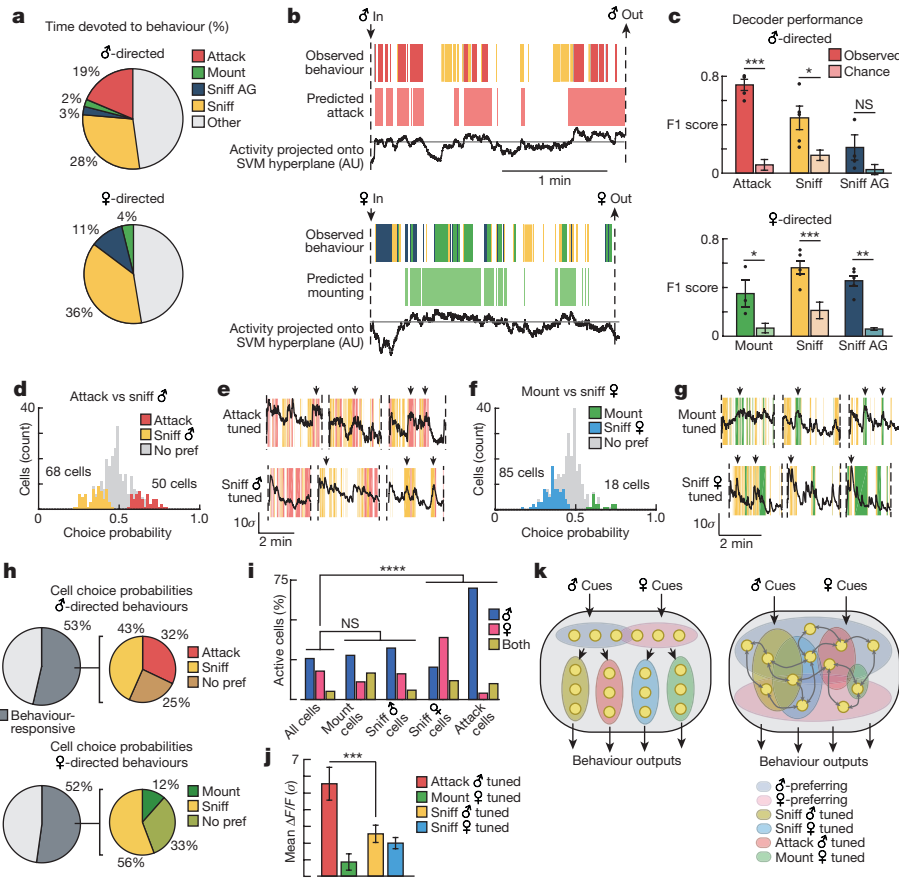
representing a particular sex, as well as in ensembles for specific appetitive or consummatory behaviours<sup>18</sup> (Fig. 2k (right) and Extended Data Fig. 4c). An exception was that ‘attack male’- and ‘sniff female’-tuned cells were enriched for male- versus female-prefering neurons, respectively (Fig. 2i). Attack-tuned cells were more strongly activated than other behaviour-tuned neurons (Fig. 2j), suggesting a higher threshold for this action and consistent with optogenetic results<sup>7,15</sup>.

The above-mentioned data were obtained from the 3rd consecutive day of social encounter testing. Before day 1, mice were naive to non-littermate conspecifics, sexually inexperienced and were individually housed following surgery for at least four weeks (see Methods). During initial trials on day 1, we observed little mounting or attack (Fig. 3j (day 1) and Extended Data Fig. 5a). Notably, imaging revealed many cells that responded to both males and females (Fig. 3a–c (day 1)). Over the next two days, more male- or female-prefering cells were observed, while the overall percentage of active cells remained fairly constant (Fig. 3a–c (day 2 and 3)). Accordingly, the PCC between the male- and female-response population vectors decreased to nearly zero by day 3 (Fig. 3e (gold curve), f). Similar changes were observed in decoder performance or the Mahalanobis distance ratio<sup>19</sup> between same- versus opposite-sex trials (Fig. 3g–i). Sex-specific ensembles were stable for months in animals isolated following these trials. Interestingly, two out of seven animals imaged failed to exhibit any separation of representations, and showed neither mounting nor attack (for example, Extended Data Fig. 6c, d). These ‘anomalous’ mice were observed at similar frequencies (around 15–20%) among unoperated control animals (data not shown). The origin of these individual differences among inbred mice is unclear.

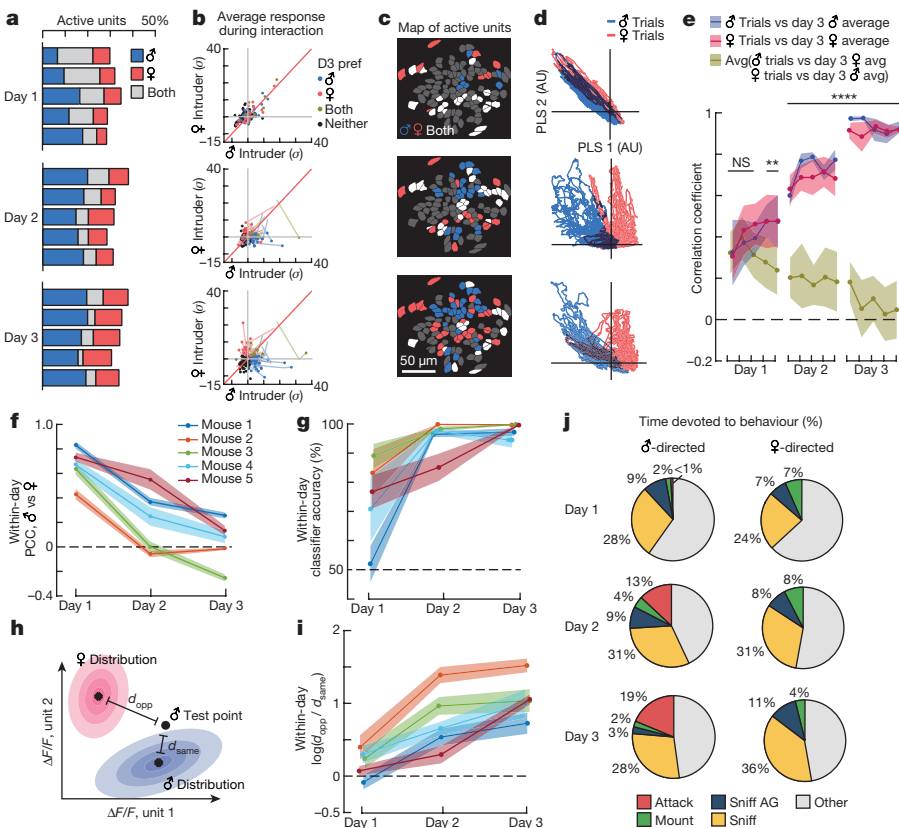
To track the origins of the sex-specific cells observed on day 3, we registered 455 cells across three days of imaging (Extended Data Fig. 7; around 40% of all cells could be registered). Twenty-five per cent of day 3 sex-prefering cells were bi-responsive on day 1, while almost half derived from cells that were initially unresponsive ( $<2\sigma$  above baseline) to either sex (Extended Data Fig. 8). A small number (10%) arose from switching sex-specificity. Conversely, 42% of cells that were active on day 1 were no longer active on day 3. These data reveal complex and dynamical changes in the response properties of VMHvl *Esr1*<sup>+</sup> neurons during the acquisition of social experience.

The progressive separation of male versus female representations across the three days of testing was accompanied by an increase in mounting and attack (Fig. 3j and Extended Data Fig. 5a). To quantify this correlation, we computed trial-by-trial changes in the male versus female PCC, and plotted the PCC for each mouse against the cumulative minutes of different social behaviours (Fig. 4a–c and Extended Data Fig. 9). The strongest correlation was with mounting plus agonential investigation (Fig. 4b) and the weakest with sniffing (Fig. 4c).

The weak correlation with sniffing suggested that olfactory cues alone might not suffice to promote the separation of male versus female representations. To test this, we suspended intruders by the tail (allowing sniffing, but preventing mounting or attack) throughout each 2-min trial of a standard three-day schedule (Fig. 4d). Under these conditions, the male versus female PCC remained high at day 3 in all three mice (Fig. 4e (teal bars)). However, when two of these mice (a third became ill and was euthanized) were allowed free interactions with females and males over two additional days, sex-specific representations became well-separated, and mounting and attacks were observed (Fig. 4f, g).

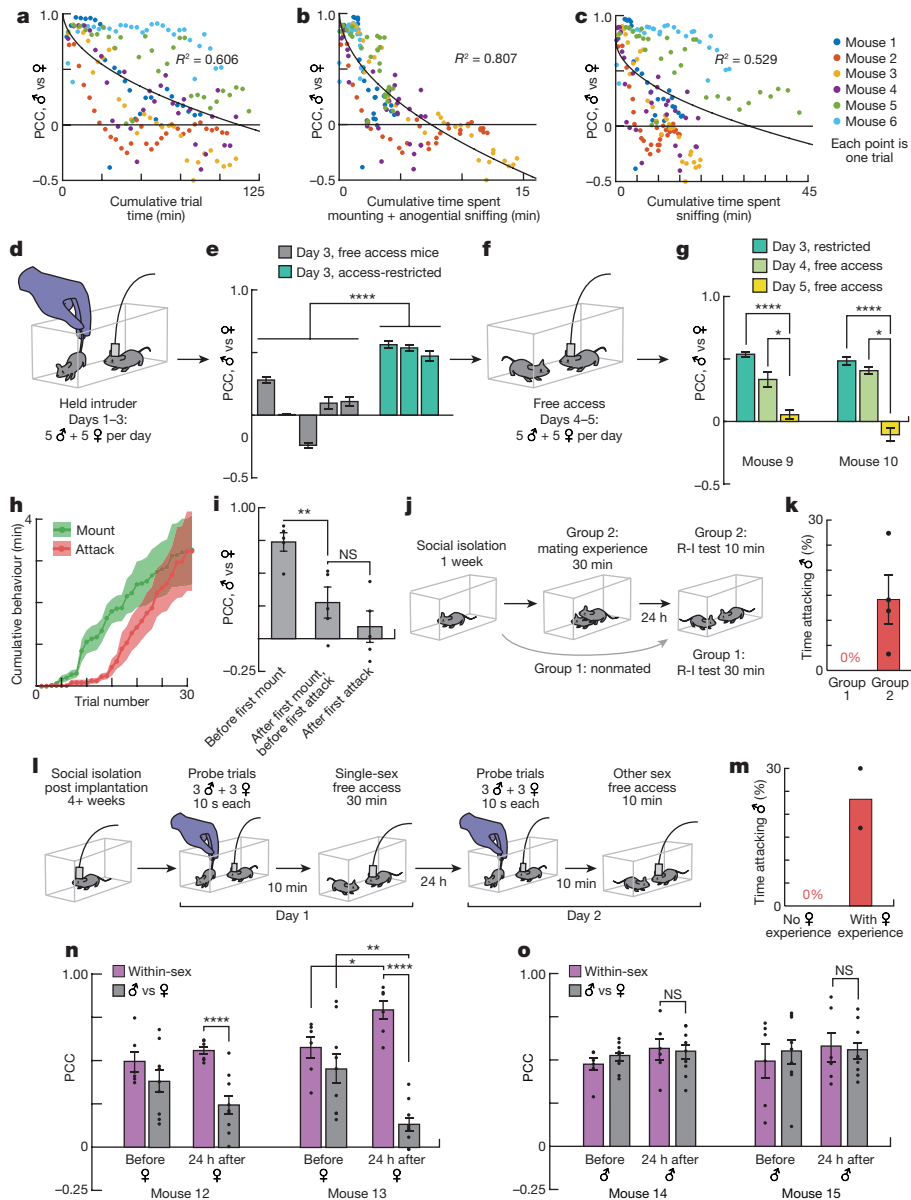


**Figure 2 | Population activity predicts social behaviour.** **a**, Behavioural responses to males and females ( $n = 5$  mice). AG, anogenital. **b**, Example decoder accuracy. **c**, Decoder performance ( $n = 5$  mice, 10 trials each; see Methods for significance testing). **d, e**, Choice probabilities of 291 neurons (**d**) and traces from two example neurons (**e**) during three separate male interactions. **f, g**, Choice probabilities for the same 291 neurons during female interactions (**f**) and two example neurons (**g**). **h**, Proportion of cells exhibiting significant ( $>2\sigma$ ) choice probabilities (that is, behavioural tuning). **i**, Proportion of sex-preferring cells within each category of behaviour-tuned cells ( $\chi^2$  test, 3 degrees of freedom). **j**, Response strength of behaviour-tuned populations, during their preferred behaviour. (corrected for multiple comparisons). **k**, Two models of ensemble representations of behaviour among  $Esr1^+$  VMHvl neurons. The data argue against the model on the left. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < .001$ ; NS, not significant by two-sided  $t$ -tests unless specified otherwise for all figures and extended data figures. Specific  $P$  values for Figs 2–4 are provided in Supplementary Table 1.



**Figure 3 | Sex-specific ensembles emerge with experience.** **a–d**, Ensemble representations of responsive cells across pairs of consecutive trials ( $n = 455$  cells from 5 mice). **a**, Fraction of responsive cells across pairs of consecutive trials. **b**, Response strength ( $\Delta F/F$ ) for 135 cells identified across three days. Day 3 preference (D3 pref.) indicates the response preference of each unit on day 3 towards males (blue), females (red) or both (grey). Grey lines in Day 2 and Day 3 plots indicate shifts in preference from Day 1 or Day 2 position, respectively. **c**, Example spatial maps of intruder sex preference. **d**, Population vectors in a partial least squares (PLS) regression space. **e**, PCC between the trial-averaged ensemble responses on day 3 and each trial (points) for five mice. Significant ensemble separation starts on day 1 (post hoc Bonferroni test to correct for multiple comparisons). **f**, Average PCC for each mouse. **g**, SVM decoder accuracy across days. **h**, Schematic description of Mahalanobis distance ( $d$ ) calculation. opp, between opposite sexes; same, within same sex. **i**, Changes in the log of the Mahalanobis distance ratio. **j**, Percentage of time spent engaging in each social behaviour ( $n = 5$  mice).





**Figure 4 | Social experience promotes ensemble separation.** **a–c**, PCC for the  $n$ th trial versus cumulative social experience (**a**), cumulative mounting and anogenital sniffing (**b**) or cumulative body/face-directed sniffing (**c**).  $R^2$  values are for fit curve  $y = a\sqrt{x} + b$  (black). **d**, Illustration of the restricted-access experiment. **e**, Ensemble separation for five mice with free access (grey) and three mice with restricted access (teal) were significantly different (one-way ANOVA). **f, g**, PCC decreased after two restricted-access mice received free access (day 4 and day 5;

one-way ANOVA); day 3 bars (teal) are duplicated from **e** to facilitate comparison. **h**, Cumulative time spent engaging in each behaviour ( $n = 5$ ). **i**, Ensembles separate after mounting commences. **j**, Priming experiment. **k**, Males primed with females (group 2,  $n = 4$ ) attacked conspecific males; unprimed males (group 1,  $n = 8$ ) did not. **l**, Priming experiment adapted for imaging. **m**, Males primed with females fought. **n, o**, Mice primed with females ( $n = 2$ ) showed ensemble separation 24 h later (**n**); mice primed with males ( $n = 2$ ) did not (**o**). R-I; resident–intruder assay.

Similar results were obtained using a wire mesh to block social interactions with intruders (Extended Data Fig. 10). These data suggested that exposure to conspecific odours is insufficient to promote sex-specific ensemble separation, implying a requirement for social behaviours.

To identify these behaviours, we next investigated the temporal relationship between mounting, attack and ensemble separation. Mounting appeared before attack, in both operated (Fig. 4h) and unoperated mice (Extended Data Fig. 5b), consistent with previous reports<sup>21</sup>. Ensemble separation increased significantly after the first mount, but before the first attack (Fig. 4i). These observations suggested that social experience with females might suffice to promote both attack<sup>22</sup> and sex-specific ensemble separation.

To test the former possibility, socially inexperienced, unoperated males were ‘primed’ in a single 30-min unrestricted interaction with

females, and then tested 24 h later in a single, 10-min resident–intruder assay with a male (Fig. 4j). Notably, this priming with females (during which mounting was observed) promoted aggressiveness in 100% of males (Fig. 4k (group 2)). By contrast, control males with no priming showed no attack towards an intruder male, even after 30 min (Fig. 4k (group 1)).

To determine whether sexual experience can also promote the appearance of sex-specific neural ensembles, we repeated this priming experiment while imaging *Esr1*<sup>+</sup> neurons in naive mice (Fig. 4l). Probe trials with intruders that were suspended by their tail (three males, three females, 10 s each) were performed 10 min before, and 24 h after, the 30-min priming interaction with females. Notably, the PCC between male versus female representations was significantly reduced, in probe trials performed after (but not before) female priming, compared to the



same-sex PCC (Fig. 4n). These primed mice also proceeded to attack males during the resident–intruder test on day 2 (Fig. 4m). By contrast, the male versus female PCC was not reduced after 24 h in mice primed with a male on day 1 (Fig. 4o). Therefore, a single 30-min priming experience with a female sufficed to promote both sex-specific ensemble separation and aggression. Whether ensemble separation itself is required for attack remains unclear.

Here we report the imaging analysis of neuronal activity in a deep hypothalamic structure during social behaviour in freely moving animals. Our observations in experienced mice revealed distinct and reproducible VMHvl *Esr1*<sup>+</sup> subpopulations activated by males versus females, suggestive of labelled lines<sup>14</sup>. However, in socially isolated and sexually inexperienced animals these divergent representations were not present, but emerged with sexual experience. How can these findings be reconciled with the ‘innate’ nature of mating and aggression<sup>1,2</sup>? One possibility is that circuits downstream of the VMH are hard-wired, but that sexual experience is necessary to direct these behaviours towards appropriate social stimuli<sup>23</sup>. We cannot exclude that plasticity in the VMHvl is ‘inherited’ from upstream regions<sup>24–27</sup>. Nevertheless, our findings suggest that some computational features characteristic of the hippocampus and cortex—including dynamic coding, mixed selectivity and experience-dependent plasticity<sup>28–30</sup>—also operate in deep subcortical circuits. More generally, our findings raise the possibility that the neural representations underlying other instinctive behaviours may also incorporate learned features.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 3 January; accepted 31 July 2017.**

1. Tinbergen, N. *The Study of Instinct* (Clarendon Press, 1951).
2. Lorenz, K. *On Aggression* (Harcourt, Brace & World, 1966).
3. Root, C. M., Denny, C. A., Hen, R. & Axel, R. The participation of cortical amygdala in innate, odour-driven behaviour. *Nature* **515**, 269–273 (2014).
4. Newman, S. W. The medial extended amygdala in male reproductive behavior. A node in the mammalian social behavior network. *Ann. NY Acad. Sci.* **877**, 242–257 (1999).
5. Veening, J. G. *et al.* Do similar neural systems subserve aggressive and sexual behaviour in male rats? Insights from c-Fos and pharmacological studies. *Eur. J. Pharmacol.* **526**, 226–239 (2005).
6. Yang, C. F. *et al.* Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* **153**, 896–909 (2013).
7. Lee, H. *et al.* Scalable control of mounting and attack by *Esr1*<sup>+</sup> neurons in the ventromedial hypothalamus. *Nature* **509**, 627–632 (2014).
8. Sano, K., Tsuda, M. C., Musatov, S., Sakamoto, T. & Ogawa, S. Differential effects of site-specific knockdown of estrogen receptor  $\alpha$  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 (2013).
9. Ghosh, K. K. *et al.* Miniaturized integration of a fluorescence microscope. *Nat. Methods* **8**, 871–878 (2011).
10. Ziv, Y. *et al.* Long-term dynamics of CA1 hippocampal place codes. *Nat. Neurosci.* **16**, 264–266 (2013).
11. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
12. Thurmond, J. B. Technique for producing and measuring territorial aggression using laboratory mice. *Physiol. Behav.* **14**, 879–881 (1975).
13. Geladi, P. & Kowalski, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* **185**, 1–17 (1986).
14. Lin, D. *et al.* Functional identification of an aggression locus in the mouse hypothalamus. *Nature* **470**, 221–226 (2011).

15. Falkner, A. L., Dollar, P., Perona, P., Anderson, D. J. & Lin, D. Decoding ventromedial hypothalamic neural activity during male mouse aggression. *J. Neurosci.* **34**, 5971–5984 (2014).
16. Falkner, A. L., Grosenick, L., Davidson, T. J., Deisseroth, K. & Lin, D. Hypothalamic control of male aggression-seeking behavior. *Nat. Neurosci.* **19**, 596–604 (2016).
17. Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
18. Jennings, J. H. *et al.* Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. *Cell* **160**, 516–527 (2015).
19. Grewe, B. F. *et al.* Neural ensemble dynamics underlying a long-term associative memory. *Nature* **543**, 670–675 (2017).
20. Efron, B., Hastie, T., Johnstone, I. & Tibshirani, R. Least angle regression. *Ann. Statist.* **32**, 407–499 (2004).
21. Ogawa, S. *et al.* Survival of reproductive behaviors in estrogen receptor  $\beta$  gene-deficient ( $\beta$ ERKO) male and female mice. *Proc. Natl Acad. Sci. USA* **96**, 12887–12892 (1999).
22. Gobrogge, K. L., Liu, Y., Jia, X. & Wang, Z. Anterior hypothalamic neural activation and neurochemical associations with aggression in pair-bonded male prairie voles. *J. Comp. Neurol.* **502**, 1109–1122 (2007).
23. Redondo, R. L. *et al.* Bidirectional switch of the valence associated with a hippocampal contextual memory engram. *Nature* **513**, 426–430 (2014).
24. Xu, P. S., Lee, D. & Holy, T. E. Experience-dependent plasticity drives individual differences in pheromone-sensing neurons. *Neuron* **91**, 878–892 (2016).
25. Stowers, L. & Liberles, S. D. State-dependent responses to sex pheromones in mouse. *Curr. Opin. Neurobiol.* **38**, 74–79 (2016).
26. Gur, R., Tendler, A. & Wagner, S. Long-term social recognition memory is mediated by oxytocin-dependent synaptic plasticity in the medial amygdala. *Biol. Psychiatry* **76**, 377–386 (2014).
27. Cooke, B. M. Steroid-dependent plasticity in the medial amygdala. *Neuroscience* **138**, 997–1005 (2006).
28. Knierim, J. J. & Zhang, K. Attractor dynamics of spatially correlated neural activity in the limbic system. *Annu. Rev. Neurosci.* **35**, 267–285 (2012).
29. Burak, Y. Spatial coding and attractor dynamics of grid cells in the entorhinal cortex. *Curr. Opin. Neurobiol.* **25**, 169–175 (2014).
30. Moser, E. I., Kropff, E. & Moser, M.-B. Place cells, grid cells, and the brain’s spatial representation system. *Annu. Rev. Neurosci.* **31**, 69–89 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank X. Wang, J. S. Chang and R. Robertson for technical help, H. Lee and P. Kunwar for experimental advice, D. Senyuz for testing behaviour in wild-type mice, D.-W. Kim for pilot experiments, M. McCardle and C. Chiu for genotyping, J. Costanza for mouse colony management, G. Stuber for advice on GCaMP6s expression, Inscopix Inc. for technical support, P. Perona for mouse tracking software, L. Abbott for comments on the manuscript, R. Axel, D. Y. Tsao and M. Meister for critical feedback, X. Da and C. Chiu for laboratory management and G. Mancuso for Administrative Assistance. D.J.A. and M.J.S. are Investigators of the Howard Hughes Medical Institute and Paul G. Allen Distinguished Investigators. This work was supported in part by NIH grant no. R01MH070053, and grants from the Gordon Moore Foundation, Ellison Medical Research Foundation, Simons Foundation and Guggenheim Foundation to D.J.A. A.K. is a fellow of the Helen Hay Whitney Foundation, M.Z. is a recipient of fellowships from the NSF and L’Oréal USA Women in Science.

**Author Contributions** R.R. designed and performed all imaging experiments, processed the data, contributed to analysis and co-wrote the manuscript; A.K. performed computational analysis, prepared figures and co-wrote the manuscript; M.Z. designed and performed behavioural experiments; M.J.S. and B.F.G. provided training for R.R., guidance on experimental design and data analysis, and critical feedback; D.J.A. supervised the project and co-wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.J.A. ([wuwei@caltech.edu](mailto:wuwei@caltech.edu)).

## METHODS

**Mice.** All experimental procedures involving the use of live animals or their tissues were carried out in accordance with the NIH guidelines and approved by the Institutional Animal Care and Use Committee and the Institutional Biosafety Committee at the California Institute of Technology (Caltech). *Esr1<sup>Cre/+</sup>* knock-in mice<sup>7</sup> backcrossed into the C57BL/6N background (>N10) were bred at Caltech. The *Esr1<sup>Cre/+</sup>* knock-in line is available from the Jackson Laboratory (Stock no. 017911). Heterozygous *Esr1<sup>Cre/+</sup>* mice were used for all imaging studies and were genotyped by PCR analysis of tail DNA. Mice used as residents (see Methods: resident–intruder assay) were individually housed. All wild-type mice used as intruders in resident–intruder assays and for behavioural experiments without imaging (unoperated controls) were of the C57BL/6N strain (Charles River Laboratories) purchased at 12–14 weeks of age, and group-housed at Caltech before use. Female mice were purchased ovariectomized and received injections of 17 $\beta$ -oestradiol benzoate (Sigma-Aldrich) in sesame oil (10  $\mu$ g at 48 h and 5  $\mu$ g at 24 h) and of progesterone (50  $\mu$ g at 4–6 h) before behavioural testing<sup>31</sup>. All mice were housed in ventilated micro-isolator cages in a temperature-controlled environment (median temperature 23 °C), under a reversed 12 h dark–light cycle, and had *ad libitum* access to food and water. Mouse cages were changed weekly on a fixed day on which experiments were not performed.

**Surgery.** Adult heterozygous *Esr1<sup>Cre/+</sup>* males were single-housed for at least five days before undergoing surgical procedures and were operated on at 12–14 weeks of age. Before this, mice were maintained with male littermates, following weaning at three weeks, that is, for 9–11 weeks. Mice were anaesthetized using isoflurane (2.5% induction, 1.2–1.5% maintenance, in 95% oxygen) and placed in a stereotaxic frame (David Kopf Instruments). Body temperature was maintained using a heating pad. An incision was made to expose the skull for stereotaxic alignment using the inferior cerebral vein and the Bregma as vertical references. We based the coordinates for the craniotomy and stereotaxic injection of VMHvl on an anatomical magnetic resonance atlas of the mouse brain (AP: –4.68 mm; ML:  $\pm$ 0.78 mm; DV: –5.80 mm), as previously described<sup>7</sup>. Virus injection and GRIN endoscope lens implantation were performed during a single surgery. Virus suspension was injected using a pulled-glass capillary at a slow rate of 8–10 nl min<sup>-1</sup>, 100 nl per injection site (Nanojector II, Drummond Scientific; Micro4 controller, World Precision Instruments). The glass capillary was withdrawn 15–20 min after the cessation of injection and a custom-built tube of hypodermic grade stainless steel was lowered to the VMHvl target coordinates to create space that would accommodate the endoscope lens. The tube was removed after 20–30 min and immediately replaced by a GRIN endoscope lens (Inscopix, 0.5 or 0.6 mm diameter, minimum 6 mm length). The lens was cemented to the skull using a small drop of dental cement (Metabond, Parkell), and subsequently curing thin layers of dental cement were applied to cover the entire exposed skull. Next, multiple layers of dental acrylic (Coltene, Whaledent) were applied into which a head-fixation bar was allowed to set. Once the dental acrylic had cured, the mice were taken off isoflurane and allowed to recover in clean, autoclaved cages to minimize chance of infection. Bupivacaine and ketoprofen were administered as local and systemic analgesics, respectively, during surgery and mice were provided motril and ibuprofen in their drinking water after surgery for a week. Mice were maintained in isolation until recording experiments were performed (typically more than 4 weeks; see below).

**Virus.** *Esr1<sup>+</sup>* neurons expressing Cre-recombinase in the VMHvl were transduced with an adeno-associated virus (AAV) expressing GCaMP6s under the human Synapsin I promoter (AAV1.Syn.GCaMP6s; Penn Vector Core). A series of titration experiments was performed to determine the highest dilution at which bright cytoplasmic, non-nuclear GCaMP6s localization could be observed in slices of fixed tissue from injected animals<sup>32,33</sup> before performing injections for *in vivo* imaging applications.

**Microscope alignment.** Mice were initially checked for epifluorescence signals four weeks after virus injection and endoscope lens implantation. Mice were anaesthetized using isoflurane, mounted in a stereotaxic frame and the head was levelled to the horizontal plane. A head-mounted miniaturized microscope (nVista, Inscopix) was then lowered over the implanted lens until GCaMP6s-expressing fluorescent neurons were in focus uniformly across the imaging plane. If fluorescent neurons were not observed, the mouse was returned to the cage and tested again on a weekly basis. If GCaMP6s-expressing fluorescent neurons were detected, the microscope was aligned and a permanent baseplate attached to the head according to published protocols<sup>33</sup>. A baseplate cover (BPC-2, Inscopix) was used to prevent bedding dust from the home cage from settling on the lens. After 48 h, the baseplate cover was replaced by a weight-matched dummy microscope (DMS-2, Inscopix), to acclimatize the mouse to the weight of the microscope for at least three days before social behavioural testing.

**Selection of animals for imaging experiments.** Imaging data were collected from a total of 25 *Esr1<sup>Cre/+</sup>* mice (out of 122 injected/implanted animals generated in multiple cohorts). Sixteen out of 25 mice were used for experiments reported here. Mice were selected for combined imaging and behaviour experiments based on the following criteria: (1) they had individually discernable neurons across at least half the imaging field; (2) the imaging plane was clear and not occluded by blood or debris; and (3) the image registration could be applied successfully to correct for motion artefacts (all except one mouse). The most frequent failure modes were mistargeting of the viral injection and/or GRIN lens implantation, inadequate levels of GCaMP6s expression and failure to obtain a clear imaging plane.

**Determination of sample size.** For imaging experiments, between three and seven mice were imaged per experimental paradigm; the number of mice was dependent on the number of usable mice and the number of animals initially implanted and injected. Selection of mice was determined by screening with the miniscope before experimentation. Selection was not based on any pre-specified effect. Data presented in Figs 1–3 is from five out of seven mice imaged that eventually showed mounting and attack behaviour by the third day of imaging. The two ‘anomalous’ mice that showed neither consummatory social behaviour (mating, fighting), nor a separation of ensemble representations, were treated separately; data from one of these mice (mouse 6) is shown in Fig. 4a–c and in Extended Data Fig. 6.

The sample size for the number of cells was 175–300 cells per mouse, and was affected by the volume of viral GCaMP6s injected, the efficiency of Cre-mediated recombination and level of expression. During the post-processing of the imaging data (after identification of individual cells by principal component and independent component analyses (PCA–ICA)), 1–5% of the identified components were discarded as artefacts and the remaining 95–99% were treated as cells and used for further analysis.

Sample sizes for behavioural experiments using wild-type unoperated mice were determined by the current standard used for mice in behavioural neuroscience experiments, based on the minimal number of mice required to detect significance with a  $\alpha$  rate set to 0.05 in a standard-powered experiment.

**Randomization.** Animals were assigned randomly to experimental and control groups.

**Blinding.** Social behaviours were manually annotated blind to the sex of the intruder mouse by a trained human annotator. Behavioral assays, imaging and computational analysis were not performed blinded.

**Resident–intruder assay.** Before resident–intruder testing, implanted mice were individually housed for at least four weeks, to allow recovery from surgery and adequate levels of GCaMP6s expression. The lens-implanted *Esr1<sup>Cre/+</sup>* mouse was always the home cage resident. Data were collected from each resident over three consecutive days of testing, beginning when the resident was still naive to adult non-littermate conspecifics (that is, housed in social isolation beginning five days before surgery). The results in Figs 1, 2 are derived from resident animals on their third day of testing, by which time most of the mice expressed normal sexual and aggressive behaviours. Data for each day were collected as a sequence of 14 individual trials, comprising two baseline trials (that is, trials performed in the absence of an intruder or object), two trials with a toy mouse (OurPets Play-N-Squeak) and five trials each with a male or female intruder. The sequence of trials with an intruder or object was interleaved; the sex of the first intruder was randomized across animals. Each trial with an object or intruder lasted 2–3 min, to minimize photobleaching of GCaMP6s, and was preceded and followed by 30 s of imaging before and after the intruder was introduced, respectively. Intruders were manually introduced into the resident’s home cage held by the tail, during which the residents were allowed to anogenitally investigate the intruders for 1–5 s. Care was taken to remove the intruder or toy from the arena only when the resident was at a distance to prevent artificially truncating any naturally occurring social interaction. The inter-trial interval was 10 min. A different intruder mouse was used for each trial with a male or female intruder. One cohort of seven implanted mice were used for imaging in these three-day assays; of these seven mice, two did not show mating or aggressive behaviour or ensemble separation and were omitted from the analyses performed in Figs 1–3 (see main text and Extended Data Fig. 6). Another cohort of 12 unoperated wild-type mice was used for behavioural analysis only (Fig. 4j, k). A third cohort of three implanted mice was used for the restricted-access assay (Fig. 4d–g); a fourth cohort of six implanted mice was used for imaging in the female priming assay in Fig. 4l–o; a fifth cohort of two implanted mice was used for imaging in the mesh container version of the restricted-access assay (Extended Data Fig. 10).

**Restricted-access assay: held-intruder.** For the cohort of implanted mice ( $n = 3$ ), the experimental design was identical to the resident–intruder assay; however, to prevent free social interaction, intruder mice were held by the experimenter for the duration of each trial. Interactions with five males, five females, two toy and two baseline trials were recorded for each resident each day, for three consecutive

days of testing. The same implanted mice were subsequently given two additional consecutive days of testing in the original (free access) resident–intruder assay (see ‘Resident–intruder assay’). One of the three implanted mice became ill and had to be euthanized after the restricted-access portion of the experiment.

**Restricted-access assay: mesh barrier.** For a second group of implanted mice ( $n = 2$ ), the experimental design was identical to the resident–intruder assay, however intruders were placed inside a wire-mesh barrier (an inverted pencil cup) when introduced into the resident’s home cage. Interactions with five males, five females, and two toy and two baseline trials were recorded for each resident each day, for three consecutive days of testing. The mice were subsequently tested for two consecutive days of social interactions as in the primary resident–intruder assay (days 4, 5), followed by an additional day of testing with intruders behind the mesh barrier (day 6).

**Priming assay (Fig. 4l–o).** Naive, implanted mice were allowed to freely interact in a single session with either a female or a male for 30 min on day 1 (priming), and then tested with a conspecific of the opposite sex on day 2 in a 10 min free-interaction trial. Before and after the priming on day 1, probe trials were performed with (different) probe animals of each sex ( $n = 3$  males,  $n = 3$  females), to test for ensemble separation before and after priming. During these probe trials, a male or female probe animal was suspended by the tail in the cage for 10 s, during which time sniffing (but no attack or mounting) by the implanted resident occurred. Three implanted mice were primed with females on day 1 and three mice were primed with males on day 1. Of the three mice primed with females, one mouse failed to mate with the female during priming, or to show aggression towards a male on day 2, and was omitted from imaging analysis, because the purpose of the experiment was to investigate whether ensemble separation correlated with the appearance of aggression following mating with females. (This mouse may have been one of the ‘anomalous’ mice we describe, which exhibit neither mating nor fighting behaviour.) Of the three mice primed with males on day 1, one that already showed a sex-specific ensemble separation on the first probe trial (day 1), that is, before priming, was omitted from the analysis, as the purpose of the experiment was to investigate whether priming with a male was sufficient to cause sex-specific ensemble separation in mice where such separation had not yet occurred.

An additional 12 wild-type unoperated mice that had been socially isolated for two weeks were tested in this priming assay only to quantify behaviour (Fig. 4j, k). Four mice were presented with a female on day 1, and eight mice only experienced a single, 30-min encounter with a male on day 2.

**Neuronal and behavioural data acquisition.** Mice were temporarily head-fixed before imaging sessions and the head-mounted dummy was replaced by the miniaturized microscope. Anaesthesia was not used at this step, because it strongly suppressed hypothalamic neuronal activity. The mice were placed in their home cages situated within the behavioural arena with head-mounted microscopes attached, and allowed to habituate 15–30 min before data acquisition. The behavioural arena was custom-built (all parts from Thorlabs) with one top-view and one side-view camera (Flea3, Point Grey) for acquiring behavioural data, which were saved as AVI files on a Windows computer. A TTL pulse from the Sync port of the data acquisition box (DAQ, Inscopix) of the microscope was used to trigger individual frames through the GPIO pins on the camera. This allowed a 1:1 correspondence between behavioural frames and neuronal imaging frames, which was necessary for synchronizing neuronal activity to behaviour. Shortly before data acquisition, the imaging parameters were configured and saved using nVista (Inscopix) control software. All but two mice used in this study were imaged at 30 Hz with 33.2 ms exposure per frame (the other two were imaged at 15 Hz with 66.2 ms exposure per frame, as they exhibited lower levels of GCaMP6s fluorescence). Microscope LED illumination was 11–14% of the maximum allowed, at  $2\text{--}2.75 \times$  gain. The field of view was cropped to the region encompassing the fluorescent neurons to maximize uninterrupted data acquisition at 30 Hz. All imaging parameters were saved and were the same for every trial for each mouse. Data were saved as stacks of frames in TIFF or RAW format.

**Position tracking.** Instantaneous positions and poses (estimated by a fit ellipse) of both mice was tracked from the top-view video, using custom software written in MATLAB<sup>34</sup>. The distance between animals was defined as the distance between the centroids of the fit ellipses.

**Data processing.** The behavioural videos were loaded into MATLAB using a custom-scripted behavioural annotator<sup>7,14</sup> and a trained individual blind to the experimental design annotated the videos frame-by-frame for attack, mounting, head and/or flank-directed sniffing (sniffing) and anogenital sniffing (AG sniffing).

Imaging frames were spatially downsampled by a factor of two (that is,  $0.5 \times$ ) in the  $x$  and  $y$  dimensions, and normalized by dividing by a spatially band-passed version of the frame (lower bound,  $100 \mu\text{m}$ ; upper bound,  $10,000 \mu\text{m}$ ). Frames collected over the course of a single day were concatenated into a single stack and registered to each other to correct for motion artefacts using the Mosaic software

package (Inscopix)<sup>18,33</sup>. Imaging data were calculated as relative changes in fluorescence,  $\Delta F(t)/F_0 = (F(t) - F_0)/F_0$  where  $F_0$  was the mean of all frames across the stack. To extract single cells and their  $\text{Ca}^{2+}$  activity traces from the fluorescent imaging frames, we applied an established computational algorithm for cell sorting based on PCA–ICA<sup>35</sup> to generate a series of spatial filters that each corresponded to a single cell. Filters were individually inspected manually, and those that did not appear to correspond to single neurons were discarded<sup>10,33</sup>. Thresholded, segmentation masks for each spatial filter (that is, cell) were projected onto the  $\Delta F(t)/F_0$  stack of frames to extract a time course of fluorescence changes for that cell. On average, we identified  $228 \pm 64$  cells per mouse per day using PCA–ICA (total of 3,396 cells imaged in seven mice; 2,379 cells from the five mice that showed mounting and attack). For each cell,  $\Delta F(t)/F_0$  traces were normalized by the standard deviation of the first baseline trial, and reported in units of standard deviation ( $\sigma$ ).

**Registration across days.** Individual spatial filters, each corresponding to a cell, for each day of imaging, were thresholded and combined into maximum intensity projection maps (Extended Data Fig. 7). Intensity-based image registration was used to identify a pair of affine transformations for aligning the day 1 and day 3 maps to the day 2 map. Overlapping filter triplets from the three days were manually screened for accuracy. All selected triplets had an average of less than  $3 \mu\text{m}$  Euclidean distance between centroids on pairs of days (day 1 versus day 2, day 2 versus day 3, and day 1 versus day 3; average centroid separation of registered filters =  $0.98 \pm 0.03 \mu\text{m}$ ). Roughly 40% of all spatial filters (cells) were registered across three days of imaging ( $n = 99 \pm 37$  cells per mouse).

**Ensemble response vector.** The time-averaged  $\Delta F(t)/F_0$  activity of each cell was computed over all frames of close social interactions (sniffing, mounting or attack) during each trial, and concatenated into a  $N \times 1$  vector ( $N =$  cells) called the population activity (or ensemble response) vector. The use of responses from periods of social interaction precluded any effects of inter-animal distance<sup>15</sup> in our representations. Nevertheless, analyses using different epochs during a trial (for example, only frames from periods of sniffing, or distance-based criteria as described below) did not fundamentally change our observations (Extended Data Fig. 3).

**Quantifying the distance modulation of cell responses.** Cells excited or inhibited by conspecifics ( $\text{mean}|\Delta F(t)/F_0| > 2\sigma$  during periods of male- or female-interaction) showed varying degrees of modulation by the distance to the intruder. To quantify the distance modulation of responsive cells, we used the distance ( $d$ ) between centroids of the two mice (see ‘Position tracking’) to identify frames in which the two animals were not interacting (and were not within  $\pm 1$  s of interaction) and were close ( $d < 35\%$  diagonal length of cage; Extended Data Fig. 3) or far ( $d > 70\%$ ).  $r_{\text{close}}$  and  $r_{\text{far}}$  were measured as the average  $\Delta F(t)/F_0$  of each cell for each distance condition and used to compute a distance modulation index ( $m$ ) as  $m = |r_{\text{close}} - r_{\text{far}}| / (|r_{\text{close}}| + |r_{\text{far}}|)$ ; a measurement of the extent to which cell activity varied with inter-animal distance. Example cells with a low value of  $m$  ( $0 < m < 0.2$ ) showed little modulation of activity depending on distance between the animals (Extended Data Fig. 3e2).

**PCC of male–female similarity.** Because on any given trial, the imaged resident encountered a conspecific of one sex, the PCC between male and female representations was computed between a given trial’s population activity vector and the average population activity vector during the two nearest neighbouring trials of the other intruder sex. For the first and last trials, only the one temporally closest other-sex trial was used.

**Fraction of variance captured by intruder sex.** The difference of covariances method<sup>36</sup> was used to identify the two largest male- and female-specific sources of variance in the ensemble response. For a resident mouse on a given day, the covariance matrix between cells on all male trials ( $C_M$ ) and on all female trials ( $C_F$ ) was generated, and the difference between them ( $C = C_M - C_F$ ) was calculated. By construction, the eigenvectors of  $C$  with the largest positive and negative eigenvalues correspond to the largest male- and female-specific sources of variance, respectively. To measure the fraction of variance captured by this pair of eigenvectors, the  $N \times T$  (neurons by time) matrix of ensemble activity on a given day of imaging was reconstructed as a weighted sum of these two vectors using linear regression. The  $R^2$  value of the reconstruction was reported as the fraction of variance captured.

**Dimensionality reduction.** Low-dimensional representations for visualizing changing ensemble dynamics over time were constructed using PCA<sup>37</sup> or partial least squares (PLS) regression<sup>38</sup> (using the `plsregress` function in MATLAB). For PLS, all traces from a day of imaging were concatenated and regressed against a  $1 \times T$  vector with entries valued at  $-1$  (if a male intruder was present),  $1$  (if there was a female intruder) or  $0$  (otherwise).

**Decoding intruder sex (Fig. 3f).** To determine whether VMHvl activity contained information about intruder sex, a linear decoder was used to decode the sex of the intruder from ensemble activity during periods of social interaction. Training data was constructed from the set of  $N \times 1$  ( $N =$  neurons) population



activity vectors from all frames occurring during social interaction. To avoid effects of dataset size across imaged mice, a random subset of 50 cells was drawn from the set of all cells that were imaged that day. This training data, along with intruder sex labels, was then used to train a linear SVM decoder<sup>39</sup>. Decoder performance was evaluated using cross-validation, by excluding all data from one intruder during training and using that data to test decoder performance; this was repeated 15 times for each intruder, each time with a different random subset of cells, and performance computed as the average accuracy across test sets. Analysis was repeated for each day of imaging, with a separate decoder trained for each day.

**Decoding intruder sex following intruder introduction (Fig. 11).** We wanted to identify how quickly, from the moment the intruder entered, a linear decoder could predict the sex of the intruder from the ensemble activity. The decoder performance was tested in half-second windows, starting 5 s before to 10 s after the moment that the intruder fully entered the resident's home cage. The training set was constructed by averaging the  $\Delta F(t)/F_0$  for each cell within a 0.5 s time bin centred at a time  $t$ , for each male and female trial on one day (typically 10 trials total). Overfitting to the small ( $n=9$ ) number of training examples was reduced by projecting training data onto the first 10 principal components of ensemble activity from that day; we found that this improved decoder performance. The projected data and intruder sex labels were then used to train a linear SVM decoder, and decoder performance was evaluated using leave-one-out cross-validation with each trial held out once. To allow dynamic changes in intruder-evoked responses, a new decoder was trained at each time point.

**Decoding behaviour (Fig. 2b).** To determine whether ensemble activity contained information about specific social behaviours, a separate binary linear decoder was trained for each behaviour, to predict whether the behaviour was occurring or not. To avoid confounding representations of behaviour with representations of intruder sex, separate decoders were trained for male- and female-directed behaviours. Training data was constructed from the set of  $N \times 1$  ensemble activity vectors from all frames of all day 3 trials with a given intruder sex, labelled positive or negative for a target behaviour. Negative training data (that is, data from periods when the behaviour of interest was not occurring) was under-sampled to produce an equal number of positive and negative training examples. These binary labels and their corresponding ensemble activity vectors were then used to train a linear SVM decoder. Decoder performance was tested by cross-validation by excluding all data from one intruder during training, and computing the average decoder performance on the held-out data using the F1 score (see below). Testing was repeated for each intruder, and for 15 iterations of each training set.

**Measuring performance of behaviour decoders (Fig. 2c).** Because most behaviours occurred on a small fraction of tested frames, reporting the accuracy (true positives + true negatives) / (total samples) of a trained behaviour decoder could be misleading. Instead, we reported the F1 score,  $F1 = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$ , where precision (also called specificity) is the fraction of predicted positives that are actual (ground-truth) positives (precision = (true positives) / (true positives + false positives)) and recall (also called sensitivity) is the fraction of actual (ground-truth) positives that are correctly predicted positive (recall = (true positives) / (true positives + false negatives)).

F1 scores of trained SVMs were computed by shuffling both the cells and the behaviour annotations. To permute behaviour annotations while preserving their temporal structure, all behaviour bouts and all inter-bout intervals were identified for a given day/intruder sex/behaviour, and synthetic behaviour rasters were constructed by randomly drawing without replacement from the set of bouts and the set of inter-bout intervals. The trained SVM weights were applied to the shuffled cells, and the F1 score was computed relative to the shuffled behaviour. Each form of shuffling was repeated  $100 \times$ , giving 10,000 shuffled observations; we considered observed F1 scores as significant if they fell above the 97.5 percentile of the distribution of chance F1 scores.

**Mahalanobis distance (Fig. 3h, i).** For each imaged mouse, the Mahalanobis distance was used as an additional measurement of male and female representation separation. Two distributions of  $N$ -dimensional ( $N$ =neurons) population activity vectors were generated from all imaging frames during which social interaction

occurred: one distribution for male trials and the other for female trials. The Mahalanobis distance ( $d$ ) between each vector (the test vector) and these two distributions were computed, to give  $d_{\text{same}}$  (distance to points of the same sex as the test vector) and  $d_{\text{opp}}$  (distance to points of the opposite sex), which were then averaged across all test vectors. Points within the distributions that came from the same trial as the test vector were omitted during distance calculation.

**Choice probability (Fig. 2d-h).** Choice probability (CP) analysis was used to measure a cell's tuning, defined here as how well two conditions could be predictively discriminated from a single cell's activity<sup>17</sup>. The CP of a given cell, for a pair of (behavioural) conditions, was computed by constructing a histogram of that cell's  $\Delta F(t)/F_0$  values under each of the two conditions. These two histograms were plotted against each other to generate an ROC (receiver-operating characteristic) curve. The integral of the area under this ROC curve generated the CP value for each cell, with respect to each of the two behavioural conditions (see Extended Data Fig. 4). This CP value is bounded from 0 to 1, where a CP of 0.5 corresponded to no difference between conditions. All frames from all behaviour bouts lasting two or more seconds were used to compute the CP, and data from the second and third day of imaging were combined. Cells from mice that spent less than 2% of time engaging in any of the four examined behaviours (attack, mounting, male-directed sniffing and female-directed sniffing) were omitted from this analysis, owing to noisiness of their CP estimates.

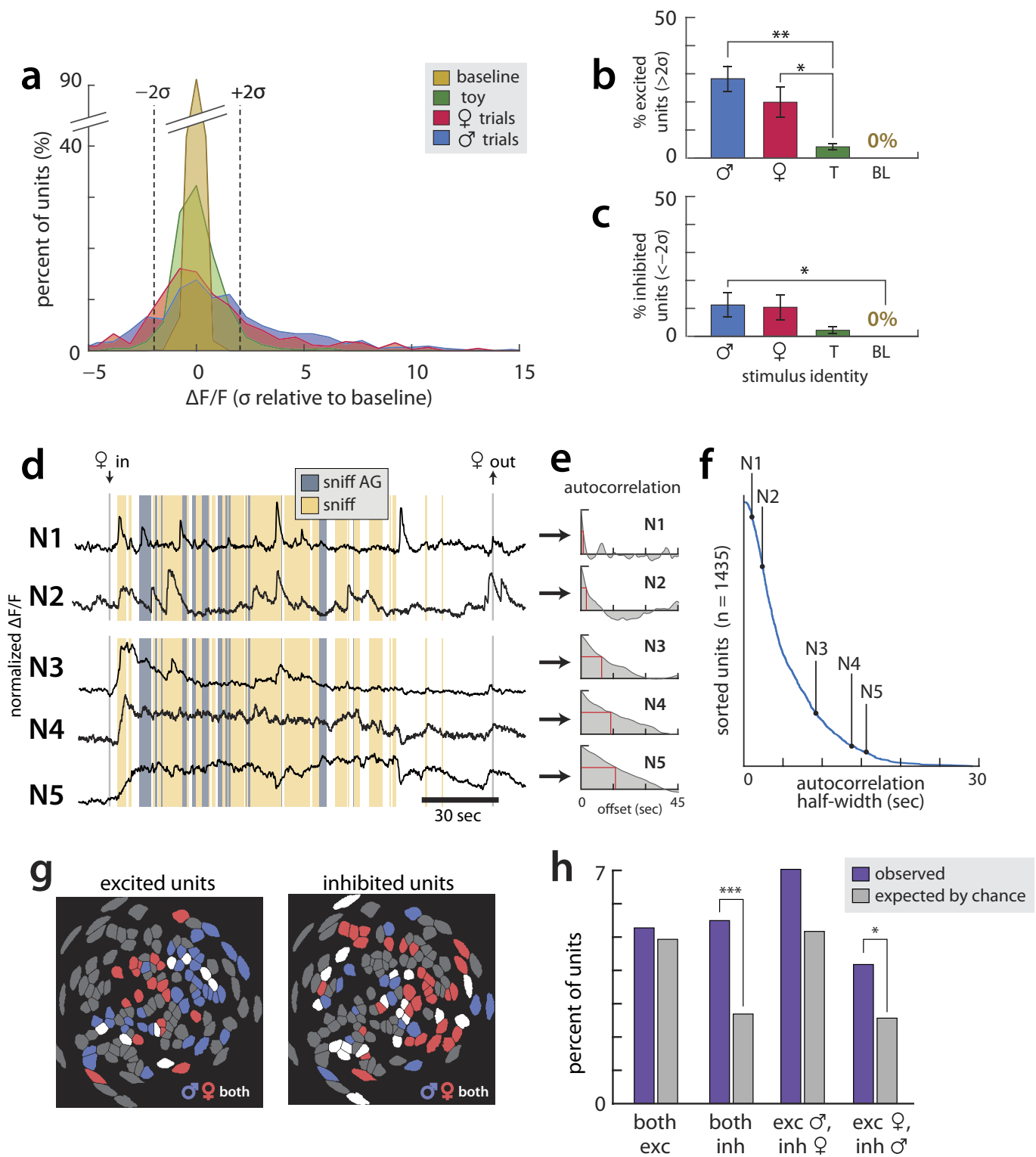
The significance of a cell's CP, relative to chance, was computed using shuffled bout timings for the two compared conditions (as was done for estimating chance decoder performance). Shuffling was repeated  $100 \times$ , from which the mean and s.d. of shuffled CPs was estimated. Any observed CP that was  $\geq 2$  s.d. above the shuffled mean was considered significant.

**Non-negative LASSO regression of behaviour against ensemble separation (Fig. 4a-c).** To identify behavioural correlates of ensemble separation, the PCC and the cumulative time spent engaging in each behaviour was computed for each animal and trial (around 30 trials per mouse;  $n=6$  mice for which imaging data was available for all three days; this includes one of the two mice that did not show mounting or attack behaviour, and whose sex representations did not segregate). The weighted sum of cumulative behaviours most correlated with the PCC was found using a non-negative least-angle regression (LARS) implementation of least absolute shrinkage and selection operator (LASSO)<sup>20</sup>, to impose sparseness and non-negativity constraints on the fit weights. The sparseness parameter of LARS was selected to minimize the cross-validated mean squared error of the fit. For all behaviours, a square-root transform of cumulative behaviour times, before LASSO regression, increased their correlation with the PCC.

**Code availability.** Custom code written for the purpose of this study is accessible at [https://github.com/DJALab/VMHv1\\_imaging](https://github.com/DJALab/VMHv1_imaging).

**Data availability.** Imaging and behavioural data will be made available by the corresponding author upon reasonable request.

- Ogawa, S. *et al.* Abolition of male sexual behaviors in mice lacking estrogen receptors  $\alpha$  and  $\beta$  ( $\alpha\beta$ ERKO). *Proc. Natl Acad. Sci. USA* **97**, 14737–14741 (2000).
- Dana, H. *et al.* Thy1-GCaMP6 transgenic mice for neuronal population imaging *in vivo*. *PLoS ONE* **9**, e108697 (2014).
- Resendez, S. L. *et al.* Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nat. Protocols* **11**, 566–597 (2016).
- Dollár, P., Welinder, P. & Perona, P. Cascaded pose regression. In *IEEE Conference on Computer Vision and Pattern Recognition* 1078–1085 (CVPR, 2010).
- Mukamel, E. A., Nimmerjahn, A. & Schnitzer, M. J. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* **63**, 747–760 (2009).
- Machens, C. K., Romo, R. & Brody, C. D. Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
- Jolliffe, I. *Principal Component Analysis* (Wiley Online Library, 2002).
- De Jong, S. SIMPLS: an alternative approach to partial least squares regression. *Chemom. Intell. Lab. Syst.* **18**, 251–263 (1993).
- Cristianini, N. & Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* (Cambridge Univ. Press, 2000).

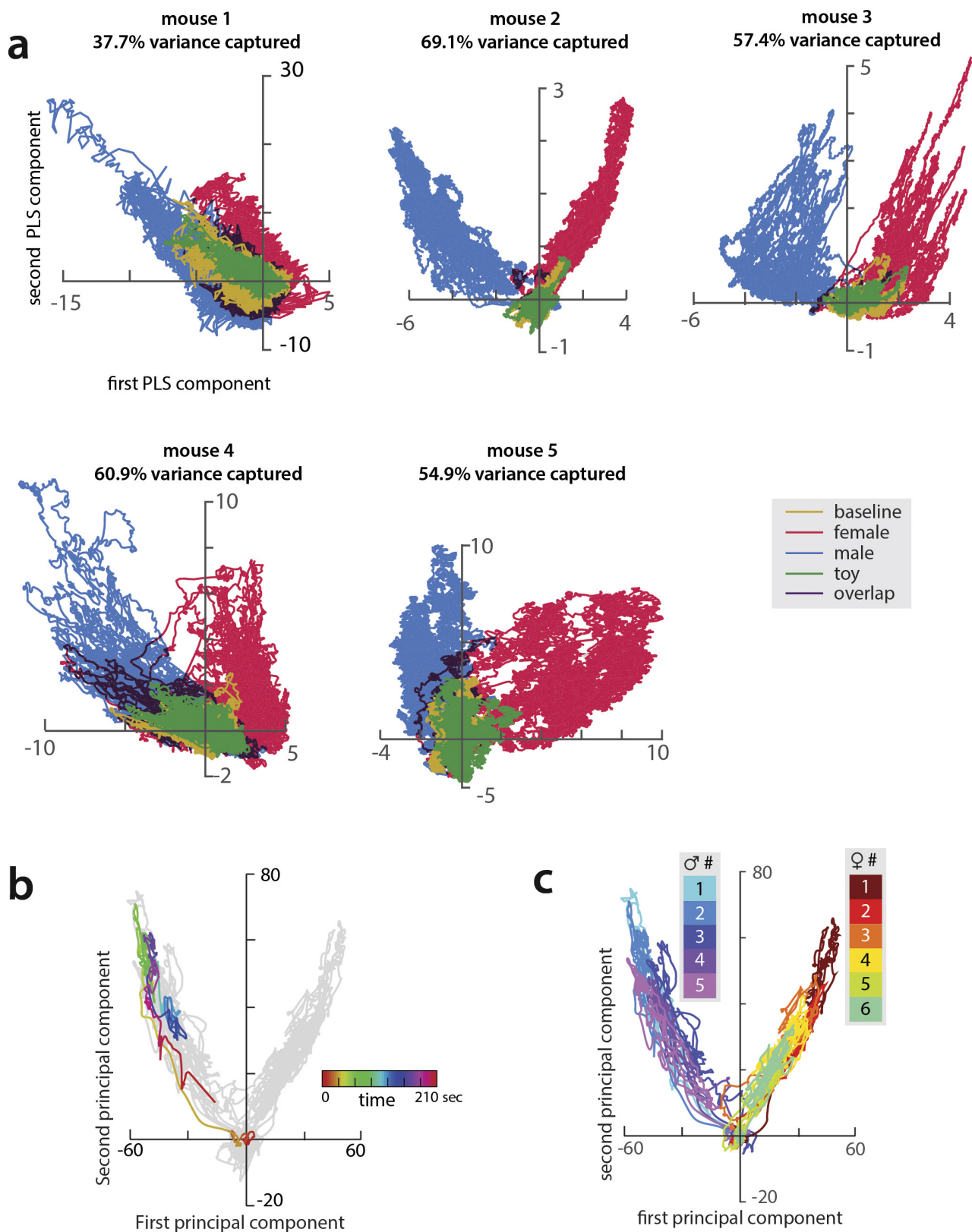


Extended Data Figure 1 | See next page for caption.

**Extended Data Figure 1 | Properties of  $Esr1^+$  neuron responses in VMHvl during free social interactions.** Data presented are from the third day of imaging. **a**, Histogram of time-averaged  $\Delta F(t)/F_0$  values observed in the 30 s following stimulus introduction (or the first 30 s of imaging, on baseline trials). Dashed lines indicate thresholds for identifying excited ( $\geq 2$  s.d. ( $2\sigma$ ) of baseline above the mean pre-intruder baseline) and inhibited ( $\leq -2\sigma$  below pre-intruder baseline) cells. ( $n = 5$  imaged mice, 5 male trials, 5 female trials, 2 toy trials, 2 baseline trials per mouse.) **b, c**, Average percentage of time cells were excited (**b**) or inhibited (**c**) in first 30 s of imaging. **d**, Example traces from single  $Esr1^+$  cells, showing example response profiles during a single trial interaction with a female conspecific. While some cells showed transient calcium responses (N1, N2), others showed slower dynamics (N3–N5), including persistent elevation or suppression of activity for the duration of the encounter with the female. We believe the latter responses reflect the ongoing activity

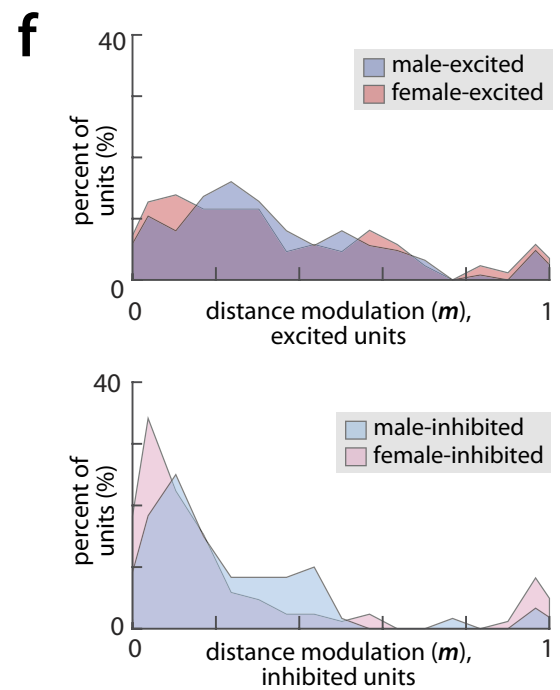
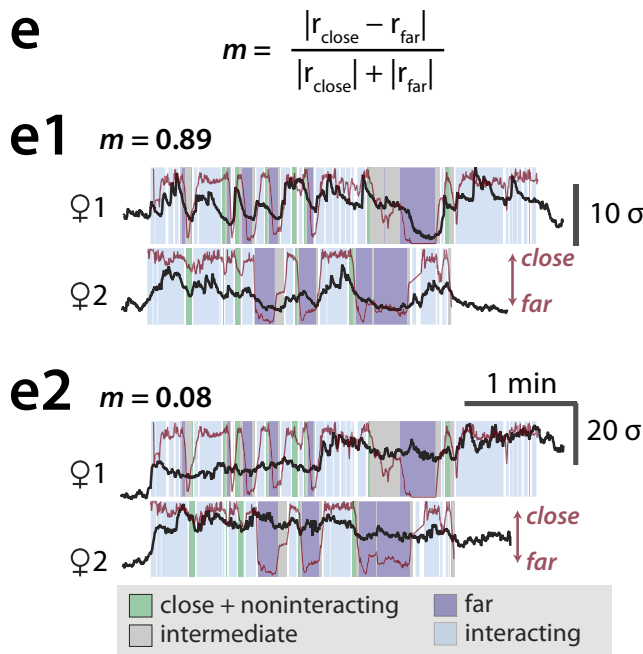
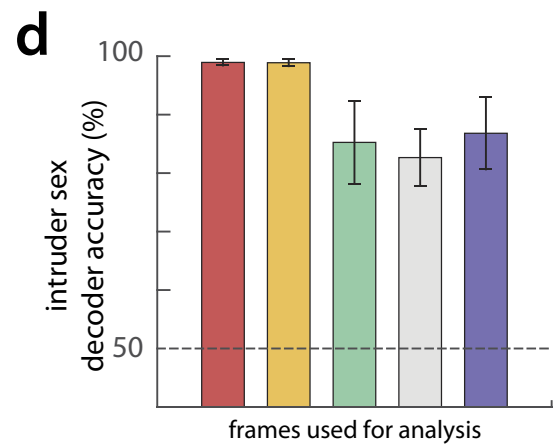
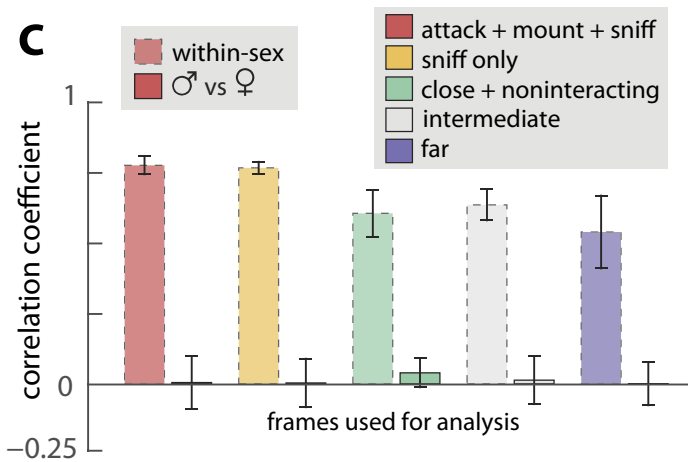
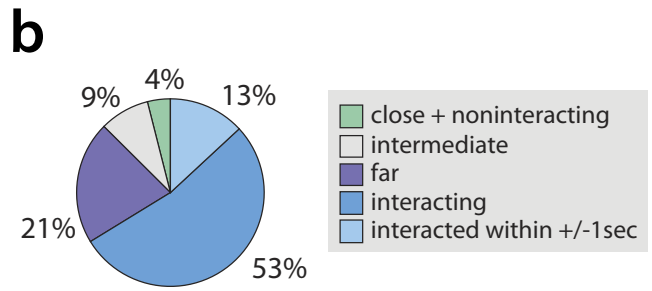
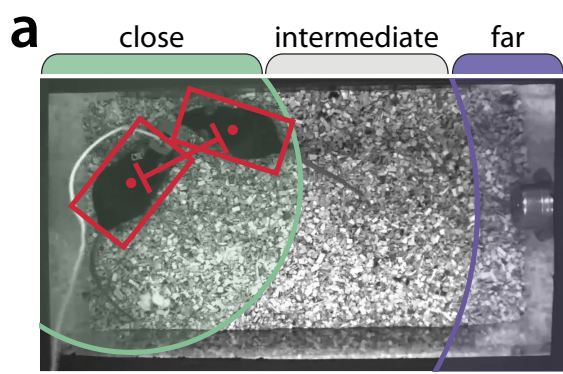
of some VMHvl cells, as filtered by the slow dynamics of GCaMP6s. **e**, Autocorrelation functions computed across all day 3 trials, for each example cell in **d**; red line indicates half-width. **f**, Autocorrelation half-width computed across all day 3 trials, for all cells from all mice ( $n = 1,435$  cells in five mice), sorted from smallest to largest, revealing a continuous range of values. Autocorrelation half-width of the five cells in **d** are indicated. **g**, Example spatial maps of excited (left; mean  $\Delta F(t)/F_0 \geq 2\sigma$  during periods of social interaction) and inhibited (right; mean  $\Delta F(t)/F_0 \leq -2\sigma$  during interaction across all male and female conspecifics on the third day of imaging) cells in the presence of male and female conspecifics. **h**, The percentage of cells across all mice ( $n = 5$ ) that were excited and/or inhibited by both males and females, compared to the percentage expected by chance assuming statistical independence. (both inhibited,  $P = 2.17 \times 10^{-4}$ ; excited by female + inhibited by male,  $P = 0.0299$ ).





**Extended Data Figure 2 | Representations of males and females in all individual mice.** **a**, Day 3 ensemble representations of intruder sex for the five mice that mounted and fought with conspecifics, projected onto the first two axes of a PLS regression against intruder sex. Traces are coloured by intruder sex identity. Percentage of variance explained by the first two

PLS components is noted for each mouse. **b**, Projection of a single trial onto the first two principal axes of one example mouse (mouse 2), colour-coded to highlight temporal trajectory of the population activity vector during the trial. Principal axes were identified using PCA (same axes as in Fig. 1k). **c**, All trials from mouse 2, coloured by trial number.

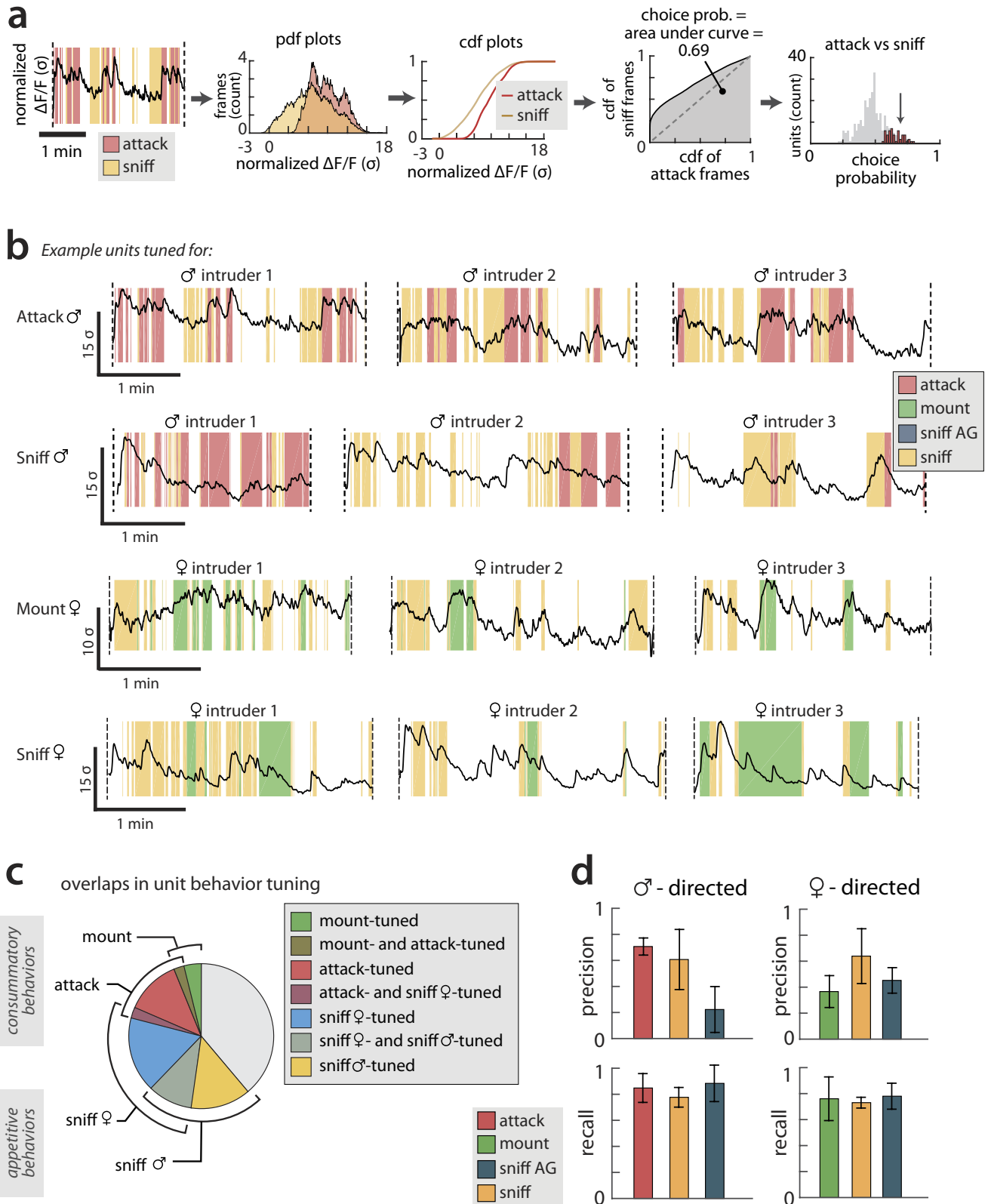


Extended Data Figure 3 | See next page for caption.

**Extended Data Figure 3 | Distance dependence of intruder sex representations.** **a**, Example video frame showing estimates of resident and intruder poses and centroids (red bounding boxes and points, respectively) produced by an automated tracker<sup>3,4</sup>. Green, grey, and purple bars mark the inter-mouse distances (relative to the resident) categorized as close, intermediate and far. **b**, Percentage of time animals spent either interacting with conspecifics or present within the three zones. Time points in which animals have recently interacted (that is, interaction has occurred within  $\pm 1$  s) are shown as their own category as a control for the effect of slow GCaMP dynamics. **c**, PCC between trials of the same intruder sex, or between male and female trials, where representations of intruders are computed by averaging each cell's  $\Delta F(t)/F_0$  across the indicated subsets of imaging time points. **d**, Accuracy of a linear SVM decoder for intruder sex, using representations computed as in **c**. Decoder

accuracy was not significantly affected by distance. **e**, Top, the distance modulation index ( $m$ ), a measurement of the extent to which cell activity is modulated by inter-animal distance, is computed from  $r_{\text{close}}$  (the average response of a cell when animals are close but have not recently interacted (within  $\pm 1$  s)), and  $r_{\text{far}}$  (the average response of that cell when animals are far apart and have not recently interacted). **e**, Bottom, example  $\Delta F/F$  traces (black) from two cells that have a high (**e1**,  $m = 0.89$ ) and low (**e2**,  $m = 0.08$ ) distance modulation index, and corresponding inter-animal distances (brown). **f**, Histograms of values of  $m$  observed in all cells that are significantly excited (above) or inhibited (below) during interaction with males or females. Note that inhibited cells are less sensitive to inter-animal distance than are excited cells, and that the distribution of  $m$  is similar for cells that are responsive to male and female intruders.

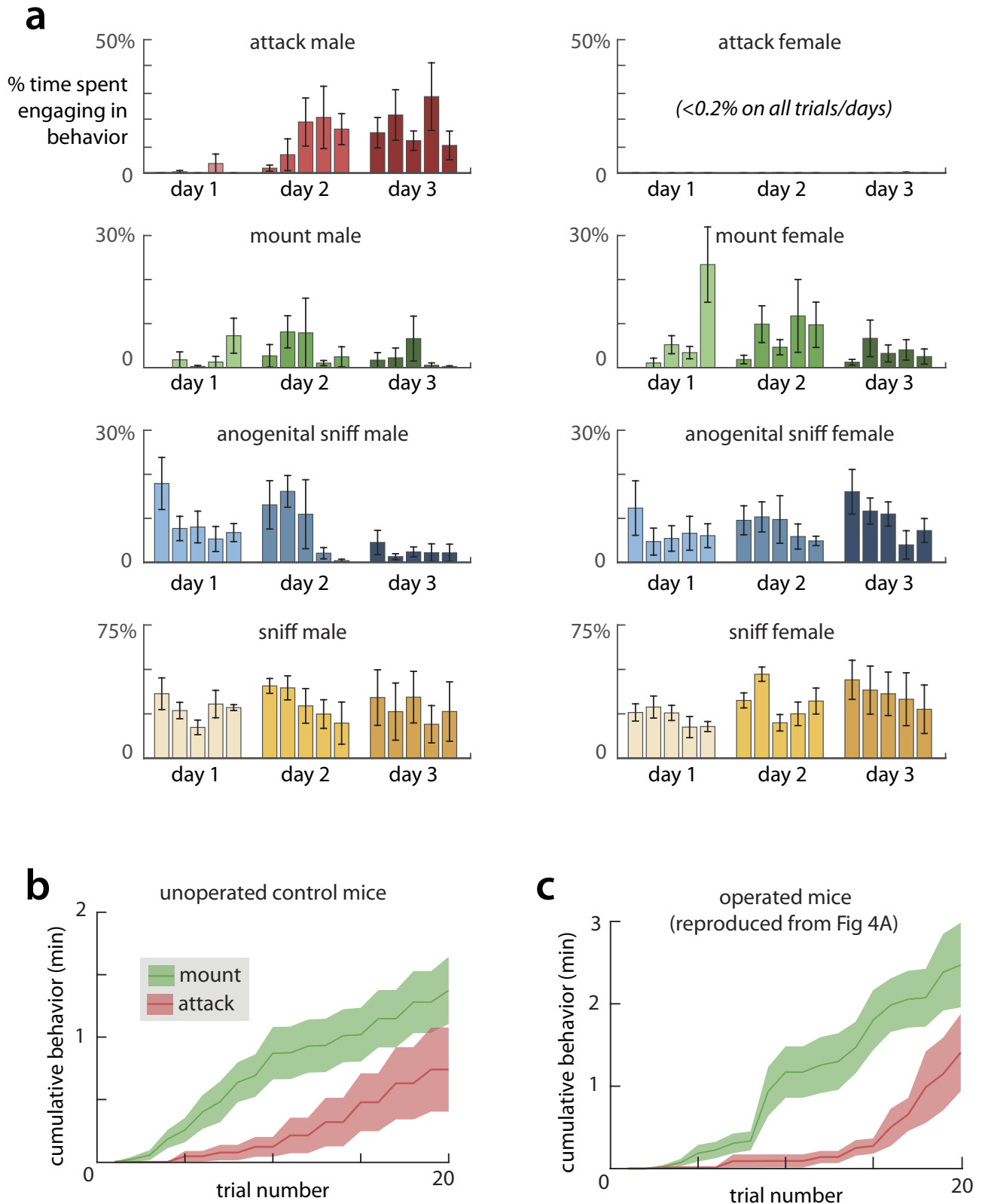




Extended Data Figure 4 | See next page for caption.

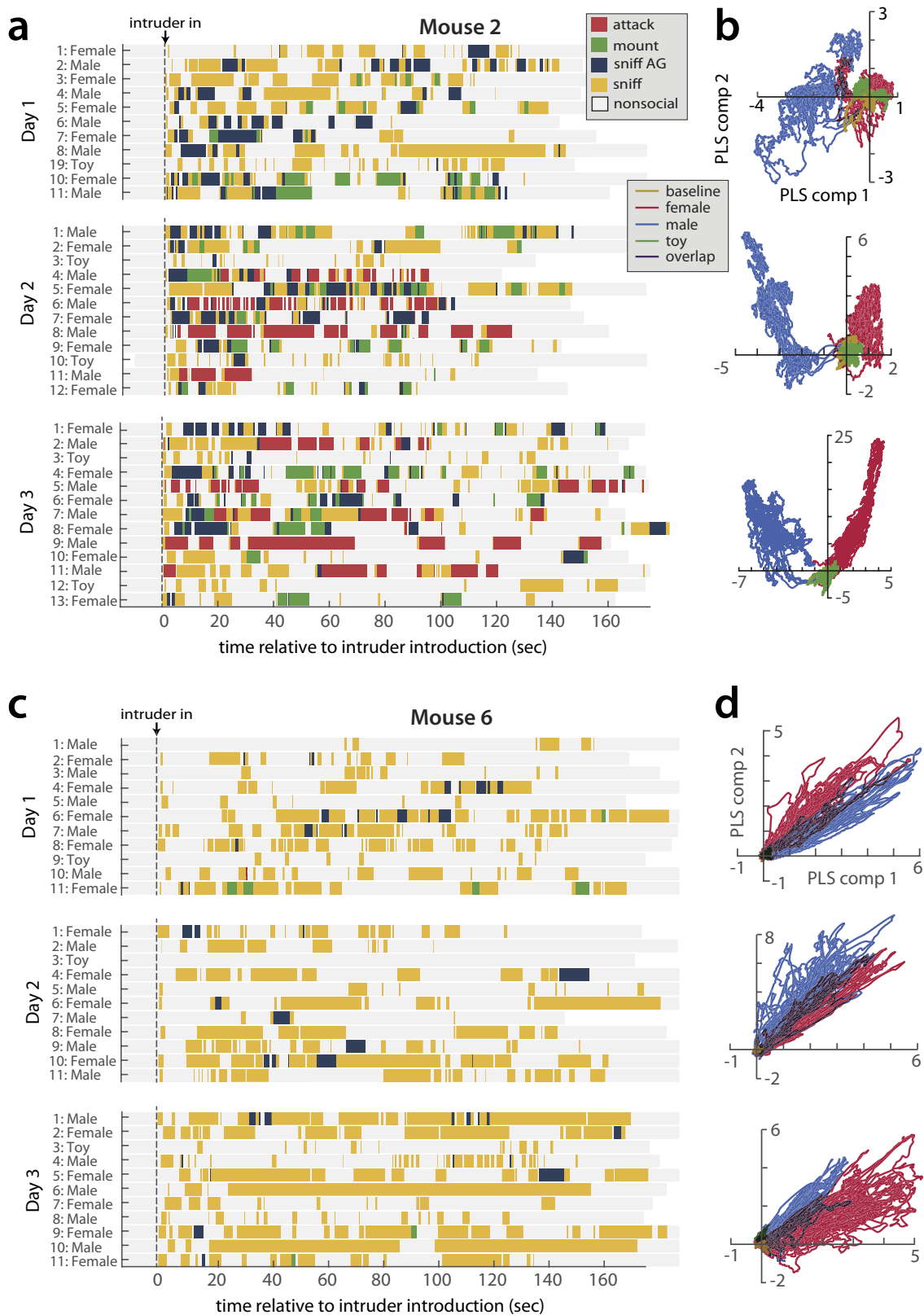
**Extended Data Figure 4 | Choice probability histograms and example cells.** **a**, Steps for computing CP, a measurement of the discriminability of two conditions, given the  $\Delta F(t)/F_0$  of a single cell. Given a pair of behaviours (here, male-directed attack and sniffing), we computed the distribution of  $\Delta F(t)/F_0$  values for this cell for each behaviour, and then integrated each distribution to produce a pair of cumulative distribution functions (cdfs). The area under the ROC curve formed from these two distributions is defined as the cell's CP; CP of this example cell was 0.69. Right-most panel shows CP for all cells (arrow indicates CP of the example cell); outlined bars denote cells for which CP was significantly higher than chance. **b**,  $\Delta F/F$  traces and corresponding behaviour rasters from example cells that showed significant CPs for the specified behaviour, illustrating the relative changes in  $\Delta F/F$  and behaviour. Three representative trials are shown (same examples as in Fig. 2 d–g). **c**, Proportion of cells significantly tuned for each examined behaviour. Overlap between blocks reflects the

proportion of cells that were significantly tuned for both behaviours. Note that by this metric, cells cannot be tuned for both attack and male-directed sniffing, nor for both mounting and female-directed sniffing, because the CP of a cell is defined by comparing  $\Delta F(t)/F_0$  values under these paired conditions. All cells significantly tuned for attack, mount or sniff also showed a significant CP for periods of close social interactions (including sniffing and/or mount or attack) versus periods of non-social interactions, by construction. However, some cells that showed a significant CP for close social interactions versus non-social periods did not show a significant CP for a specific behaviour (see Fig. 2h). **d**, Separate precision (true positives / (true positives + false positives)) and recall (true positives / (true positives + false negatives)) scores for the behaviour decoders presented in Fig. 2b, c. The F1 score (presented in Fig. 2c) is defined as  $2(\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$ .



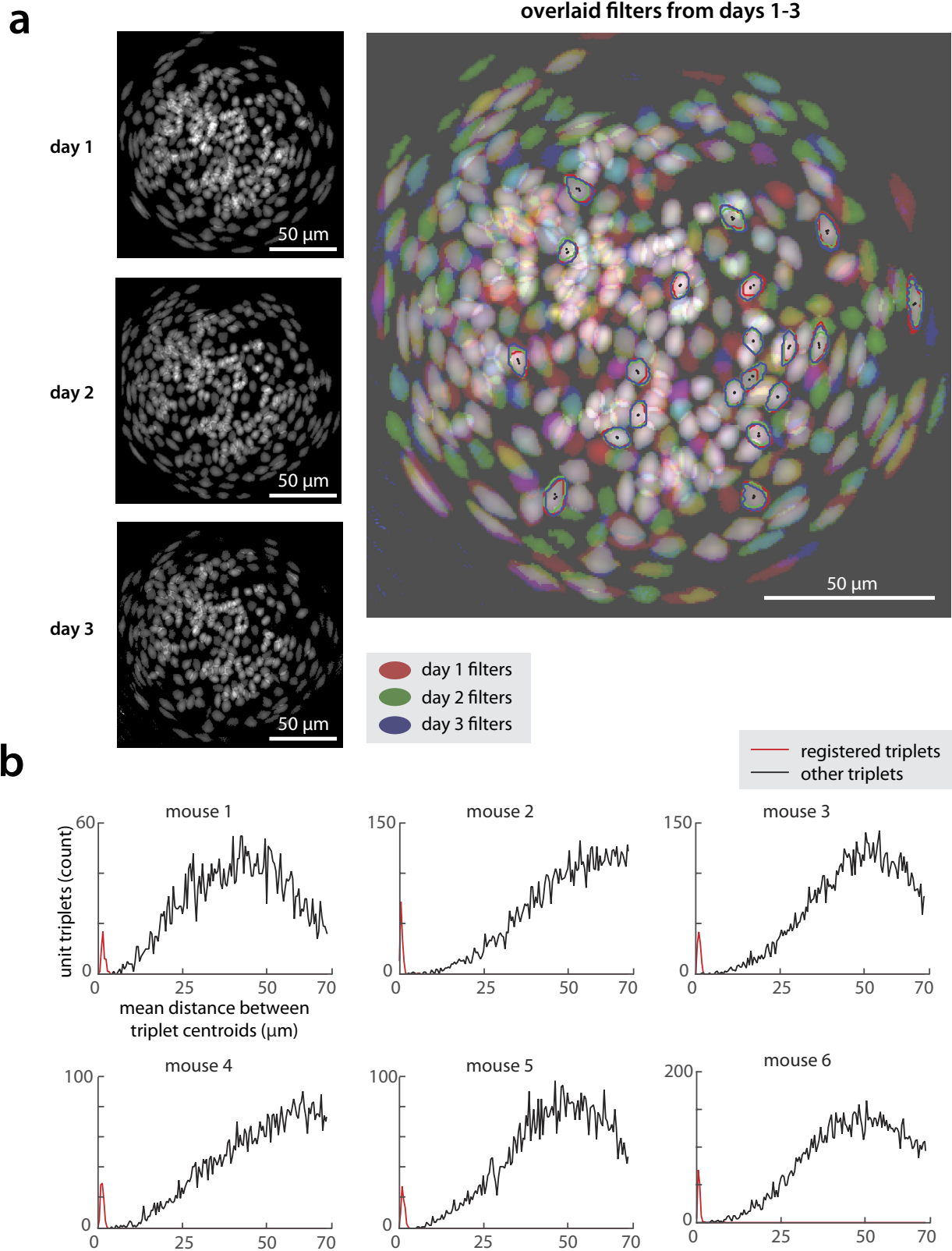
**Extended Data Figure 5 | Expression of social behaviours across trials and days. a,** Percentage of time mice spent engaging in male- and female-directed behaviours on each of five trials on three days of imaging.  $n = 5$  mice, mean  $\pm$  s.e.m. **b,** Cumulative time spent mounting and attacking by a cohort of unoperated, socially isolated mice that underwent two days of the standard social experience paradigm depicted in Fig. 1d ( $n = 8$  mice).

The unoperated mice exhibited a gradual appearance of mounting, and showed mounting behaviours before attack, indicating that changes in behaviour were not due to the effects of surgery or the presence of the scope. **c,** Data from implanted mice ( $n = 5$ ), reproduced from Fig. 3g and restricted to the first 20 trials for direct comparison.



**Extended Data Figure 6 | Comparison of the social behaviours and neural representations in example mice that showed and did not show aggression. a–d, Behaviour rasters (a, c) and the corresponding first two PLS components (b, d) for example mice that showed (a, b) and did not (c, d) show aggression to conspecific males over the course of the experiment.**





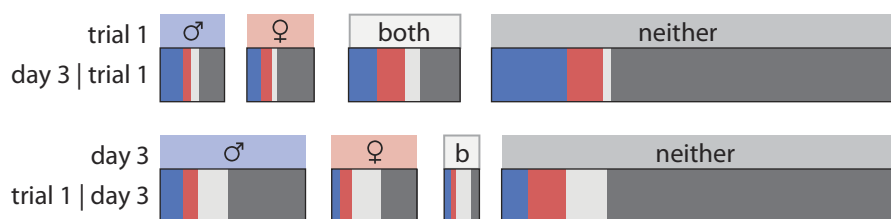
**Extended Data Figure 7 | Registration of cells across three days of imaging.** **a**, Left, maximum projection maps of spatial filters of all cells from each of three days of imaging for an example mouse. Right, RGB-composite image of the three left images (day 1 = red channel, day 2 = green channel, day 3 = blue channel), showing overlap of registered filters (overlap in all three channels appears in white). Outlined are 20 example cells (out of 135) that could be identified in all three days of

imaging: red/green/blue lines indicate filter outlines on first/second/third day, respectively; black points mark filter centroids. **b**, Histograms of average distance between filter centroids from days 1–3, in cells that could be tracked across days (red) as compared to random triplets of cells (black). Day 1–3 centroids from tracked cells were separated by an average of  $2.15 \pm 0.06 \mu\text{m}$  (mean  $\pm$  s.e.m.,  $n = 593$  cells tracked across days in six mice, during the standard resident–intruder assay).

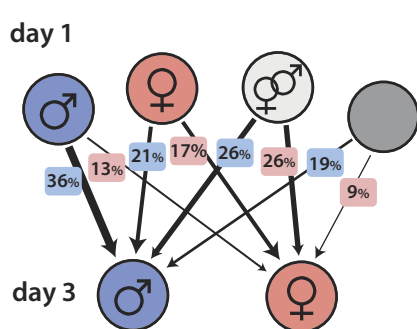
**a**

		day 3 active units				TOTAL
		♂	♀	both	neither	
trial 1 active units	♂	16	6	5	18	45
	♀	10	8	3	26	47
	both	20	20	10	28	78
	neither	53	25	6	201	285
	TOTAL	99	59	24	273	455

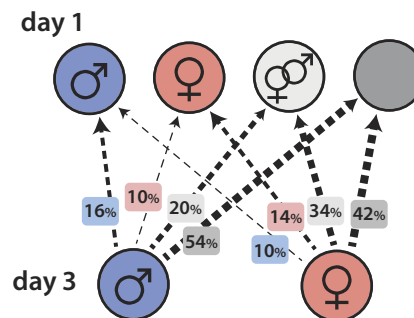
**b**



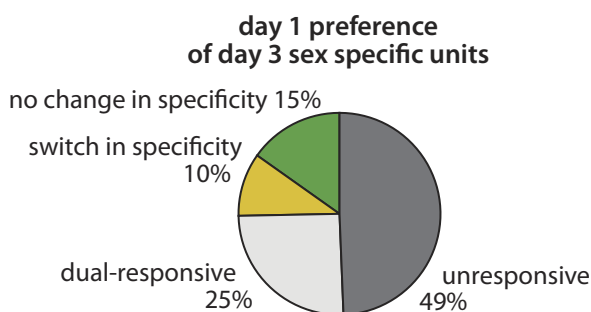
**c**



**d**

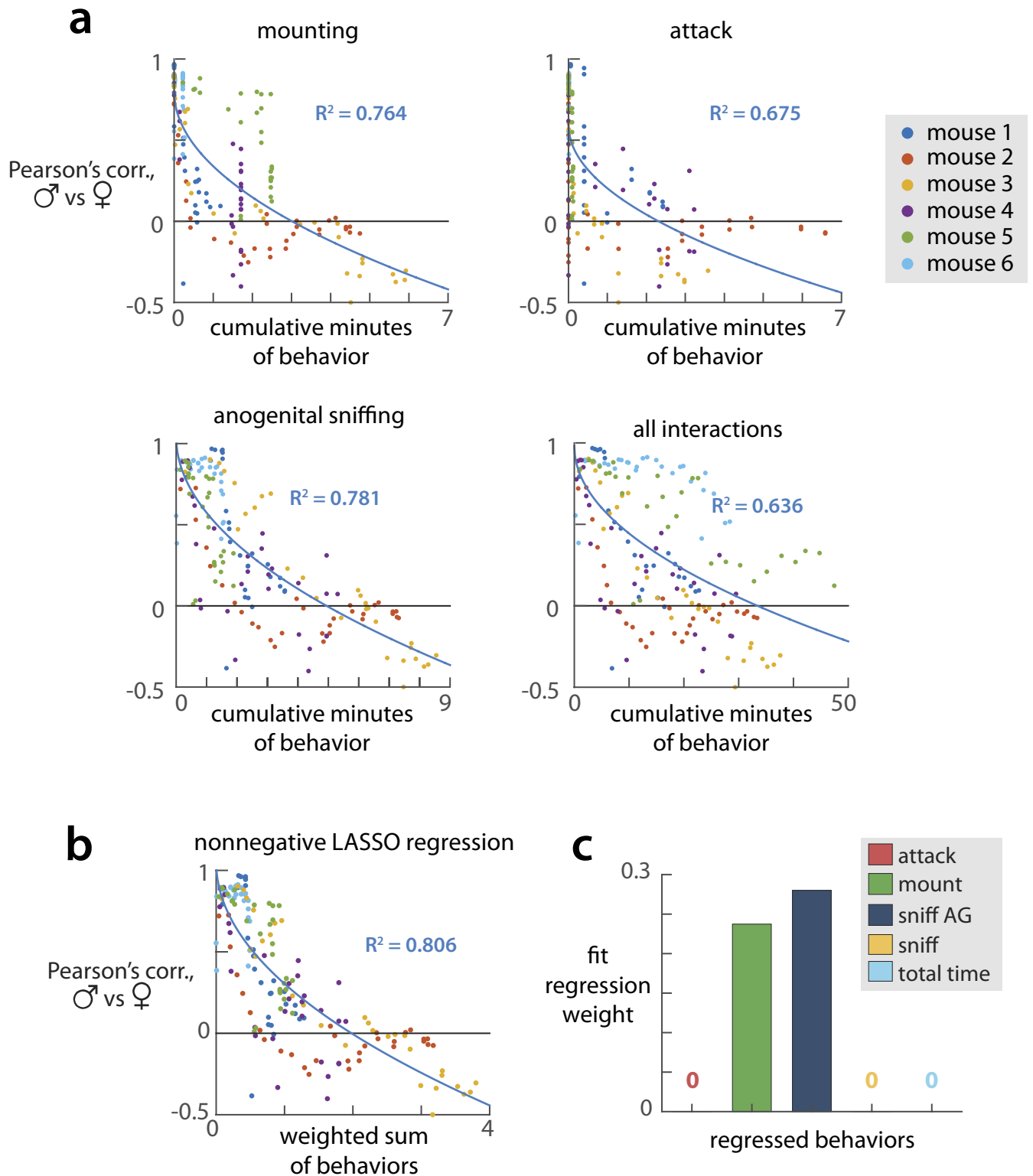


**e**



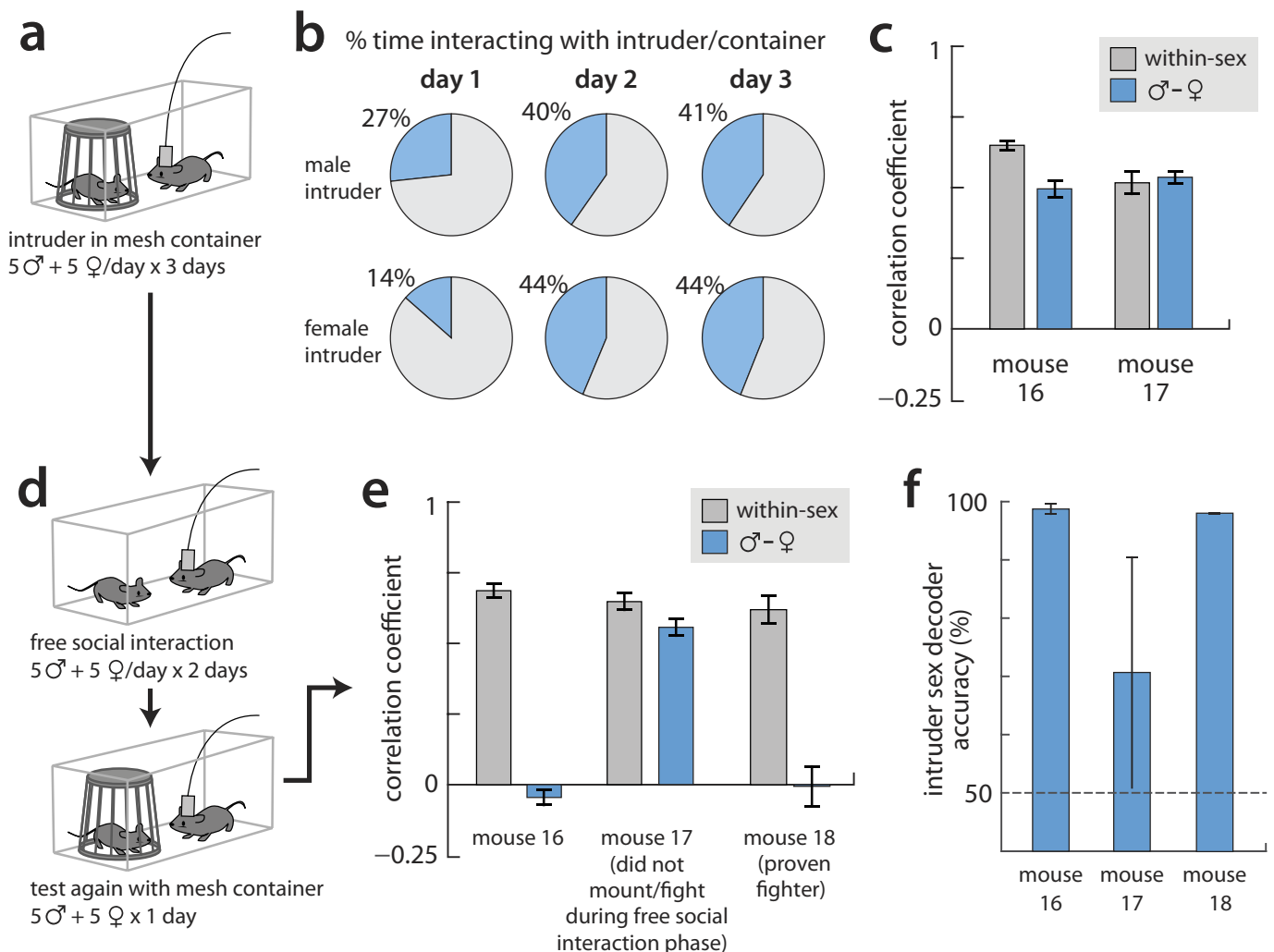
**Extended Data Figure 8 | Preference changes of *Esr1*<sup>+</sup> neurons during the acquisition of social experience.** **a**, Responses on day 1 trial 1 are compared to active cells on day 3, for all mice and all cells that were registered across three days of imaging ( $n = 455$  cells analysed in the five mice that showed attack/mounting by day 3 of the standard resident-intruder assay). **b**, Top, response properties of cells on day 3, conditioned on their responses on day 1 trial 1, grouped according to the response types on trial 1. For example, among cells that responded to both males and females on day 1 trial 1, around 26% (20 out of 78) specifically responded to males on day 3, around 26% responded specifically to females; additionally, around 35% responded to neither sex

and around 13% responded to both. The percentages of these categories are summarized in **c**. Bottom, response properties of cells on day 1 trial 1 conditioned on their response properties on day 3, grouped according to response types on day 3. The percentages in different categories are summarized in **c** and **d**. For example, of the cells that showed male-specific responses on day 3, around 54% is derived from cells that responded to neither sex on day 1 trial 1; 20% is derived from cells that responded to both sexes, 10% is derived from initially female-specific cells and 16% is derived from initially male-specific cells. Numbers used to calculate the percentages are from **a**. **e**, Analysis of the day 1 preferences of cells that responded only to males or only to females on day 3.



**Extended Data Figure 9 | Behaviour correlation with male/female PCC.** **a**, Cumulative minutes of each behaviour by the  $n$ th trial plotted against the PCC between male and female representations on that trial, for six mice (behaviours already presented in Fig. 4a–c have been omitted). All trials from a given mouse are shown in the same colour. Solid lines are square-root fits (of the form  $y = a\sqrt{x} + b$ ) of the plotted points from all mice. **b**, A weighted sum of cumulative minutes of each recorded

behaviour (attack, mounting, anogenital sniffing and other sniffing) as well as cumulative time spent interacting with conspecifics, plotted against PCC between male and female representations; weights fit by non-negative LASSO regression. **c**, Bar plot of weights used to generate plot in **b**, fit by non-negative LASSO with sparseness parameter chosen to minimize mean-squared error on held-out data (see Methods).



#### Extended Data Figure 10 | Restricted access using a mesh container.

To exclude the possibility that the lack of representation separation shown in Fig. 4l–o was due to the presence of the experimenter's hand during access-restricted trials, we repeated this experiment with the intruder mouse inside a wire mesh container. **a**, Diagram of experimental setup. **b**, Percentage of time the imaged mouse spent interacting with the intruder on each of the three days ( $n = 2$  mice). Aside from day 1, the barrier to free interactions presented by the container did not reduce the time the resident spent investigating the intruder. **c**, PCC between male and female representations on the third day of the assay (blue bars) showed that the separation of representations did not occur in these two imaged mice. Gray bars show the average PCC between pairs of male trials or pairs of female trials, for comparison. **d**, Following the three days of interactions with the intruder behind the barrier, mice were given two additional days

of free social interaction, before a final day (day 6) in which intruders were again presented inside the mesh container. A third, experienced animal (mouse 18) was also tested with the mesh container. **e**, PCC between representations of males and females presented within the mesh container on day 6; grey bars show average PCC between pairs of male trials or pairs of female trials. Two out of three tested mice showed clear separation of male and female representations; the third (mouse 17) did not, but failed to fight or mate with conspecifics during the free social interactions on the two days preceding the day 6 test. **f**, Performance of an SVM decoder trained to predict intruder sex from the data on day 6, showing high accuracy in the two mice that had previously fought and mounted. These data provide additional evidence that olfactory cues from conspecifics are not sufficient to produce ensemble separation, and that behaviors occurring during free social interactions are required.