Article

Fast odour dynamics are encoded in the olfactory system and guide behaviour

Check for updates

Tobias Ackels[1,2,3], Andrew Erskine[1,2,3], Debanjan Dasgupta[1,2,3], Alina Cristina Marin[1], Tom P. A. Warner[1], Sina Tootoonian[1,2], Izumi Fukunaga[1], Julia J. Harris[1,2] & Andreas T. Schaefer[1,2✉]

Odours are transported in turbulent plumes, which result in rapid concentration fluctuations[1,2] that contain rich information about the olfactory scenery, such as the composition and location of an odour source[2–4]. However, it is unclear whether the mammalian olfactory system can use the underlying temporal structure to extract information about the environment. Here we show that ten-millisecond odour pulse patterns produce distinct responses in olfactory receptor neurons. In operant conditioning experiments, mice discriminated temporal correlations of rapidly fluctuating odours at frequencies of up to 40 Hz. In imaging and electrophysiological recordings, such correlation information could be readily extracted from the activity of mitral and tufted cells—the output neurons of the olfactory bulb. Furthermore, temporal correlation of odour concentrations[5] reliably predicted whether odorants emerged from the same or different sources in naturalistic environments with complex airflow. Experiments in which mice were trained on such tasks and probed using synthetic correlated stimuli at different frequencies suggest that mice can use the temporal structure of odours to extract information about space. Thus, the mammalian olfactory system has access to unexpectedly fast temporal features in odour stimuli. This endows animals with the capacity to overcome key behavioural challenges such as odour source separation[5], figure–ground segregation[6] and odour localization[7] by extracting information about space from temporal odour dynamics.

The turbulent nature of air[1,2,4,8] and water[9,10] flow results in complex temporal fluctuations in odour concentration that depend on the distance and direction of odour sources[1–8,10]. Insects are thought to use the temporal structure of odour plumes to infer, for example, the location[4,7,11–13] or composition[13–15] of an odour source. Mammalian olfaction, by contrast, has generally been considered a slow sense. Individual sniffs are thought to be the unit of information[16], implying that fast changes in odour concentration (at sub-sniff resolution) should be inaccessible to the mammalian olfactory system. However, the neural circuitry of, for example, the mouse olfactory bulb (OB) is in principle capable of millisecond-precise action potential firing[17,18], and is endowed with rich computational resources that could be used to extract fine temporal information from dynamic inputs[19]. Here we show that the mouse olfactory system has access to fast, sub-sniff temporal patterns in the odour scenery and that mice can use this information to detect high-frequency odour correlations, thereby enabling source separation.

Fast odour dynamics encoded in OB inputs

Normal airflow is characterized by complex, often turbulent, flow patterns and imposes a rich temporal structure on odour concentration profiles, with substantial power in frequencies well above typical sniff rates (Fig. 1a). To assess whether the mouse olfactory system has access to this frequency regime, we designed an odour delivery system that can reliably present odours with a bandwidth beyond 50 Hz (Fig. 1b, Supplementary Fig. 1). As prototypical, simplistic high-frequency stimuli, we used two 10-ms square pulses of odour separated by 10 or 25 ms (Fig. 1c). Olfactory sensory neurons (OSNs) are known to be slow in responding to odour stimuli[20]. Both epithelial mucus and the biochemical transduction cascade act as low-pass filters[16,20,21], which suggests that individual OSNs cannot directly follow rapidly fluctuating odour stimuli. However, axons from up to tens of thousands of OSNs that express the same olfactory receptor converge onto one or a few glomeruli in the OB[22]. This organization resembles the auditory system, in which—despite the relatively low temporal resolution of individual cells—population responses faithfully report high-frequency signals[23]. Thus, we built a model of populations of noisy integrate-and-fire neurons with stimulus filtering and neuronal dynamics that matched experimental data to investigate whether this large convergence could aid in detecting high-frequency stimuli in OSNs (Extended Data Fig. 1). Our simulation results suggest that across the thousands of OSNs that express the same OR—although still not directly following the odour profile—the population can faithfully discriminate between such 10-ms or 25-ms stimuli (Extended Data Fig. 1d, f, h). Key high-frequency information in the odour profile might, therefore, be preserved in the inputs to the OB.

[1]Sensory Circuits and Neurotechnology Laboratory, The Francis Crick Institute, London, UK. [2]Department of Neuroscience, Physiology & Pharmacology, University College London, London, UK. [3]These authors contributed equally: Tobias Ackels, Andrew Erskine, Debanjan Dasgupta. ✉e-mail: andreas.schaefer@crick.ac.uk
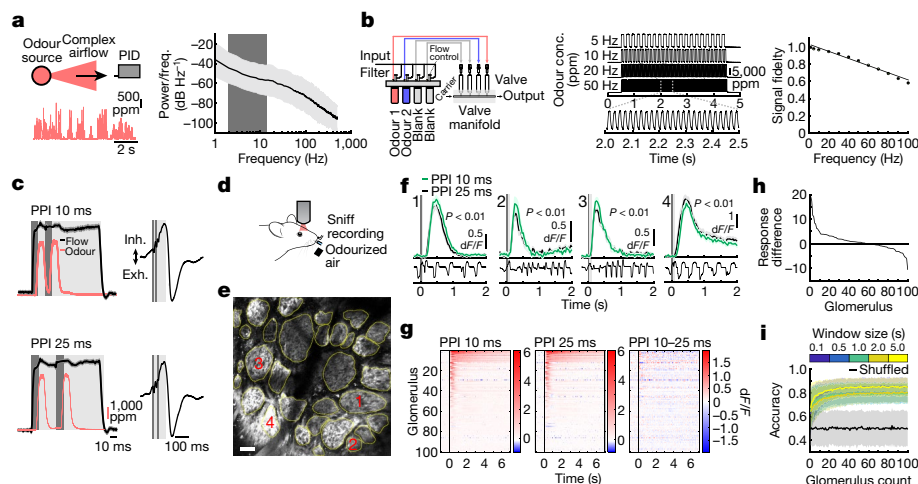
Fig. 1 | Sub-sniff detection of odour signals in olfactory bulb inputs. a, Left, example odour plume recorded outdoors under natural, complex airflow conditions using a photoionization detector (PID). Right, averaged power spectrum of all recorded odour plumes ($n = 37$ plumes, mean ± s.d. of log power); dark grey, typical range of mouse sniff frequencies. **b**, Left, multi-channel high-bandwidth odour delivery device. Middle, representative odour pulse recordings at command frequencies between 5 and 50 Hz. ppm, parts per million. Right, relationship of frequency and odour pulse signal fidelity (Methods; $n = 5$ repeats for each frequency, mean ± s.e.m., Supplementary Fig. 1). **c**, Left, odour and flow traces of stimuli with paired pulse interval (PPI) of 10 ms (top) and 25 ms (bottom); dark grey shading, valve commands. Right, stimuli are presented during the inhalation phase of the respiration cycle. **d**, The two-photon imaging approach. **e**, GCaMP6f fluorescence recorded in OB glomeruli (maximum projection of 8,200 frames, numbered glomeruli correspond to example traces in **f**). Scale bar, 50 μm. **f**, Top, example calcium traces in response to odour stimuli with PPIs of 10 or 25 ms (mean of ten trials ± s.e.m., unpaired two-sided t-test for 2-s response integral from odour onset). Bottom, example respiration trace. **g**, Calcium transients as colour maps for odour stimuli with PPI 10 ms or 25 ms, and the difference between the two. Glomeruli are sorted by response magnitude to the PPI 10-ms stimulus. **h**, Glomerular responses sorted by magnitude of difference in response to PPI 10 ms and PPI 25 ms. **i**, Classifier accuracy over all glomeruli when a linear classifier was trained on several response windows (colours; black, shuffled control) to stimuli with PPI 10 ms versus 25 ms ($n =$ up to 100 glomeruli from 5 mice; mean ± s.d. of 500 repetitions). Throughout, ethyl butyrate was used as the odour stimulus.

To test this experimentally, we performed Ca^{2+} imaging experiments in anaesthetized and awake mice expressing the calcium indicator GCaMP6f in OSNs (Fig. 1c–i, Extended Data Fig. 2) while we delivered odour pulses locked to inhalation (Fig. 1c, d). Overall, responses for all glomeruli were highly correlated between the two stimuli (Fig. 1f, g). Glomerular activity did not directly follow the 10-ms or 25-ms pulses (Fig. 1f). However, in one-third of glomeruli ($n = 33$ of 100, $P < 0.01$), responses were consistently and significantly different for the two stimuli (Fig. 1f–h, Extended Data Fig. 2), mirroring the simulation results (Extended Data Fig. 1). Notably, just a few dozen randomly chosen glomeruli were sufficient to discern between the stimuli at more than 80% success rate with a linear classifier (Fig. 1i, Extended Data Fig. 1h). Expanding the stimulus set to different concentrations and multiple pulses (Extended Data Fig. 2) confirmed that information about concentration and temporal patterns with features exceeding the 25-ms timescale is reliably and independently preserved in the population of OSNs.

Discrimination of correlation structure

To investigate whether mice can base their behaviour on such high-frequency stimuli, we trained mice in an automated go/no-go operant conditioning system (AutonoMouse[24]) (Fig. 2a, Supplementary Video 1) to discriminate between high-frequency stimuli. To ensure that the brief odour pulses were delivered during inhalation in freely moving mice, we opted for 2-s pulse trains at different frequencies with constant airflow (Fig. 2b). Mice could discriminate whether an odour was presented at, for example, 4 Hz or 20 Hz, but the apparent 'critical flicker frequency' (Fig. 2c, Extended Data Fig. 3) was significantly lower than frequencies that are readily represented by OSNs (Fig. 1, Extended Data Figs. 1, 2). However, in both visual and auditory systems, conventional flicker fusion frequency or gap detection thresholds substantially underestimate the temporal sensitivity, particularly for

tasks with multiple stimuli present[23,25,26]: In vision, for example, flicker fusion frequency is around 60 Hz, whereas thresholds for detecting synchrony between stimuli have been reported to be 3 ms[26]. Therefore, we wanted to probe whether olfactory tasks involving multiple odours could also reveal behavioural access to higher frequencies. We presented stimuli composed of two odours that fluctuated in a correlated or anti-correlated manner as the rewarded and unrewarded stimulus, respectively (and vice versa) (Fig. 2d–f). Mice readily learned to respond differentially to correlated or anti-correlated odours (Fig. 2h–k). A gradual increase in the correlation frequency showed that mice could reliably detect the correlation structure of stimuli at frequencies of up to 40 Hz (Fig. 2h, j, k). As a population, performance decreased by approximately 5% per octave, with performance significantly above chance at frequencies of up to 40 Hz ($n = 33$ mice in two cohorts of 14 and 19 mice) (Fig. 2k). To mitigate the risk of mice using unintended cues for discrimination, odours were presented from changing valve combinations (Fig. 2g, Extended Data Fig. 4), and odour flow was carefully calibrated (Fig. 2e, Extended Data Fig. 4d, e) and varied randomly between trials so that neither flow, valve clicking noises nor average concentration provided any information about the nature of the stimulus (Extended Data Fig. 4d–h). Consistent with this, when valve identities were scrambled, mice performed at chance (Fig. 2k). Finally, when odour presentation was changed to a new set of valves, performance levels were maintained (Fig. 2g–i, Extended Data Fig. 4i–k), indicating that only intended cues (the temporal structures of odours) were used for discrimination. Performance was independent of the odour pair used (Extended Data Fig. 3g) and was maintained for tasks in which the mice had to distinguish correlated from uncorrelated (rather than anti-correlated) odours (Extended Data Fig. 3e, f).

Mice tended to take more time to detect the correlation structure of stimuli with higher fluctuation frequencies (Extended Data Fig. 5j–l). This was most pronounced for mice with higher overall performance (Extended Data Fig. 5j). Accuracy strongly correlated with reaction time
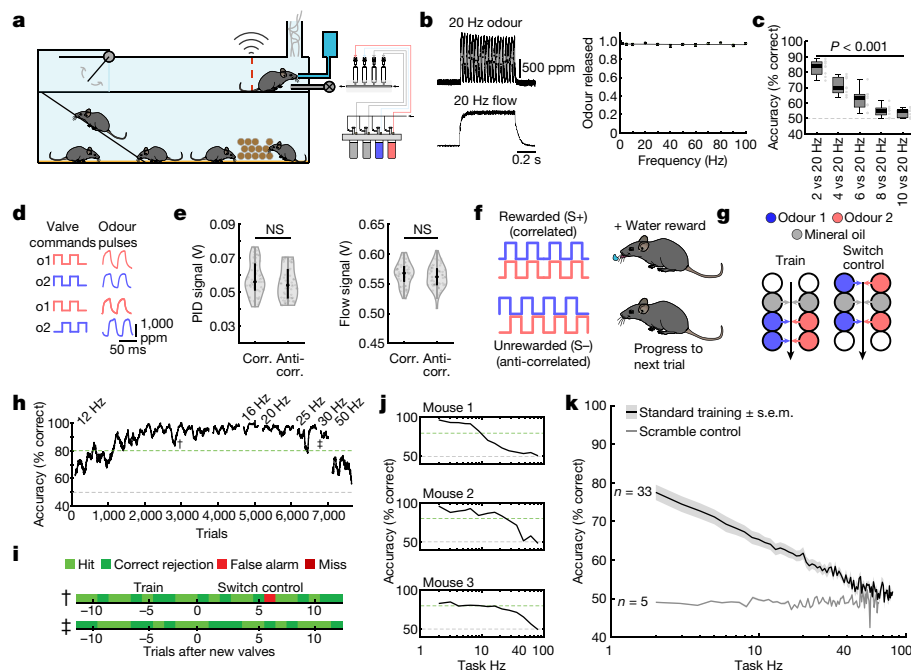
Fig. 2 | Mice can discriminate odour correlation structure at frequencies up to 40 Hz. a, The automated operant conditioning system (AutonoMouse) housing cohorts of up to 25 animals. **b**, Left, representative trace of a 20-Hz odour pulse train (top) and corresponding stable airflow (bottom). Right, relationship between frequency and total amount of odour released ($n = 5$ repeats for each frequency, mean ± s.e.m.). **c**, Group accuracy in frequency discrimination task ($n = 10$ mice, $P < 0.001$ for all stimuli compared to chance (paired two-sided t-test); Extended Data Fig. 3). Thick lines, medians; boxes, 25th–75th percentiles; whiskers, most extreme data points not considered outliers (Methods). **d**, Left, valve commands to release two odours fluctuating at 20 Hz in a correlated (top) or anti-correlated (bottom) manner. Right, resulting odour concentration changes measured using dual-energy PIDs (Supplementary Fig. 2). **e**, Odour (left) and flow (right) signals for correlated and anti-correlated stimuli fluctuating at 20 Hz ($n = 60$ trials for each condition; odour, $P = 0.19$; flow, $P = 0.23$; unpaired two-sided t-test). Black dots, medians;

black bar, second and third quartiles. **f**, Discrimination stimuli; mice were trained to discriminate between two odours presented simultaneously in either a correlated (top) or anti-correlated (bottom) fashion in a standard go/no-go paradigm. **g**, Valve combinations for stimulus production. Train: six valves are used to produce the stimulus through varying valve combinations. Switch control: two extra valves are introduced and odour presentation is switched over to the newly introduced valves. **h**, Data for example mouse performing the correlation discrimination task at different frequencies. †, ‡, time of new valve introduction as in **i**. **i**, Trial response maps before and after switch to control valves ($n = 12$ trials before and 12 trials after new valve introduction; Extended Data Fig. 4). **j**, Accuracy of three representative mice when stimulus pulse frequency is randomized from trial to trial. **k**, Group accuracy for the experiment in **j** ($n = 33$ training mice, $n = 5$ control mice, $n = 9.3 \times 10^5$ trials). Throughout, isoamyl acetate and ethyl butyrate were used as odour stimuli.

across all stimuli and mice (Extended Data Fig. 5k), despite the fact that total time of odour delivery was the same across all trials regardless of stimulus frequency. Consequently, when analysis was restricted to trials in which mice sampled the stimuli for long enough—for example, for at least 750 ms—performance significantly increased across frequencies (Extended Data Fig. 5l). This indicates that the measured performance might not be the psychophysical limit for discrimination of fluctuating odour stimuli. Furthermore, this suggests that mice integrate information across large portions of the presented stimuli, rather than, for example, detecting simultaneity of odour onset[14] to determine whether odours were correlated or not. To directly test this possible strategy, we interleaved training trials with probe trials in which the onset characteristics were flipped (Extended Data Fig. 5f–i). Notably, performance did not drop substantially (Extended Data Fig. 5h, i), consistent with a strategy that relies primarily on discerning the high-frequency correlation structure of the stimulus over several hundreds of milliseconds, rather than the onset only (Extended Data Fig. 5f, g, i). Sniff rate, in turn, was independent of the correlation frequency of stimuli presented (Extended Data Fig. 5a–e).

Odour correlation encoded in OB output

To assess how this high-frequency information is represented and reformatted in the olfactory system, we imaged neural activity in response to high-frequency stimuli (Fig. 3). Ca^{2+} imaging of OSN

responses to correlated and anti-correlated stimuli showed that—unlike for two pulses with variable gaps (Fig. 1)—the correlation structure of odour pulse trains was difficult to discern on the level of inputs to the OB using simple linear classifiers (Extended Data Fig. 6). Directly imaging from the output neurons of the OB—mitral and tufted cells (M/TCs) (Fig. 3a–g, Extended Data Fig. 7)—showed that overall, M/TCs also responded similarly to correlated and anti-correlated stimuli (Extended Data Fig. 7j–l). However, 17% of all M/TCs showed significantly different integral responses (0–5 s after odour onset, $P < 0.01$) to the two stimuli (114 of 680 regions of interest (ROIs)) (Fig. 3d–f). As a result, correlated and anti-correlated odours were reliably discriminated by a linear classifier using the M/TC population responses (Fig. 3g (somatic response), Extended Data Fig. 7d, i (dendritic response)), in contrast to the OSN population response (Extended Data Fig. 6k, l). This finding is consistent with the idea that the OB circuitry implements a nonlinear transformation of OSN input such that the representation of correlation becomes more readily accessible in the OB output.

We used odour stimuli that fluctuated rapidly at frequencies that substantially exceeded the temporal resolution of Ca^{2+} imaging, which captures a low-pass filtered signal of neural activity. Although the Ca^{2+} signal does not follow individual stimulus frequencies, the M/TC population response contained enough information to determine whether a correlated or anti-correlated stimulus was presented. To investigate whether additional information about stimuli is present in the output of the OB at finer time scales, we turned to extracellular unit recordings
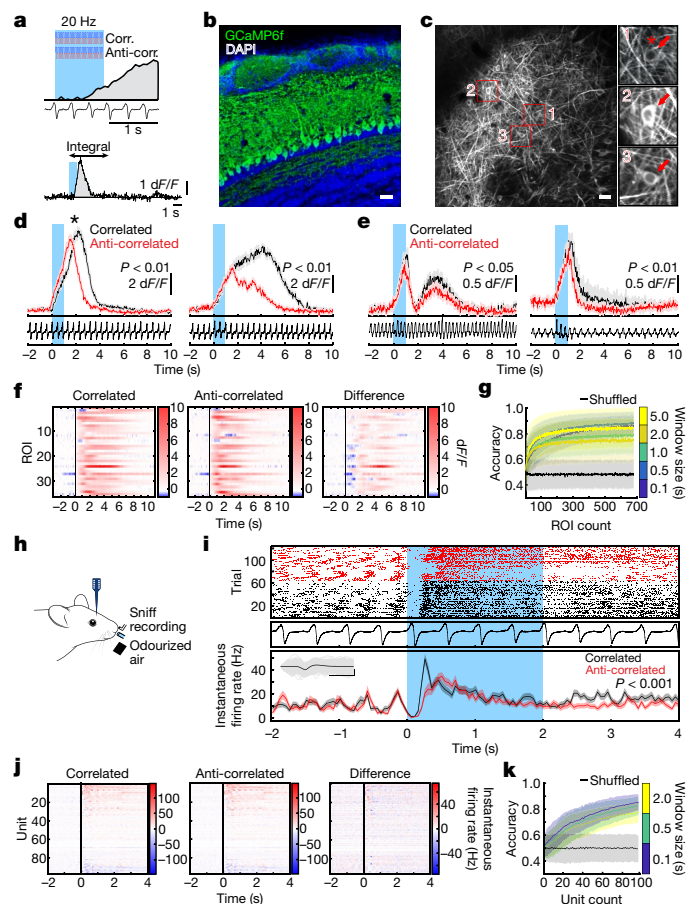
Fig. 3 | Odour correlation structure is encoded by OB output neurons.
a, The two-photon imaging approach (Extended Data Fig. 7e). **b**, Coronal OB section showing GCaMP6f (green) expressed in projection neurons. Scale bar, 20 μm. **c**, GCaMP6f fluorescence from mitral and tufted cells (maximum projection of 8,000 frames). Responses from ROI marked with asterisk in inset 1 are shown in **d**. Scale bar, 20 μm. **d**, **e**, Example traces of ROIs that show differential response kinetics to correlated (black) and anti-correlated (red) 20-Hz stimulation in anaesthetized (**d**; mean ± s.e.m. of 24 trials, unpaired two-sided t-tests on 5-s response integrals) and awake mice (**e**; mean ± s.e.m. of 16 trials, unpaired two-sided t-tests). Light blue, odour presentation. **f**, Calcium transients as colour maps for correlated and anti-correlated averaged trials and for the difference between the two for the 5% of ROIs with the largest differential responses. **g**, Accuracy of linear classifier trained on several response windows (colours; black, shuffle control) to correlated versus anti-correlated stimuli at 20 Hz (n = up to 680 ROIs from 6 mice; mean ± s.d. of 500 repetitions). **h**, The extracellular recording approach. **i**, Example single unit response to correlated and anti-correlated stimuli shown as raster plot (top) and peristimulus time histogram (PSTH) (mean ± s.e.m.) of spike times binned every 50 ms (bottom); inset, average spike waveform (black) and 1,000 individual spike events (grey). Scale bars, 100 μV, 1 ms. Light blue, odour presentation. Two-sided Mann–Whitney U test comparing spike time distributions of correlated and anti-correlated trials during 4 s after odour onset. **j**, Binned spike discharge over time shown as colour maps for all units (correlated, anti-correlated and the difference between the two). **k**, Accuracy of linear classifier trained on the average 2-s response to correlated versus anti-correlated stimuli at 20 Hz (yellow) (n = up to 97 units from 6 mice; mean ± s.d. of 1,000 classifier repetitions; Methods, Extended Data Fig. 8).

(Fig. 3h–k, Extended Data Fig. 8) and whole-cell patch-clamp recordings (Extended Data Fig. 9). Despite the kilohertz temporal resolution, single units also did not directly follow high-frequency stimuli. Average activity (summed spike count during 500 ms after odour onset) was, however, significantly different between correlated and anti-correlated stimuli in 24% of single units (23 of 97; $P < 0.01$, Mann–Whitney U test)

(Fig. 3i, j, Extended Data Fig. 8b), consistent with the Ca^{2+} imaging results. As few as 60 randomly selected units were sufficient to classify the odour stimuli with more than 80% accuracy (Fig. 3k). Additional information was contained at finer time scales, as increasing the temporal resolution of analysis improved discriminability (Fig. 3k, Extended Data Fig. 8e–g). Together, these results demonstrate that information about high-frequency correlation structure in odours is accessible to animals for behavioural decisions and readily available in the output of the OB.

Correlations allow source separation

We next considered what the detection of high-frequency correlations could be useful for. Natural odours consist of multiple different types of molecules, and a typical olfactory scene contains several sources[6]. To make sense of the olfactory environment, the brain must be able to separate odour sources, attributing the various chemicals present to the same or different objects[5]. Motivated by the turbulent nature of odour transport, Hopfield suggested that the temporal structure of odour concentration fluctuations might contain information about the locations of odour sources[5]—that is, that chemicals belonging to the same source would co-fluctuate in concentration. Detecting correlations in odour fluctuations would thus allow mice to discern which odours arise from the same object. To experimentally probe the potential of odour correlation structure to facilitate odour source separation in air, we devised a dual-energy fast photoionization detection method to simultaneously measure the odour concentrations of two odours with high temporal bandwidth (Methods; Fig. 4a, b, Extended Data Fig. 10a–e, Supplementary Fig. 2). When an odour was presented in a laboratory environment with artificially generated complex airflow patterns (Fig. 4a), to mimic the outdoor measurements (Fig. 1a), odour concentration fluctuated with a spectrum extending beyond 40 Hz (Extended Data Fig. 10a). When two odours were presented from the same source, these fluctuations were highly correlated (Fig. 4a, b, Extended Data Fig. 10b). When we separated odour sources and presented the two odours 50 cm apart, odour dynamics were uncorrelated (Fig. 4a, b) with intermediate correlations for closer distances (Fig. 4b). This pattern of almost perfect correlation for the same source and virtually uncorrelated dynamics for separated sources was maintained at closer and farther distances between odour source and sensor (Extended Data Fig. 10d), independent of the odours used (Extended Data Fig. 10c), and was mirrored outdoors (Extended Data Fig. 10e). Thus, the correlation structure of odorant concentration fluctuations contains reliable information about odour objects—for example, whether odours emerge from the same or different sources.

To investigate whether mice can make use of this information, we trained a new cohort of mice in a modified AutonoMouse setting, presenting odours that corresponded to the 'same source' or 'source separated' cases as rewarded or unrewarded stimuli (Fig. 4c, d, Extended Data Fig. 10). Mice could learn to discriminate these stimuli (Fig. 4d, e). Once the mice had acquired the task, we probed their performance with artificially generated stimuli (Extended Data Fig. 10f–k) that were derived from previous measurements with natural airflow but were perfectly correlated (Fig. 4e). Notably, the mice reliably responded to these probe trials with correlated stimuli as they did to the 'same source' stimuli they had been trained on (Fig. 4e, Extended Data Fig. 10m). To further ascertain that the mice were using the correlation structure to make these decisions, we probed them with artificial square pulse stimuli (as in Fig. 2, 3) at different frequencies. Mice performed significantly above chance in probe trials at frequencies of up to 40 Hz (Fig. 4e, Extended Data Fig. 10s), which implies that learning about source separation stimuli directly translates to distinguishing temporal features in correlated or uncorrelated stimuli.
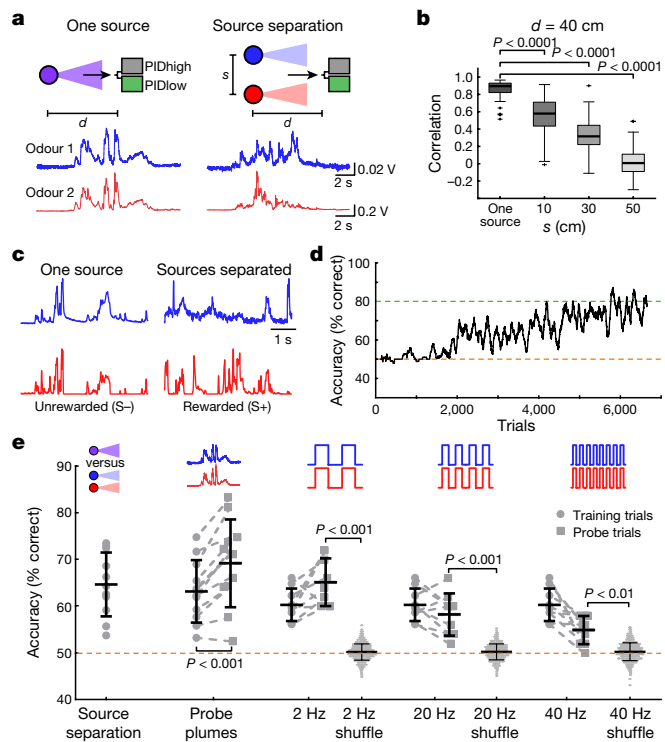
Fig. 4 | Source separation using correlations of odour concentration fluctuations. a, Simultaneous measurement of two odours (odour 1: α-terpinene (AT); odour 2, ethyl butyrate (EB)) using a dual-energy PID (Extended Data Fig. 10a–e, Supplementary Fig. 2) at $d = 40$ cm, presented from one source or from two sources separated by $s = 50$ cm, with complex airflow in the laboratory. **b**, Correlation coefficients over all recordings for odours from the same source or from two sources separated by $s = 10$–50 cm (EB versus AT; $n = 61$ for one source, $n = 71$ for each individual distance; unpaired two-sided t-test). Thick lines, medians; boxes, 25th–75th percentiles; whiskers, most extreme data points not considered outliers (Methods). **c**, Example plumes used for training mice on a virtual source separation task to discriminate between odour stimuli derived from the same source (unrewarded, S−) and from separated sources (rewarded, S+). **d**, Example learning curve for a mouse trained to perform the virtual source separation task. Isoamyl acetate and ethyl butyrate were used as odour stimuli. **e**, Average accuracy over different variants of the task, calculated over the last 2,400 trials of virtual source separation training ($n = 11$ mice, $P < 0.0001$, unpaired t-test, compared to chance performance), and subsequent stages in which probe trials containing novel plume types were interleaved with the training set. Responses are compared between probe and training trials within each stage. Probe plumes: odours fluctuate in a perfectly correlated manner, with a novel temporal structure (120 probe trials in a segment of 2,400 trials, $n = 11$ mice, paired t-test). Probe 2 Hz, 20 Hz, 40 Hz: correlated or anti-correlated square pulse trains (50 probe trials per frequency in a segment of 1,650 trials, $n = 9$ mice). Responses to 2 Hz, 20 Hz and 40 Hz probe trials were shuffled 10,000 times to calculate chance performance; data are mean ± s.d.; unpaired two-sided t-test.

Discussion

We have shown that the mammalian olfactory system has access to temporal features of odour stimuli at frequencies of at least up to 40 Hz. We have demonstrated that mice have access to information in rapid odour fluctuations using different behavioural experiments (Figs. 2, 4). We have shown reliable decoding from imaging and unit recordings from different stages of the olfactory system using both correlated odour concentration fluctuations (Fig. 3, Extended Data Figs. 7, 8a–g) and simplistic paired pulse stimuli with gaps as small as 25 ms (Fig. 1, Extended Data Figs. 2, 8h–l), corroborated by computational modelling (Extended Data Fig. 1). Our results are consistent with previous findings that the olfactory bulb circuitry not only enables

highly precise odour responses[17,18] but also enables the detection of optogenetically evoked inputs with a precision of 10–30 ms[27–29], with different projection neurons displaying distinct firing patterns in response to optogenetic stimulation[29]. Although behavioural and physiological responses to precisely timed odour stimuli have been observed in insects[13,15,30], in mammals the complex shape of the nasal cavity was generally thought to low-pass filter any temporal structure of the incoming odour plume. Our results show that while the low-pass filtering in the nose and by OSNs might reduce the ability of neurons to directly follow high-frequency stimuli, sufficient information about high-frequency content is preserved and available that mice can readily make use of this information.

We considered what such high bandwidth could be useful for. We have shown that odour sources even in close proximity differ in their temporal correlation structure. Thus, the ability to detect whether odorants are temporally correlated could allow mice to perform source separation, solving the 'olfactory cocktail party problem'[5,6] without prior knowledge about the odour scenery. We show that mice can indeed discriminate between stimuli derived from a single source or separated sources. They readily translate this discrimination to artificial, correlated pulse trains, thereby demonstrating that they are using correlation structure to make this distinction. Distinguishing between other environmental features, such as the distance or direction of an odour source, could also be achieved by extracting temporal features from odour fluctuations[1–3,8] possibly in combination with strategies that compare information reaching the brain through the two nares[31,32].

We also considered how exactly this temporal information is extracted. Although insects can detect the simultaneity of onset of two odours[14,33,34], this strategy is unlikely to be the dominant means that mice use to detect correlation (Extended Data Fig. 5). Similarly, mice do not show adjustment of sniff strategies to discriminate high-frequency odour correlations (Extended Data Fig. 5, Supplementary Video 2). Although individual mammalian OSNs are thought to be quite slow and unreliable[20], the large convergence of OSN axons provides a substrate to create the needed high temporal bandwidth[35] (Extended Data Fig. 1). Biophysical heterogeneity of OSNs might improve how the population encodes temporally structured stimuli[36,37]. Intrinsic cellular biophysics (Extended Data Fig. 9), local interneurons or long-range lateral inhibition[5,38] might permit the extraction of temporal correlation within the olfactory bulb circuitry and possibly result in tuning of individual projection neurons to specific temporal structures.

The turbulence of odour plumes has often been viewed as a source of noise for mammals. By contrast, we find that the mouse olfactory system has access to high-frequency temporal features in odour stimuli. This opens up a new perspective on how mice could make use of natural turbulence to obtain information about their spatial environment. This in turn provides new computational challenges for the mammalian olfactory system and an entry point into how information about space is extracted from sensory inputs.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03514-2.

1. Fackrell, J. & Robins, A. Concentration fluctuations and fluxes in plumes from point sources in a turbulent boundary layer. *J. Fluid Mech.* **117**, 1–26 (1982).
2. Mylne, K. R. & Mason, P. J. Concentration fluctuation measurements in a dispersing plume at a range of up to 1000 m. *Q. J. R. Meteorol. Soc.* **117**, 177–206 (1991).
3. Schmuker, M., Bahr, V. & Huerta, R. Exploiting plume structure to decode gas source distance using metal-oxide gas sensors. *Sens. Actuators B Chem.* **235**, 636–646 (2016).
4. Murlis, J., Elkington, J. S. & Carde, R. T. Odor plumes and how insects use them. *Annu. Rev. Entomol.* **37**, 505–532 (1992).

5. Hopfield, J. J. Olfactory computation and object perception. *Proc. Natl Acad. Sci. USA* **88**, 6462–6466 (1991).
6. Rokni, D., Hemmelder, V., Kapoor, V. & Murthy, V. N. An olfactory cocktail party: figure-ground segregation of odorants in rodents. *Nat. Neurosci*. **17**, 1225–1232 (2014).
7. Vergassola, M., Villermaux, E. & Shraiman, B. I. 'Infotaxis' as a strategy for searching without gradients. *Nature* **445**, 406–409 (2007).
8. Celani, A., Villermaux, E. & Vergassola, M. Odor landscapes in turbulent environments. *Phys. Rev. X* **4**, 041015 (2014).
9. Crimaldi, J. P. & Koseff, J. R. High-resolution measurements of the spatial and temporal scalar structure of a turbulent plume. *Exp. Fluids* **31**, 90–102 (2001).
10. Moore, P. A. & Atema, J. Spatial information in the three-dimensional fine structure of an aquatic odor plume. *Biol. Bull*. **181**, 408–418 (1991).
11. Mafra-Neto, A. & Cardé, R. T. Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. *Nature* **369**, 142–144 (1994).
12. Vickers, N. J. Mechanisms of animal navigation in odor plumes. *Biol. Bull*. **198**, 203–212 (2000).
13. Riffell, J. A. et al. Flower discrimination by pollinators in a dynamic chemical environment. *Science* **344**, 1515–1518 (2014).
14. Szyszka, P., Stierle, J. S., Biergans, S. & Galizia, C. G. The speed of smell: odor-object segregation within milliseconds. *PLoS ONE* **7**, e36096 (2012).
15. Szyszka, P., Gerkin, R. C., Galizia, C. G. & Smith, B. H. High-speed odor transduction and pulse tracking by insect olfactory receptor neurons. *Proc. Natl Acad. Sci. USA* **111**, 16925–16930 (2014).
16. Kepecs, A., Uchida, N. & Mainen, Z. F. The sniff as a unit of olfactory processing. *Chem. Senses* **31**, 167–179 (2006).
17. Shusterman, R., Smear, M. C., Koulakov, A. A. & Rinberg, D. Precise olfactory responses tile the sniff cycle. *Nat. Neurosci*. **14**, 1039–1044 (2011).
18. Cury, K. M. & Uchida, N. Robust odor coding via inhalation-coupled transient activity in the mammalian olfactory bulb. *Neuron* **68**, 570–585 (2010).
19. Burton, S. D. Inhibitory circuits of the mammalian main olfactory bulb. *J. Neurophysiol*. **118**, 2034–2051 (2017).
20. Duchamp-Viret, P., Chaput, M. A. & Duchamp, A. Odor response properties of rat olfactory receptor neurons. *Science* **284**, 2171–2174 (1999).
21. Munger, S. D., Leinders-Zufall, T. & Zufall, F. Subsystem organization of the mammalian sense of smell. *Annu. Rev. Physiol*. **71**, 115–140 (2009).
22. Bressel, O. C., Khan, M. & Mombaerts, P. Linear correlation between the number of olfactory sensory neurons expressing a given mouse odorant receptor gene and the total volume of the corresponding glomeruli in the olfactory bulb. *J. Comp. Neurol*. **524**, 199–209 (2016).
23. Carr, C. E. & Amagai, S. Processing of temporal information in the brain. *Adv. Psychol*. **115**, 27–52 (1996).
24. Erskine, A., Bus, T., Herb, J. T. & Schaefer, A. T. AutonoMouse: high throughput automated operant conditioning shows progressive behavioural impairment with graded olfactory bulb lesions. *PLoS ONE* **14**, e0211571 (2019).
25. Brown, J. L. Visual sensitivity. *Annu. Rev. Psychol*. **24**, 151–186 (1973).
26. Westheimer, G. & McKee, S. P. Perception of temporal order in adjacent visual stimuli. *Vision Res*. **17**, 887–892 (1977).
27. Smear, M., Shusterman, R., O'Connor, R., Bozza, T. & Rinberg, D. Perception of sniff phase in mouse olfaction. *Nature* **479**, 397–400 (2011).
28. Rebello, M. R. et al. Perception of odors linked to precise timing in the olfactory system. *PLoS Biol*. **12**, e1002021 (2014).
29. Li, A., Gire, D. H., Bozza, T. & Restrepo, D. Precise detection of direct glomerular input duration by the olfactory bulb. *J. Neurosci*. **34**, 16058–16064 (2014).
30. Geffen, M. N., Broome, B. M., Laurent, G. & Meister, M. Neural encoding of rapidly fluctuating odors. *Neuron* **61**, 570–586 (2009).
31. Rajan, R., Clement, J. P. & Bhalla, U. S. Rats smell in stereo. *Science* **311**, 666–670 (2006).
32. Catania, K. C. Stereo and serial sniffing guide navigation to an odour source in a mammal. *Nat. Commun*. **4**, 1441 (2013).
33. Baker, T., Fadamiro, H. & Cosse, A. Moth uses fine tuning for odour resolution. *Nature* **393**, 530 (1998).
34. Stierle, J. S., Galizia, C. G. & Szyszka, P. Millisecond stimulus onset-asynchrony enhances information about components in an odor mixture. *J. Neurosci*. **33**, 6060–6069 (2013).
35. Abeles, M. Time is precious. *Science* **304**, 523–524 (2004).
36. Padmanabhan, K. & Urban, N. N. Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nat. Neurosci*. **13**, 1276–1282 (2010).
37. Park, I. M., Bobkov, Y. V., Ache, B. W. & Príncipe, J. C. Intermittency coding in the primary olfactory system: a neural substrate for olfactory scene analysis. *J. Neurosci*. **34**, 941–952 (2014).
38. Fukunaga, I., Herb, J. T., Kollo, M., Boyden, E. S. & Schaefer, A. T. Independent control of gamma and theta activity by distinct interneuron networks in the olfactory bulb. *Nat. Neurosci*. **17**, 1208–1216 (2014).

Article

Methods

Ethical compliance

All animal procedures performed in this study were approved by the UK government (Home Office) and by the Crick Institutional Animal Welfare Ethical Review Panel.

Mice

All mice used for behavioural experiments were C57/Bl6 males, as long-term group housing precluded the use of mixed-sex cohorts (Figs. 2, 4, Extended Data Figs. 3–5, 10). In vivo imaging experiments were performed in 12–20-week-old heterozygous OMP-cre[39] (JAX stock no. 006668; Fig. 1, Extended Data Figs. 2, 6) or Tbet-cre[40] (Jax stock no. 024507; Fig. 3, Extended Data Fig. 7) mice crossed with the Ai95(RCL-GCaMP6f)-D line[41] (JAX stock no. 028865) of either sex. Extracellular unit (Fig. 3, Extended Data Fig. 8) and whole-cell patch-clamp recordings (Extended Data Fig. 9) were performed in 5–8-week-old C57/Bl6 males. Mice were housed up to 5 per cage under a 12–12 h light–dark cycle. Food and water were provided ad libitum.

Reagents

All odours were obtained in their pure form from Sigma-Aldrich. Unless otherwise specified, odours were diluted 1/5 with mineral oil in 15-ml glass vials (27160-U, Sigma-Aldrich).

Statistical analysis and data display

To test for statistical significance between groups where appropriate we used either paired or non-paired student t-tests or, for non-parametric data, the Mann–Whitney U test, or the Kolmogorov–Smirnov test to test the equality of probability distributions. Statistical test details and P values are provided in figures and/or legends. Unless specified otherwise, boxplots were plotted using the MATLAB boxplot function with the median depicted as a thick line and default maximal whisker length of $1.5 \times (q3 - q1)$ where q3 and q1 indicate the 75th and 25th percentile, respectively. If points were located outside this whisker range they were displayed individually as outliers. Violin plots show the median as a black dot and the second and third quartiles by the bounds of black bars. Mouse cartoons were adapted from https://scidraw.io/drawing/123 and https://scidraw.io/drawing/49.

High-speed odour delivery device

The odour delivery device was based on a modular design of four separate odour channels, and consisted of an odour manifold for odour storage, a valve manifold for control of odour release and hardware for controlling and directing airflow through the system (Fig. 1b). The odour manifold was a $12.2 \times 3.2 \times 1.5$-cm³ stainless steel block with four milled circular indentations (1-cm radius). Within each of these indentations was a threaded through-hole for installation of an input flow controller (AS1211F-M5-04, SMC) and an output filter (INMX0350000A, The Lee Company). For each inset, the cap of a 15-ml glass vial (27160-U, Sigma-Aldrich) with the centre removed was pushed in and sealed with epoxy resin (Araldite Rapid, Huntsman Advanced Materials). This meant that glass vials could be screwed in and out of the insets for rapid replacement.

Solenoid valves typically limit high-fidelity odour stimulation, resulting in odour rise times of several tens of milliseconds under optimal conditions[42]. We therefore used high-speed micro-dispense valves with custom electronics for pulse-width modulation to maximize bandwidth: four VHS valves (INKX0514750A, The Lee Company) were installed in a four-position manifold (INMA0601340B, The Lee Company) with standard mounting ports (IKTX0322170A, The Lee Company). Each valve was connected to a corresponding odour position in the odour manifold with 10-cm Teflon tubing (TUTC3216905L, The Lee Company). Each valve was controlled by digital commands via a spike-and-hold driver. Each digital pulse delivered to the spike-and-hold driver delivered a 0.5-ms, 24-V pulse to the valve (to open it), followed by a 3.3-V holding pulse lasting the rest of the duration of the digital pulse. This spike-and-hold input allowed fast cycling of the valve without switching between 0 and 24 V at high frequencies, to prevent overheating of the valve. Each valve was controlled by an individual spike-and-hold driver. Up to four drivers could be controlled and powered with a custom-made power supply unit consisting of a 24-V power input and a linear regulator to split the voltages into 24-V and 3.3-V lines, as well as control inputs that take digital signal input and route it to the appropriate valve. Pulse profiles for calibration and stimulus production were generated with custom Python software (PyPulse, PulseBoy; https://github.com/RoboDoig), allowing us to define pulse parameters across multiple valves using a graphical user interface.

To generate airflow through the olfactometer, a pressurised air source was connected to a demister (AME250C-F02, SMC) and filter (AMF250CF02, SMC) and then split into two separate lines, the input line and carrier line. Both lines were then connected to a pressure regulator (AR20-F01BG-8, SMC) and flow controller (FR2A13BVBN, Brooks Instrument). The main line was then connected to the input of the valve manifold. The input line was split into four separate lines and connected to the input flow controllers (set to 0.25 l/min) on each odour position of the odour manifold. The output of the valve manifold was fitted with MINSTAC tubing (TUTC3216905L, The Lee Company). Where the design was scaled up (for example, to include eight odour positions) the valve manifold outputs were connected and consolidated to a single output with three-way connectors (QSMY-6-4, Festo). The shape and reliability of odour pulses depended strongly on low volume headspace and low pressure levels (0.05 MPa). Flow change due to odour pulses was always compensated by mineral oil presentation (for example, light grey in Fig. 1c).

Odour characterization. Signal fidelities were calculated by first subtracting the average amplitude of troughs from the average amplitude of peaks during a pulse train and then subsequently dividing this peak-to-trough value by the difference in the average peak amplitude subtracted by baseline amplitude (SignalFidelity = (meanPeak − meanTrough)/(meanPeak − baseline)). This results in a value between 0 and 1, with 1 being perfectly modulated odour pulses.

Behaviour

Automated operant conditioning of cohorts of mice (Autono-Mouse). In AutonoMouse, groups of mice (up to 25) implanted with a radiofrequency identification (RFID) chip are housed in a common home cage[24] (Fig. 2a). Within the common home cage of AutonoMouse, mice have free access to food, social interaction and environmental enrichment. Water is not freely available in the system, but can be gained at any time by completion of an operant conditioning go/no-go task. To access these behavioural tasks, mice must leave the home cage and enter a behavioural area. This behavioural area contains the odour port and a lick port through which water rewards can be released. The lick port is also connected to a lick sensor, which registers the animal's response (its lick rate) in response to the task stimuli. As animals can gain their daily water intake only by completing behavioural tasks, mice are motivated to complete long sequences of trials without manual water restriction. Sample sizes for each cohort were determined by the availability of simultaneously weaned male mice to be group-housed. The minimum sample size for each cohort was determined by requiring at least three mice in each subgroup. Mice were randomly assigned to different test subgroups based on their performance during an initial simple odour discrimination task until performance levels between the subgroups were statistically indistinguishable by a one-way ANOVA. After this initial allocation, the experimenter was blinded as to which group mice belonged to since the only distinguishing feature between mice was their RFID chip code.

Training on temporally structured odours. We aimed to probe whether mice could perceive a particular temporal feature of naturally occurring odour signals: temporal correlations between odour signals. In particular, we aimed to investigate this question with the simplest possible case: whether mice could discriminate perfectly correlated from perfectly anti-correlated odour stimuli.

All tasks followed a standard go/no-go training paradigm. Mice were presented with two odours presented in either a correlated pattern or an anti-correlated pattern (Fig. 2d, Extended Data Fig. 4a–c). For roughly half of all mice, the correlated pattern was S+ (rewarded) and the anti-correlated pattern was S− (unrewarded); in the other half of the group this reward valence was reversed. All stimuli were 2 s long. A water reward could be gained by licking so that licking was detected for at least 10% of the stimulus time during an S+ presentation (a 'Hit'). Licking for the same amount of time during S− presentation resulted in a timeout interval of 7 s. In all other response cases, the inter-trial interval was 3 s and no water reward was delivered.

Stimulus structure. All anti-correlated and correlated stimuli on each trial followed a common pattern in their construction. Generally, wherever an odour position is inactivated a blank position should be activated to compensate for flow change. There should also be no consistent differences in the amount of odour or flow released during the stimulus between correlated and anti-correlated stimuli. The detailed algorithm for stimulus generation is as follows:

I. Correlated or anti-correlated/uncorrelated odour pulses (Fig. 2d, k, Extended Data Figs. 3, 5)

1. The stimulus is chosen to be correlated or anti-correlated/uncorrelated.

2. A set of 1–2 positions each for odour 1 and odour 2 and 2–3 positions for blank are randomly chosen from a pre-defined subset of 6 of the 8 total positions. For example, a valid combination could be odour 1 at position 1, 2; odour 2 at position 5; and blank at position 3 and 7 (Fig. 2g, Extended Data Fig. 4b).

3. A guide pulse is created at the desired frequency (for example, 2-Hz pulse with 50% duty; Supplementary Fig. 1c) for all positions that follows the chosen stimulus structure.

4. The relative contributions of each position to the total stimulus are randomly generated. At each time point in the stimulus, only two position types should be active (for example, odour 1 and blank for an anti-correlated stimulus) so the maximum contribution for any position type is 50% of the total release amount. Where two positions have been chosen for a position type, their relative contributions should add to 50% (Extended Data Fig. 4b).

5. The guide pulses are pulse-width modulated according to the relative contributions of each position (Supplementary Fig. 1c). Pulse-width modulation (PWM) is at 500 Hz with some added jitter in the duty to avoid strong tone generation.

6. For uncorrelated pulses, temporal offsets are added in one channel according to a distribution of time delays that follow the desired correlation structure between the two odour pulses (Extended Data Fig. 3e, f).

II. 'One source' and 'source separated' naturalistic plumes (Fig. 4c–e, Extended Data Fig. 10)

1. The stimulus is chosen to be 'one source' or 'source separated'.

2. A plume bank of plume pairs obtained from indoor PID recordings is created. Each trial will contain two plumes, each representing one odour recording originating from one source or from sources positioned 50 cm apart. To maintain consistency in trial length between behaviour experiments, a 2-s time window from each plume was selected from the middle of each 5-s recording, such that odour was always present in the first 500 ms of the trial. Trials where the correlation of the 2-s window was vastly different from the original 5 s were excluded from the plume bank. This procedure resulted in a plume bank containing 72 plume pairs for the separated source condition and 48 plume pairs for the one source condition.

3. An odour plume pair is randomly selected from the plume bank, from the corresponding category.

4. The odour that will be used to replicate each plume in a pair and the positions in the odour delivery device that will be used for that purpose are randomly assigned, as described previously. For each odour valve active, a blank valve will also be activated to produce an 'anti-plume' structure, to compensate for the changes in flow created by odour delivery.

5. Plumes are recreated from the chosen PID recordings. Each trace is normalized to between 0 and 1, and then converted into a series of binary opening and closing times. The length of the openings and closings relate directly to the value of the normalized signal: a value of one translates to a continuous opening, and a value of zero translates to continuously closed. This series of openings and closings are relayed to the valves and the resulting output resembles the original plume.

III. Perfectly correlated plume trials (probe trials, Fig. 4e)

1. A 2-s window is chosen from the source separated plume bank (plume structures previously associated with opposite reward valence; 5 trials) or from independent plume recordings obtained in a different environment than the original recordings (completely novel plume structures; 10 trials).

2. The chosen plume structure is replicated using both odour channels, resulting in a plume where both odour components fluctuate in a perfectly correlated manner.

IV. Frequency discrimination pulses (Fig. 2c, Extended Data Fig. 3a–d)

1. Two frequencies are chosen for discrimination (for example, 2 Hz versus 20 Hz).

2. For each trial one of the frequencies is chosen for presentation.

3. Valves are selected for presentation of both odours.

4. A guide pulse is created for each odour channel that pulses at the desired trial frequency with 50% duty, such that pulse alternates between channels at the given frequency.

5. Guide pulses are pulse-width-modulated as for correlated and anti-correlated stimuli.

Task structure for the correlation experiment. This is relevant to Fig. 2d–k, Extended Data Figs. 3–5. Task frequency was randomized from trial to trial in a range between 2 and 81 Hz. The choice of frequency was with weighted probability divided into three frequency bands. For example, this task could be arranged such that 2–20 Hz would be chosen with $P = 0.6$, 21–40 Hz with $P = 0.3$ and 41–81 Hz with $P = 0.1$. Within each of these frequency bands, the choice of individual task frequency was based on a uniform distribution. Thus, few trials were performed for frequencies exceeding 40 Hz, resulting in more 'noisy' behavioural performance data in Fig. 2k.

Onset detection. For the onset detection experiments (Extended Data Fig. 5f–i), mice were trained to discriminate perfectly correlated (for example, S+) from perfectly anti-correlated stimuli (for example, S−) and probed with partially altered stimuli where the onset (first cycle) of the probe S+ stimuli was anti-correlated and probe S− stimuli where the onset (first cycle) was correlated. Performance during these probe trials was then compared to the average performance during training ($perf_{train}$).

We calculated the expected average mouse performance on the probe trials based on two models (prediction data, Extended Data Fig. 5f, g). Model 1 assumed the mice were taking any part of the stimulus into account equally when making a decision. Model 2 assumed that only the onset of the stimulus would contribute to discrimination. Thus, for Model 1, a stimulus of frequency f (for example, 10 Hz) that was sampled for time t_{sample} consisted of a 'shifted' onset component of one cycle for S+ ($1/f$) and half a cycle for S− ($0.5/f$) corresponding to a fraction of $frac_{onset} = 1/f/t_{sample}$ of the entire stimulus and a 'normal' residual ($frac_{res} = 1 − frac_{onset}$). Thus, the predicted probe trial performance would be: predicted probe trial $perf_{entire} = perf_{train} \times frac_{res} + (1 − perf_{train}) \times frac_{onset}$.

Article

In Extended Data Fig. 5i, this prediction was calculated for the following parameters: sniff frequency, 6 Hz; inhalation fraction, 0.2; stimulus sampling time, 0.7 s. In Extended Data Fig. 5g, sampling time was varied as indicated.

For Model 2, ignoring inhalation timing, the prediction would be that preference would be reversed (as onset correlations during probe trials are reversed). However, this ignores the fact that odour stimuli during the exhalation period might not be detected. Thus, to more accurately predict the performance of mice for Model 2, we assume that the part of the stimulus that is detected as the 'onset' is the first odour pulse during an inhalation phase. During the probe trial, this will be the 'inverted' first cycle if the stimulus begins either during the inhalation phase or at most $1/f$ before the inhalation (then inhalation would start during the inverted first cycle of the probe trial). The probability of this occurring is $perf_{onset} = (dur_{inh} + 1/f)/dur_{sniff}$, with dur_{inh} and dur_{sniff} being the inhalation and sniff durations, respectively (provided $dur_{inh} + 1/f < dur_{sniff}$). Predicted probe trial performance for an 'onset only' model would therefore be: predicted probe trial $perf_{onset} = perf_{train} \times (1 - perf_{onset}) + (1 - perf_{train}) \times perf_{onset}$.

Extended Data Figure 5f, g displays the predictions of these two models in comparison to the experimental data for a broad range of respiration patterns. The 'prediction data' in Extended Data Fig. 5i show model predictions assuming typical sniff and sampling parameters as indicated above (sniff frequency, 6 Hz; inhalation fraction, 0.2; stimulus sampling time, 0.7 s).

Controls. Control valves could be automatically added to the random frequency task. These tasks produced their stimuli based on a subset of six valves and control valves could be added automatically after a set period of trials to force the algorithm to produce stimuli from all eight valves (Fig. 2g, i, Extended Data Fig. 4i–k).

A subgroup of mice was created in which the valve map was scrambled, as an ongoing control against mice learning extraneous variables in the task (Fig. 2k). The valve map was scrambled in the following way: one blank to odour 1, one odour 2 to blank, one odour 1 to odour 2 and one odour 1 to blank. Every few days all odour bottles were cleaned and replaced, odour positions changed and valves re-assigned[43].

Airflow and sound recordings. Airflow and sound were recorded in AutonoMouse during trials at different frequencies to ensure that the temporal structure of the odour was the only parameter that varied over trials and that no tactile or auditory cues were present in the stimulus. A flow sensor (AWM5101VN, Honeywell, USA) and a microphone (NTG1, RØDE) were placed in close proximity to the AutonoMouse odour port. In total, 286 trials were recorded (2 Hz: $n = 75$ correlated, $n = 70$ anti-correlated; 40 Hz: $n = 69$ correlated, $n = 72$ anti-correlated) using Audacity for sound and Spike2 (Cambridge Electronic Design) for flow signals. Airflow and sound signals underwent spectral analysis (Fourier transform), as well as linear classification analysis (Extended Data Fig. 4d–g).

Training on naturalistic plumes. One group of mice ($n = 12$) were trained to discriminate between plumes derived from stimuli originating from one source (S−, unrewarded) or from separated sources (S+, rewarded), using 2-s-long stimuli produced as described above from the recordings shown in Fig. 4a. An additional 12 mice trained simultaneously on the reverse reward valence did not pass the performance criterion within the given timeframe and were not carried forward to probe trials.

To test whether correlation structure was a feature used by mice to perform the virtual source separation task, probe trials were introduced randomly at a frequency of approximately 1 in 11 trials, with every instance of a probe trial repeated every 330 trials. Probe trials consisted of perfectly correlated plumes or correlated or uncorrelated square pulses produced as described above, presented at three frequencies:

2 Hz, 20 Hz, 40 Hz. The feedback for probe trials was the same as for a training trial, with a reward or time-out given based on the response of the mouse. No change in performance across repeated presentation of the probe trials was observed, indicating that performance was not due to putative rapid re-learning. Of the 12 mice exposed to this protocol, a total of 9 mice reached all phases of the experiment.

Cohorts. The correlation discrimination experiment was performed in three separate experimental cohorts (Fig. 2, Extended Data Figs. 3–5: group 1, $n = 14$; group 2, $n = 25$ (one mouse did not successfully pass the pre-training); Fig. 4, Extended Data Fig. 10: group 3, $n = 24$, see above). Each cohort was organized into several subgroups, which performed slight variations of the behavioural tasks in terms of reward valence and valves used, but with the same underlying task aim. Half of the mice in each subgroup were trained on correlated stimuli as the S+ rewarded condition, with the other half trained on anti-correlated stimuli as rewarded. Mice were further subdivided into groups, which were trained on different subsets of valves as standard in the eight-channel olfactometer. For each cohort, mice were once assigned to each of these subgroups based on performance in a simple pure odour discrimination at the beginning of the experiment—group membership was randomized until no significant (ANOVA, Tukey–Kramer) differences in performance could be extracted between these subgroups on this task.

Data analysis. AutonoMouse behavioural data were converted to MATLAB data format using the Conversion module of the Python autonomouse-control package (https://github.com/RoboDoig). All subsequent analysis was performed with custom-written MATLAB scripts unless otherwise specified.

All behavioural performance within a specified trial bin was calculated as a weighted average of S+ versus S− performance:

$$performance = \frac{(Hit/S+) + (CR/S-)}{2}$$

in which S+ is the total number of rewarded trials, S− is the total number of unrewarded trials, Hit is the total number of rewarded trials in which a lick response was detected, and CR (correct rejection) is the total number of unrewarded trials in which no lick response was detected.

For random stimulus pulse frequency experiments (for example, Fig. 2j, k), trials were binned approximately by half-octave for performance analysis. The exact intervals were f(Hz) = [2, 3, 4, 5, 6:7, 8:10, 11:13, 14:17, 18:22, 23:29, 30:37, 38:48, 49:62, 63:81]. Reaction time (Extended Data Fig. 5) was calculated from S+ trials for each mouse as the time to the first lick after stimulus onset. For presentation of learning curves (Figs. 2h, 4d) accuracy was calculated over 100-trial sliding windows.

Motion magnification of the respiration camera video recordings (Extended Data Fig. 5, Supplementary Video 2) was performed with phase-based video motion processing with correction for large body movements based on MATLAB scripts[44] (phaseAmplifyLargeMotions). Parameters for phase amplification were: blurring $\sigma = 1$, magnification $a = 50$, amplification in frequency band between 2 and 13 Hz. Following magnification, static ROIs for each video were selected in Bonsai[45] (http://www.kampff-lab.org/bonsai/) over the mouse flank. An adaptive binary threshold was applied to the ROI to segment the mouse body from the video background. Respiration rate was extracted from the total size of the ROI occupied by the body over time.

Olfactory sensory neuron population model

Overview. We modelled the OSN population as noisy integrate-and-fire neurons integrating a filtered odour pulse and with independent (cell-specific) noise to qualitatively match experimental data[46]. The square of the resulting mean population firing rate was convolved with a calcium imaging filter to produce a model of the observed calcium

imaging signal. All code and related data for the model can be found at https://github.com/stootoon/crick-osn-model-release.

Odour input current. The olfactory input current I_t to each OSN was modelled as a filtered version of the odour pulse input O_t with filter time constant τ_C (encompassing, for example, filtering of the nasal cavity as well as olfactory transduction):

$$\tau_C \frac{\mathrm{d}I_t}{\mathrm{d}t} = -I_t + O_t$$

This models filtering by the nasal cavity, transport through the mucous, and chemical transduction from odour concentration to receptor channel opening.

Olfactory sensory neurons. Each OSN was modelled as a noisy integrate-and-fire neuron. Each OSN membrane performs a noisy integration of the olfactory input current I_t so that the membrane voltage V_t satisfies the following stochastic differential equation: $\tau_V \mathrm{d}V_t = (I_t - V_t)\,\mathrm{d}t + \sigma \mathrm{d}B_t$. Here, B_t is standard Brownian motion and σ is the standard deviation of the membrane voltage noise. The OSN generates a spike whenever its membrane voltage exceeds a spiking threshold θ:

$$S_t = \begin{cases} 1 & \text{If } V_t \geq \theta; \\ 0 & \text{otherwise.} \end{cases}$$

Upon spiking, the membrane voltage is clamped to a refractory voltage V_{ref} for a period of t_{ref} seconds. The mean instantaneous firing rate of a population of N neurons is computed as

$$\overline{S}_t = \frac{1}{N} \sum_{n=1}^{N} S_t^n$$

in which S_t^n is the spiking activity of OSN n.

Calcium imaging signal. To model the calcium imaging signal, the mean firing rate is squared and convolved with the imaging kernel h_t to form the calcium imaging signal C_t: $C_t = (\overline{S}_t^2) * h_t$. The calcium imaging kernel is an alpha function: $h_t = t\, e^{-t/\tau_h}$.

A list of parameters is given in Supplementary Table 1. All parameters were fit manually: parameters τ_C, τ_V and σ were set to produce a qualitative match in time courses between model membrane voltage traces and the suction current traces in figure 2 of ref. [46]. The remaining parameters were adjusted to produce a qualitative fit between the model and the dynamics of the observed calcium imaging traces.

Generating model glomeruli
Overview. We generated 100 model glomeruli by randomly varying a subset of the model OSN parameters described in the previous section (τ_V, τ_C, σ, θ, a). Specifically, we picked the parameters of each glomerulus by selecting uniformly within $\pm 25\%$ of the centre value of each parameter. All 5,000 OSNs within each glomerulus had the same parameters, and differed only because of the random noise applied to their membrane voltages. The range of variation is shown in Supplementary Table 2.

The effect of concentration was modelled by linear scaling of the input waveforms. For each setting of PPI and concentration, the model was run to simulate 25 consecutive trials of length 2.5 s each, with the odour onset at 0.1 s into each trial. The first five trials of data were discarded to allow the model to 'settle', yielding 20 trials for each condition that were used in subsequent analyses (Extended Data Figs. 1g, h, 2j, k).

Classifying glomerular outputs
Predictors. The predictors used for classification were the response integrals for each glomerulus, defined as the instantaneous mean firing rate of the OSNs in the glomerulus, filtered by the Ca^{2+} imaging filter (see section 'Olfactory sensory neuron population model') and summed over the 2 s following odour onset in each trial. As the scale of the responses is arbitrary, we scaled the response integrals by their overall standard deviation, computed over glomeruli and trials.

Labels. Trials were labelled by their PPI, or a combination of PPI and concentration, depending on the task.

Classifiers. The classifiers used were support vector machines with linear kernels and l2 regularization as implemented by the LinearSVC function of the Python scikit-learn library. The setting of the penalty parameter C and whether or not to learn an intercept were determined by cross-validation with scikit-learn's GridSearchCV. The values of C considered ranged in powers of 10 from 10^{-4} to 10^4.

Computing decoding accuracy. The decoding accuracy for a given subset of n glomeruli was computed as the average accuracy over 10 cross-validation trials for the results in Extended Data Fig. 1h, and 40 cross-validation trials for the results in Extended Data Fig. 2j, k. In each cross-validation trial, the classifier was trained on a random 90% of the trials and tested on the remaining 10%, and the accuracy recorded. The random subsets were stratified; that is, constrained to have the same fraction of trials from each class as the full dataset when possible. The mean accuracy across cross-validation trials was recorded as the accuracy for that subset. To compute the shuffled performance, the labels of the training and test trials were shuffled in each cross-validation trial before the classifier accuracy was computed.

Decoding PPI from the responses of model glomeruli. To determine how decoding accuracy was affected by the size of the population used we selected a random subset of n glomeruli and computed the decoding accuracy as described above. This was repeated for 256 random subsets of n glomeruli, generating 256 unshuffled and 256 shuffled accuracies. The subset size n was varied from 1 (using only a single glomerulus) to 100 (using the full population; Extended Data Fig. 1h). For $n \geq 99$, some subsets are likely to have been repeated because there are fewer than 256 possible subsets of size 99 and 100. The observed variability in accuracy in those cases is then due mainly to the random determination of training and testing trials.

Decoding PPI and concentration from the responses of model glomeruli. To compute the decoding accuracy when decoding PPI and concentration, we followed a very similar procedure to the previous section, but fixed the population size at the maximum of 100 and varied the stimulus concentration from 0.5 to 5 in steps of 0.5 to cover a factor of 10 range in concentration as used in the experimental data (Extended Data Fig. 2j–m). For comparison, the results in Extended Data Fig. 1h were for a concentration of 1. Decoders were trained to extract just concentration, or PPI and concentration.

In vivo two-photon imaging
Surgical and experimental procedures. Before surgery, all used surfaces and apparatus were sterilized with 1% trigene. Mice were anaesthetized using a mixture of fentanyl, midazolam and medetomidine (0.05 mg/kg, 5 mg/kg and 0.5 mg/kg, respectively). Depth of anaesthesia was monitored throughout the procedure by testing the toe-pinch reflex. The fur over the skull and at the base of the neck was shaved away and the skin cleaned with 1% chlorhexidine scrub. Mice were then placed on a thermoregulator (DC Temperature Controller, FHC) heat pad controlled by a temperature probe inserted rectally. While on the heat pad, the head of the mouse was held in place with a set of ear bars. The scalp was incised and pulled away from the skull with four arterial clamps at each corner of the incision. A custom head-fixation implant was attached to the base of the skull with medical super glue (Vetbond, 3M)

Article

such that its most anterior point rested approximately 0.5 mm posterior to the bregma line. Dental cement (Paladur, Heraeus Kulzer; Simplex Rapid Liquid, Associated Dental Products) was then applied around the edges of the implant to ensure firm adhesion to the skull. A craniotomy over the left OB (approximately 2 × 2 mm) was made with a dental drill (Success 40, Osada) and then immersed in artificial cerebrospinal fluid (ACSF: NaCl (125 mM), KCl (5 mM), HEPES (10 mM), pH adjusted to 7.4 with NaOH, MgSO$_4$.7H$_2$O (2 mM), CaCl$_2$.2H$_2$O (2 mM), glucose (10 mM)) before the skull was removed with forceps. The dura was then peeled back using fine forceps. A layer of 2% low-melt agarose diluted in ACSF was applied over the exposed brain surface before a glass window cut from a cover slip (borosilicate glass 1.0 thickness) using a diamond knife (Sigma-Aldrich) was placed over the craniotomy. The edges of the window were then glued with medical super glue (Vetbond, 3M) to the skull.

Following surgery, mice were placed in a custom head-fixation apparatus and transferred to a two-photon microscope rig along with the heat pad. The microscope (Scientifica Multiphoton VivoScope) was coupled with a MaiTai DeepSee laser (Spectra Physics) tuned to 940 nm (<50 mW average power on the sample) for imaging. Images (512 × 512 pixels) were acquired in SciScan (Scientifica) with a resonant scanner at a frame rate of 30 Hz using a 16× 0.8 NA water-immersion objective (Nikon). The output of a four-channel version of the temporal olfactometer described above was adjusted to approximately 1 cm away from the ipsilateral nostril to the imaging window, and a flow sensor was placed next to the contralateral nostril for continuous respiration recording.

Awake recordings. For implantation of the head-plate, mice were anaesthetized with isoflurane in 95% oxygen (5% for induction, 1.5–3% for maintenance). Local (mepivacaine, 0.5% subcutaneously (s.c.)) and general analgesics (carprofen 5 mg/kg s.c.) were applied immediately at the onset of surgery. After surgery, mice were allowed to recover for 7 days with access to wet diet and, after recovery, habituated to the head-fixed situation for at least 15 min on three consecutive days preceding the imaging experiment.

Odour stimulation. For paired-pulse experiments, ethyl butyrate was diluted in mineral oil at a ratio of 1:5 and installed into a four-channel version of the high-speed odour delivery device (15 ml per vial) along with two blank positions (15 ml mineral oil). Odour concentration range was adjusted over 10 steps on a logarithmic scale with a factor of 1.25 by modulating odour pulse-width.

For correlated versus anti-correlated stimulus experiments, stimuli were generated from mixtures of physically mixed monomolecular odorants to ensure a high probability of finding odour-responsive cells in the dorsal OB using custom Python software (PulseBoy). Binary mixtures were diluted in mineral oil at a ratio of 1:5 and installed into a four-channel version of the high-speed odour delivery device (15 ml per vial) along with two blank positions (15 ml mineral oil). Mix 1: ethyl butyrate + 2-hexanone, mix 2: isoamyl acetate + cineole. During glomerular imaging experiments (Extended Data Fig. 6), six odours (A–F) were presented either individually or in pairs: A (ethyl butyrate), B (2-hexanone), C (isoamyl acetate), D (cineol), E (ethyl tiglate) and F ((+)-fenchone). For all stimuli, odour valve offsets were compensated by opening a corresponding blank position valve to ensure no global flow changes occurred over the course of the stimulus. All stimuli were repeated between 16 and 50 times with at least 15 s inter-stimulus interval.

Data analysis. For M/TC imaging, motion correction, segmentation and trace extraction were performed using the Suite2p package[47] (https://github.com/MouseLand/suite2p). Putative neuronal somata and dendritic segments were automatically identified by segmentation and curated manually. Soma and neuropil fluorescence traces were extracted and neuropil fluorescence was subtracted from the corresponding soma trace. Further analysis was performed with custom written scripts in MATLAB.

M/TCs were recorded in 17 fields of view (FOV) from 6 individual *Tbet-cre: Ai95(RCL-GCaMP6f)-D* mice, with 40 ± 9.23 (mean ± s.d.; range 27–48) cells per FOV and 30.25 ± 12.97 (mean ± s.d.; range 7–53) M/TC dendrites,

For glomerular imaging experiments, ROIs corresponding to glomeruli were manually delineated based on the mean fluorescence image. Fluorescence signal from all pixels within each ROI was averaged and extracted as time series. $\Delta F/F = (F - F_0)/F_0$, in which F is raw fluorescence and F_0 was the median of the fluorescence signal distribution.

Glomerular signals from a total of 15 individual *OMP-cre: Ai95(RCL-GCaMP6f)-D* mice were recorded with 28 ± 4.34 (mean ± s.d.; range 20–36) glomeruli per mouse (Extended Data Fig. 6a).

Given the nature of these population imaging experiments, sample size was not pre-determined and the investigators were not blinded to study conditions. Odour stimulus sequence was pseudo-randomized within experiments.

Where the odour stimulus was not inhalation-triggered, traces were post hoc aligned to the first inhalation after odour onset. Calcium response integrals were calculated for a range of window durations starting from odour onset (100–5,000 ms). To analyse how well odour responses predicted stimulus correlation on a trial-to-trial basis, we generated a linear discriminant classifier from the dataset and analysed prediction accuracy. For the classifier, we performed 50% holdout validation, splitting the data randomly into a training set and a test set with equal numbers of samples. We then performed linear discriminant analysis on the training dataset to determine the best linear boundary between 10 versus 25 ms pulse interval stimulations or correlated versus anti-correlated data. Classifier performance was then validated on the test dataset. To determine the effect of the number of ROIs used on classifier performance, we iteratively trained multiple classifiers on random subsets of ROIs with increasing numbers of ROIs within each set. For each ROI subset size, 100 classifiers were trained and the mean ± s.d. of their performance accuracy was calculated. All classifier analysis was performed on individual, unaveraged trials.

Glomerular imaging classifiers. The classifiers used in Extended Data Fig. 6k, l were trained separately for each odour pair, each frequency, and each time window. The inputs for classification were the averaged responses of the 145 glomeruli in a given time window for 24 odour presentation trials, where the odours were fluctuating in a correlated manner in half of the trials, and in an anti-correlated manner in the remaining half. Within the correlated and anti-correlated subsets of trials, half had the first odour in the pair phase-shifted by 180° and the remaining half had no phase shift. The classification task was to determine whether the glomerular responses in a given trial were evoked by correlated or anti-correlated odour fluctuations.

Because we had far fewer trials (24) than glomeruli (145) it was important to use regularized classifiers to avoid overfitting. To promote interpretability of the decision boundaries learned by the classifiers, we opted for sparsity-promoting regularizers and settled on the Lasso, evaluated as a classifier by taking the sign of its output computed after the addition of a small amount of noise (to decide ambiguous classifications). The implementation of the Lasso we used was LassoLarsCV provided by the Python scikit-learn library because it converged readily, gave very good classification performance, and automatically tuned the weighting of the sparsity penalty. Inputs to the classifier were standardized to have mean zero and unit variance across trials. We found that it was important to learn the classification weights without intercept to avoid overfitting.

The performance of a classifier was determined by cross-validation, where in each cross-validation iteration, the classifier was trained on a random approximately 90% of the trials (21 trials) and tested on the remaining approximately 10% (3 trials), and the test accuracy recorded.

The random subsets were selected in a stratified manner, meaning that the fractions of correlated and anti-correlated trials in the subset were kept as close as possible to their fractions in the full dataset (50/50). This meant that 10 of the 21 training trials were of one type and 11 of the other. This procedure was performed for ten cross-validation iterations, and the average performance over these repeats was recorded as the performance of the classifier. The shuffled performance was computed the same way but with training and test labels shuffled in each iteration. The entire procedure was then repeated for each of 100 different random seeds to produce a distribution of classification accuracies, the means and standard deviations of which are plotted in Extended Data Fig. 6k, l.

Because we used a sparsity-promoting classifier, it was straightforward to determine which glomeruli were contributing to a particular classification decision. We found that if we used all 145 glomeruli available then frequently glomeruli would be selected for noisy fluctuations of their responses that were by chance 'informative' for the classification. To avoid the inclusion of such noisy responses, we filtered glomeruli for responsivity. To determine the responsivity of a glomerulus, the mean $\overline{X}_{\text{glom,baseline}}$ and standard deviation $\sigma_{\text{glom,baseline}}$ of its responses pooled across all baseline bins (defined as the 3 s before odour onset) and across all trials for the given odour pair and frequency were first computed. A Z-score was then computed for its averaged response for the given time bin and for each trial by comparing this response to the baseline activity according to

$$Z_{\text{glom,trial}} = \frac{X_{\text{glom,trial}} - \overline{X}_{\text{glom,baseline}}}{\sigma_{\text{glom,baseline}}/\sqrt{n_{\text{wnd}}}},$$

in which $X_{\text{glom,trial}}$ is the response of the glomerulus in the given time window and n_{wnd} is the number of time bins constituting the window. The scaling of the baseline standard deviation is to account for the reduction in variance due to the averaging over time bins used to compute the response. A glomerulus was considered responsive in a given trial if the absolute value of its Z-score as computed above was greater than 1 on three-quarters or more of the trials. Such a thresholding ensured that the number of responsive glomeruli was almost always zero before odour onset, but rose to a peak of approximately 125 of the 145 glomeruli available when 2-s windows were used. Reducing the window size reduced the peak number of glomeruli, but at least 25 glomeruli were used during the peak responsive period in all cases, and frequently many more. This filtering also meant that some time windows late in the response contained no responsive glomeruli for some window sizes, which explains the 'patchiness' observed in Extended Data Fig. 6k, l.

Extracellular recordings
Surgical and experimental procedures. We anaesthetized 5–8-week-old C57BL/6Jax mice using a mixture of ketamine and xylazine (100 mg/kg and 10 mg/kg, respectively) by intraperitoneal (i.p.) injection. An i.p. line was inserted after the initial injection to allow easier and more regular subsequent injections of anaesthetics. Surgery was carried out as described above for two-photon imaging, up until the application of agar and cranial window.

Following surgery, the mice and custom platform were transferred to the extracellular recording set up. A flow sensor (A3100, Honeywell) was placed in front of the contralateral nostril while an output from the temporal olfactometer was positioned in front of the ipsilateral nostril. An Ag/Ag$^+$Cl$^-$ reference coil was immersed in the well, over the left hemisphere of the skull. The reference wire was connected to both the reference and ground of the amplifier board (RHD2132, intan), which was connected (Omnetics) to a head-stage adaptor (A32-OM32, NeuroNexus). A 32-channel probe (A32-Poly3, NeuroNexus) was connected to the adaptor, and the tip of the probe was manoeuvred to be positioned 1–2 cm above the craniotomy. The adaptor and probe were held above the craniotomy using a micromanipulator (PatchStar, Scientifica) set at 90° to the surface of the brain. The probe was moved towards the surface of the OB, while being observed through a surgical microscope. Once the probe was in contact with the surface, but had not entered the brain, the manipulator's Z position was set to zero. The signal from the probe was streamed through OpenEphys acquisition board and software (OpenEphys). The probe was inserted at <4 μm/s until the number and amplitudes of spikes began to decrease on deeper channels, indicating that the tip of the probe was exiting the MC layer. This was found to be between 400 and 600 μm from the surface of the OB. From here, the probe was left for 10 min for neural activity to stabilize before recording began.

Odour stimulation. Odours were presented using an eight-channel version of the high-speed odour delivery device, four of which contained odours (A: ethyl butyrate, B: 2-hexanone, C: isoamyl acetate, D: eucalyptol) and four contained blank (mineral oil) which were used to compensate for flow changes. Trials paired either A and B or C and D together. Stimuli were repeated 64 times and had an 8-s inter-trial interval. The onset of odour was recorded using TTL pulses passed through additional channels in the OpenEphys acquisition board. Trial starts were triggered on inhalation as detected by the flowmeter.

Data analysis. Spikes were sorted using Kilosort2[48] (github.com/MouseLand/Kilosort2) and classified as 'good' when they displayed a strong refractory period visible in their auto-correlogram, a typical waveform and a stable firing rate, as 'MUA' (multi-unit activity) if they presented a typical waveform but a weak refractory period, or as 'noise' if they were suspected of being electrical or mechanical interference. For a first-pass analysis, units were classified as 'differentially responding' to correlated and anti-correlated stimuli if units were found to have significantly ($P < 0.01$, Mann–Whitney U test) different spike time distributions during the 4 s after odour onset. However, the cut-off for such distinction will always be somewhat arbitrary. For the majority of the analysis, we therefore pooled all good units across experiments in a pseudo-population. All classifiers used for unit recording analysis were support vector machines (SVMs) with linear kernels with a low regularization parameter, which translates to greater freedom for a classifier to vary weights for any given component. Data were split into training and test sets before classification. Test sets consisted of either 26 trials (summed spike classifiers for correlated versus anti-correlated) or 2 trials (PCA classifiers and short odour pulse combinations). Data passed to the summed spike classifiers were pre-processed in one of two ways before classification.

First, a rolling sum of detected spikes within variable window sizes was used. The window sizes varied from 10 ms to 2,000 ms. In addition to window size, window starts were also varied. Each window size was trialled with every possible window start from zero to four seconds minus the window size from odour onset, with 10-ms incremental changes. For example, a 500-ms window was tested with starts varying from 0 to 3,500 ms after odour onset.

Second, the coefficients of principal components (PCs) for units in each trial were used for classification. The PCs were found by applying principal component analysis (PCA) across all units and all training trials. Each PC represented a time series and hence the coefficients signify the strength at which that time series was followed by a given unit for a given trial. The two holdout trials were not used to find the PCs but were then projected onto them and their coefficients used as the test for the classifiers. All classifiers were repeated 1,000 times with a random selection of holdout trials each time.

Finally, for the short odour pulse classification, classifiers were trained on summed spikes in windows of 500 ms after odour onset. Each classifier was trained on all but two hold-out trials. To account for a varying number of trials between mice, training data were bootstrapped to 1,000 trials of each type. Each trial was randomly selected

from the initial pool of training trials, and each unit was bootstrapped independently. These classifiers were tested on the initial two hold-out trials. This was repeated 1,000 times with different bootstrapped datasets and different hold-out trials.

Training data for all classifiers were scaled such that each feature (unit spike count/PC coefficient) had a mean value of zero and a standard deviation of one using the following equation: $z = (x - u)/s$, in which x is the initial value of the feature, u is the mean, s is the standard deviation, and z is the scaled value. Means and standard deviations were calculated using the training data so no information from the testing data could influence the scaling. The testing data were scaled using the same values as for the training. Scaling was applied during every repeat in this manner.

Whole-cell patch-clamp recordings
Experimental procedures. Whole-cell recordings were performed as previously described[38,49]. Borosilicate pipettes (2×1.5 mm) were pulled and filled with (in mM) $KMeSO_3$ (130), HEPES (10), KCl (7), ATP-Na$_2$ (2), ATP-Mg (2), GTP-Nax (0.5), EGTA (0.05) (pH = 7.3, osmolarity about 290 mOsm/kg). The OB surface was submerged with ACSF containing (in mM) NaCl (135), KCl (5.4), HEPES (5), MgCl$_2$ (1), CaCl$_2$ (1.8) (pH = 7.4 and about 300 mOsm/kg). Signals were amplified and low-pass filtered at 10 kHz using an Axoclamp 2B amplifier (Molecular Devices) and digitized at 40 kHz using a Micro 1401 analogue-to-digital converter (Cambridge Electronic Design).

After zeroing the pipette tip position at the OB surface, the tip was advanced to reach a depth of about 200 µm from the surface. Whole-cell patch-clamp recordings were obtained as described[50]. Series resistance was compensated and monitored continuously during recording. Neurons showing series resistance >25 MΩ were discarded from further analysis. To estimate the input resistance, a −50 pA current step was delivered at the start and end of each recording.

The vertical depth of recorded neurons was calculated as the vertical distance from the brain surface. Respiration was recorded using a mass flow sensor (A3100, Honeywell) and digitized at 10 kHz. Odours were prepared and delivered as described above and triggered to the beginning of inhalation.

Data analysis. *Change in membrane potential.* Recordings were spike-clipped using a custom script written in Spike2 (Cambridge Electronic Design) and analysed in MATLAB (Mathworks). All recordings were baseline subtracted as previously described[51]. The average change in membrane potential was defined as the difference between the average membrane potential over a 2 s period before odour onset and the average membrane potential in the first 500 ms (about 2 sniffs) after odour onset.

Change in spike frequency. Action potentials were counted from raw traces, converted into spike frequency in 50-ms bins and plotted as PSTHs. The net change in spike frequency was defined as the difference in the average spike frequency between 2 s before onset and 500 ms after onset.

Arithmetic sum. Baseline-subtracted traces obtained from independent component odour (A and B) presentations were either summed and averaged in an in-phase manner to generate the arithmetic sum equivalent of the correlated response or phase-shifted to generate the equivalent of the anti-correlated response.

Arithmetic sum (correlated) = $(Vm_{odourA} + Vm_{odourB})/2$

Arithmetic sum (anticorrelated) = $(Vm_{odourA} + \text{shifted } Vm_{odourB} + \text{shifted } Vm_{odourA} + Vm_{odourB})/4$

Dual-energy fast photoionization detection (defPID)
Two PIDs (200B miniPID, Aurora Scientific; Supplementary Fig. 2) fitted with UV lamps of emission energy 10.6 eV (PID high) and 8.4 eV (PID low) were used to discriminate ethyl butyrate (EB, ionization energy = 9.9 eV) from α-terpinene (AT, ionization energy = 7.6 eV) or ethyl valerate

(EV, ionization energy = 10.0 eV) from tripropyl amine (TA, ionization energy = 7.2 eV). To accommodate the lower-voltage UV lamp, resonance circuitry in the PID headstage electronics was adjusted according to the manufacturer's recommendations. Specifically, potentiometer 'PT1' was adjusted up to the point where the 8.4 eV lamp began to glow. Further, we tested whether the now converted PID low was now sensitive to only AT and TA while not detecting EB and EV. The PID inlets were connected with a three-way connector to detect incoming odours by both PIDs simultaneously from a common point. PID heads were held on lab stands with the PID inlet at approximately 4 cm above ground level.

Odour delivery. Odours were held in ceramic crucibles (5 cm diameter, 6 ml volume) covered in an air-tight fashion using glass lids. Odours were released for 5 s with an inter-trial interval of 15 s by Arduino-based robots programmed to lift the lids from the crucibles using a servo motor (TowerPro SG-5010, Adafruit). Lid lifting events were triggered by the Arduino board, recorded in Spike2 and defined as the onset of odour for analysis. Both the Arduino board and Spike2 were controlled by a portable computer and used the same clock for synchronization. Experiments were carried out in a large open space, both indoors and outdoors (Supplementary Fig. 2f, g).

Outdoors, PIDs and the odour delivery system as described above were used to record for multiple trials in different conditions on a day with low wind (about 8–12 mph ≈ 3–5 m/s, recorded with a two-axis ultrasonic wind sensor (Gill Instruments)). Outdoor experiments were performed on a ~6 m × 10 m wooden patio structure surrounded by trees. Measurement of odour correlations in the outdoor setting were complicated by the presence of background odours: if background odours are detectable by both PIDs, measured correlation will be artificially inflated; if they can be ionized only by the PIDhigh, they will artificially decrease the measured correlation.

Indoors, a digitally controlled fan (2214F/2TDH0, ebm-papst) was placed at a distance of 325 cm facing the PID inlet. An exhaust line was situated behind the PID inlet to ensure the flow of air from the fan towards the PID inlet. During a recording, the fan was set to maximum speed such that it pushed approximately 550 cf/min (cubic feet per minute, about 260 l/s) of air towards the PID inlet. A 25 × 25 × 25-cm Thermocool box was placed 200 cm downwind of the fan to act as an obstacle to air movement, promoting complex air movement patterns at the PID inlet. The pump at each PID was set to about 0.02 l/s suction speed, which is unlikely to perturb overall airflow dynamics substantially.

Recording conditions. We placed 6 ml of the desired odour(s) in two crucibles and placed them in different locations based on the experimental conditions as described below.

1. Low energy only: the 'low-energy odour' (AT or TA) was placed 40 cm (radial distance d) away from the PID inlet, and displaced either 25 cm left or 25 cm right of the midline (the line between the PID inlet and the centre of the fan). The odour source was alternated between left and right positioning relative to the midline to remove any possible bias from positioning in the air stream. The purpose of this recording condition was to generate data to calculate the linear transformation from the low energy signal to the high energy signal (Supplementary Fig. 2c, d).

2. Mix: 3 ml EB + 3 ml AT (or 3 ml EV + 3 ml TA) was pipetted into one crucible and placed either 25 cm left or 25 cm right of the midline at radial distances of 20 cm, 40 cm and 60 cm. The purpose of this recording condition was to determine how the temporal structure of individual odours in a plume behaved when the odours were emitted from the same source.

3. Separate: 3 ml EB and 3 ml AT (or 3 ml EV and 3 ml TA) were individually pipetted into two different crucibles and placed at a radial distance of 40 cm from the PID inlet. For the $s = 50$ cm apart condition,

one odour source was placed 25 cm left of the midline while the other was 25 cm on the right of the midline and vice-versa (equal number of trials for both cases), separating the odour sources by 50 cm. This procedure was repeated for lateral distances of $s = 30$ cm and $s = 10$ cm. The '50 cm apart' case was repeated for radial distances of $d = 20$ cm and $d = 60$ cm. The purpose of this recording condition was to determine how the temporal structure of individual odours in a plume behaved when the odours were emitted from separated sources but were still free to mix in air.

Data analysis. For the decomposition procedure, we recorded the low energy odour (AT) using both PIDs as described above. Assuming a linear relation between the recorded signals from the two PIDs, we plotted the recorded events with a linear regression fit (Supplementary Fig. 2c) and calculated slope and R^2 value of the fit. The scaling factor (6.82 ± 0.356, mean $\pm$ s.d.) was calculated as the average slope of all linear fits for $R^2 \geq 0.9$.

The 'PID low' traces were multiplied by this scaling factor, which was termed 'estimated low energy odour' (Supplementary Fig. 2e). The 'estimated high energy odour' was calculated by subtracting the estimated low energy odour from the 'PID high' traces.

For correlation calculation, custom-written scripts in MATLAB (Mathworks) were used to calculate the correlation coefficient between the estimated low energy odour and the estimated high energy odour for all conditions. Box plots were obtained from these values using Igor Pro 6 (WaveMetrics, USA).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data related to the OSN model (Extended Data Fig. 1) are available at https://github.com/stootoon/crick-osn-model-release. Data related to the glomerular classifier analysis (Extended Data Fig. 6) are available at https://github.com/stootoon/crick-osn-decoding-release. The remaining data that support the findings of this study will be made available by the authors upon request.

Code availability

All custom Python scripts to generate pulses (PyPulse, PulseBoy) are available at https://github.com/RoboDoig and https://github.com/warnerwarner. Code for controlling AutonoMouse is available at https://figshare.com/articles/AutonoMouse_Code/7616090. Code related to the OSN model is available at https://github.com/stootoon/crick-osn-model-release. Code related to the glomerular classifier analysis is available at https://github.com/stootoon/crick-osn-decoding-release.

39. Ishii, T., Hirota, J. & Mombaerts, P. Combinatorial coexpression of neural and immune multigene families in mouse vomeronasal sensory neurons. *Curr. Biol.* **13**, 394–400 (2003).
40. Haddad, R. et al. Olfactory cortical neurons read out a relative time code in the olfactory bulb. *Nat. Neurosci.* **16**, 949–957 (2013).
41. Madisen, L. et al. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* **85**, 942–958 (2015).
42. Raiser, G., Galizia, C. G. & Szyszka, P. A high-bandwidth dual-channel olfactory stimulator for studying temporal sensitivity of olfactory processing. *Chem. Senses* **42**, 141–151 (2017).
43. Abraham, N. M. et al. Maintaining accuracy at the expense of speed: stimulus similarity defines odor discrimination time in mice. *Neuron* **44**, 865–876 (2004).
44. Wadhwa, N., Rubinstein, M., Durand, F. & Freeman, W. T. Phase-based video motion processing. *ACM Trans. Graph.* **32**, 1–10 (2013).
45. Lopes, G. et al. Bonsai: an event-based framework for processing and controlling data streams. *Front. Neuroinform.* **9**, 7 (2015).
46. Ghatpande, A. S. & Reisert, J. Olfactory receptor neuron responses coding for rapid odour sampling. *J. Physiol. (Lond.)* **589**, 2261–2273 (2011).
47. Pachitariu, M. et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at https://doi.org/10.1101/061507 (2016).
48. Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M. & Harris, K. D. Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv* 061481 (2016) doi:.
49. Jordan, R., Fukunaga, I., Kollo, M. & Schaefer, A. T. Active sampling state dynamically enhances olfactory bulb odor representation. *Neuron* **98**, 1214–1228.e5 (2018).
50. Margrie, T. W., Brecht, M. & Sakmann, B. In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflugers Arch.* **444**, 491–498 (2002).
51. Abraham, N. M. et al. Synaptic inhibition in the olfactory bulb accelerates odor discrimination in mice. *Neuron* **65**, 399–411 (2010).

Extended Data Fig. 1 | Distinguishing fast odour stimuli with slow OSNs.
a, Membrane voltage relative to baseline of a single model OSN in response to a 10-ms odour pulse. Black traces are individual trials; red trace is average over 20 trials. OSN spike threshold has been set high enough to prevent spiking to illustrate the subthreshold voltage time course. **b**, Membrane voltages (grey traces) of ten OSNs from a population of 5,000 in response to a paired odour pulse with pulse width 10 ms and PPI of 25 ms. The voltage time course for one example OSN is in black. Several OSNs reach the OSN spike threshold (dashed red line) and are temporarily reset to the refractory voltage of −1. The population average membrane voltage (red) reveals membrane charging in response to odour stimulation and the subsequent discharging and refractory period. **c**, Raster showing the spike times (dots) of the full population from **b** and the corresponding mean firing rate (trace) estimated in 1-ms bins. **d**, Mean firing rates computed over 20 trials in response to paired odour pulses of width 10 ms and PPIs of 10 ms (green) and 25 ms (black). **e**, Model calcium signals are produced by squaring the instantaneous mean firing rate and filtering the result with a calcium imaging kernel. **f**, Model calcium responses to the paired odour stimulus with a PPI of 10 ms (green) and 25 ms (black). Thin traces are single trials, thick traces are averages over 15 trials. **g**, Schematic of the OSN model. Variables in dashed bounding boxes are changed for each glomerulus (Methods). **h**, Linear classifier analysis over an increasing subset size of glomeruli (1–100; plotted is mean ± s.d., 256 repeats for random subsets of n glomeruli generating 256 unshuffled and 256 shuffled accuracies).
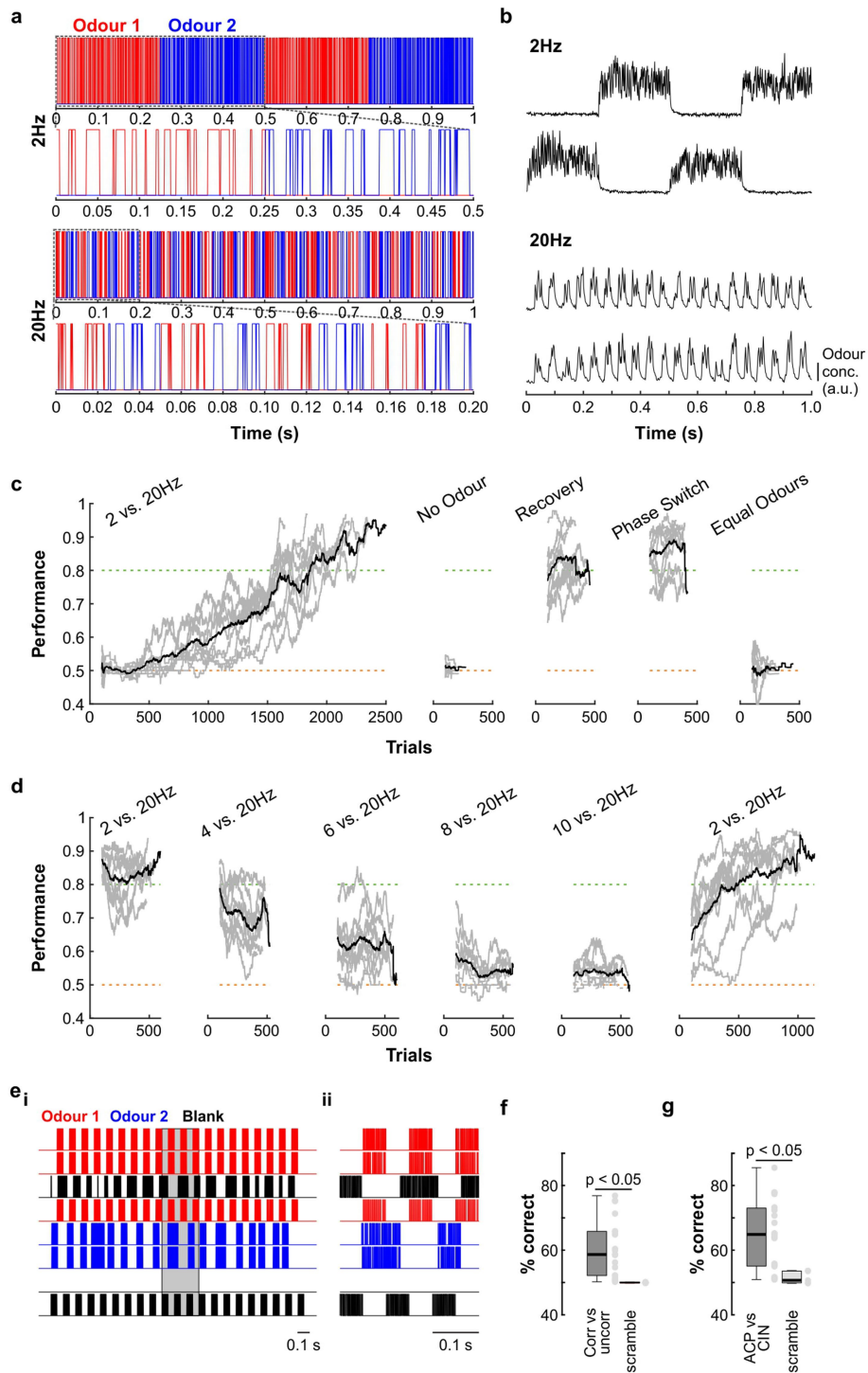
Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | Sub-sniff odour information in the olfactory bulb input layer. a, GCaMP6f fluorescence recorded in OB glomeruli in an anaesthetized *OMP-cre: Ai95(RCL-GCaMP6f)-D* mouse (maximum projection of 8,200 frames, glomerulus marked with red asterisk corresponds to first example trace shown in **b**). Scale bar, 50 μm. **b**, Example calcium traces in response to 10 and 25 ms PPI odour stimuli (mean of 50 trials ± s.e.m.). Bottom, example respiration traces. *P* values derived from unpaired two-sided *t*-tests comparing responses of individual trials integrated over 2-s windows to paired odour pulse stimulation. **c**, Classifier accuracy over an increasing number of glomeruli when a linear classifier was trained on several response windows (colour-coded; black, shuffle control) to PPI 10 versus 25 ms stimuli (mean ± s.d. of up to 93 glomeruli from 4 individual animals; 500 repetitions). **d_i**, Classifier accuracy when trained on all glomeruli in response to PPI 10 versus 25 ms stimuli recorded in anaesthetized animals (n = 93 glomeruli, mean ± s.d. from 4 individual animals) with a sliding window of different durations (colour-coded; black, shuffle control; 100 repetitions) starting at 2 s before odour onset (left) and time period between −0.5 and 0.5 s from odour onset shown at higher magnification (right). **d_{ii}**, Same as **d_i** for awake animals (n = 100 glomeruli, mean ± s.d. from 5 individual animals). **e, f**, Odour (**e**) and flow (**f**) signals integrated over 2 s for PPI 10 ms and PPI 25 ms stimuli (10 repeats each; odour, P = 0.1841; flow, P = 0.1786; unpaired two-sided *t*-test). **g**, Correlation coefficients of glomerular calcium responses to PPI 10 versus 25 ms in anaesthetized (n = 93 glomeruli from 4 individual animals) and awake (n = 100 glomeruli from 5 individual animals) mice (P = 0.3187, unpaired two-sided *t*-test, measured as in Fig. 1 from *OMP-Cre: Ai95(RCL-GCaMP6f)-D* mice). Violin plots show the median as a black dot and the first and third quartiles by the bounds of the black bar. **h_i**, Example respiration traces recorded using a flow sensor from awake mice. Inhalation goes in the upwards direction, exhalation downwards. **h_{ii}**, Average instantaneous sniff frequency from one example animal plotted as a function of time (n = 24 trials, mean ± s.e.m.). The odour stimulus consisted of two 10-ms odour pulses either 10 or 25 ms apart (Fig. 1c). **h_{iii}**, Distribution of sniff intervals during a 2-s window before (grey) and a 5-s window after (blue) odour stimulus onset (P = 1.02 × 10^{-189}, two-sample Kolmogorov–Smirnov test). **$h_{iv–vi}$**, Same as top row but for the anaesthetized condition (P = 0.3952, two-sample Kolmogorov–Smirnov test). **i**, Mean odour signal for PPI 10 and 25 ms for 10 increasing concentration steps defined by modulating valve pulse duty (Methods and Supplementary Fig. 1). There were no significant differences in odour concentration between both stimuli (unpaired two-sided *t*-tests). **j**, Modelled response integrals to PPI 10 versus 25 ms stimulations over a tenfold concentration range pooled over all 20 trials and 100 glomeruli (Methods). Box plots show median and extend from the 25th to 75th percentiles, whiskers extend to the 5th and 95th percentiles. **k_i**, Confusion matrix of SVM-based classification results of modelled glomerular signals in response to a range of ten odour concentrations ranked and colour-coded (n = 100 glomeruli). **k_{ii}**, Shuffle control with labels assigned randomly. **k_{iii}**, Confusion matrix showing the ranked and colour-coded results of glomerular responses independently classified for 10 ms versus 25 ms PPI and across the range of ten odour concentrations. **k_{iv}**, shuffle control for **k_{iii}** with labels assigned randomly. **l**, As in **j** but with 2-s response integrals derived from Ca^{2+} imaging data (10 repeats for each concentration). **m**, As in **k** for Ca^{2+} imaging data (n = 57 glomeruli, from 2 individual animals, 10 repeats for each concentration). Note that 10 ms PPI could be reliably distinguished from 25 ms PPI with only few instances where a response to, for example, a 10 ms PPI stimulus was misclassified as 25 ms or vice versa (compare light red quadrants to light green quadrants). **n**, Shifting the position of 10 ms PPI within a single inhalation. PPI 10 ms at position 1 (**n_i**) or at position 2 (**n_{ii}**) of three 10-ms odour pulses. Odour pulses as recorded with a PID shown in red, valve commands are shown in dark grey. Light grey area shows additional compensatory blank valve command to keep the flow profile indistinguishable between stimuli. **n_{iii}**, Total odour concentration was independent of the pulse profile (10 repeats, P = 0.57, unpaired two-sided *t*-test). **o**, The 10 ms PPI at both position 1 (**o_i**) and position 2 (**o_{ii}**) are presented during the inhalation phase (respiration shown in black, inhalation upwards, exhalation downwards). **p**, Example calcium traces in response to 10 ms PPI at position 1 (black) and position 2 (red), shown is the mean of 10 trials ± s.e.m. *P* values derived from unpaired *t*-tests comparing 2 s integrated responses of individual trials to odour pulses. **q**, Classifier accuracy over increasing number of glomeruli when a linear classifier was trained on the 2-s response to PPI 10 ms at position 1 versus position 2 (mean ± s.d. of up to 57 glomeruli, from 2 individual animals, 500 repetitions; blue: PPI 10 ms at position 1 versus position 2; black: shuffle control). For box plots, boxes indicate 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers (Methods).
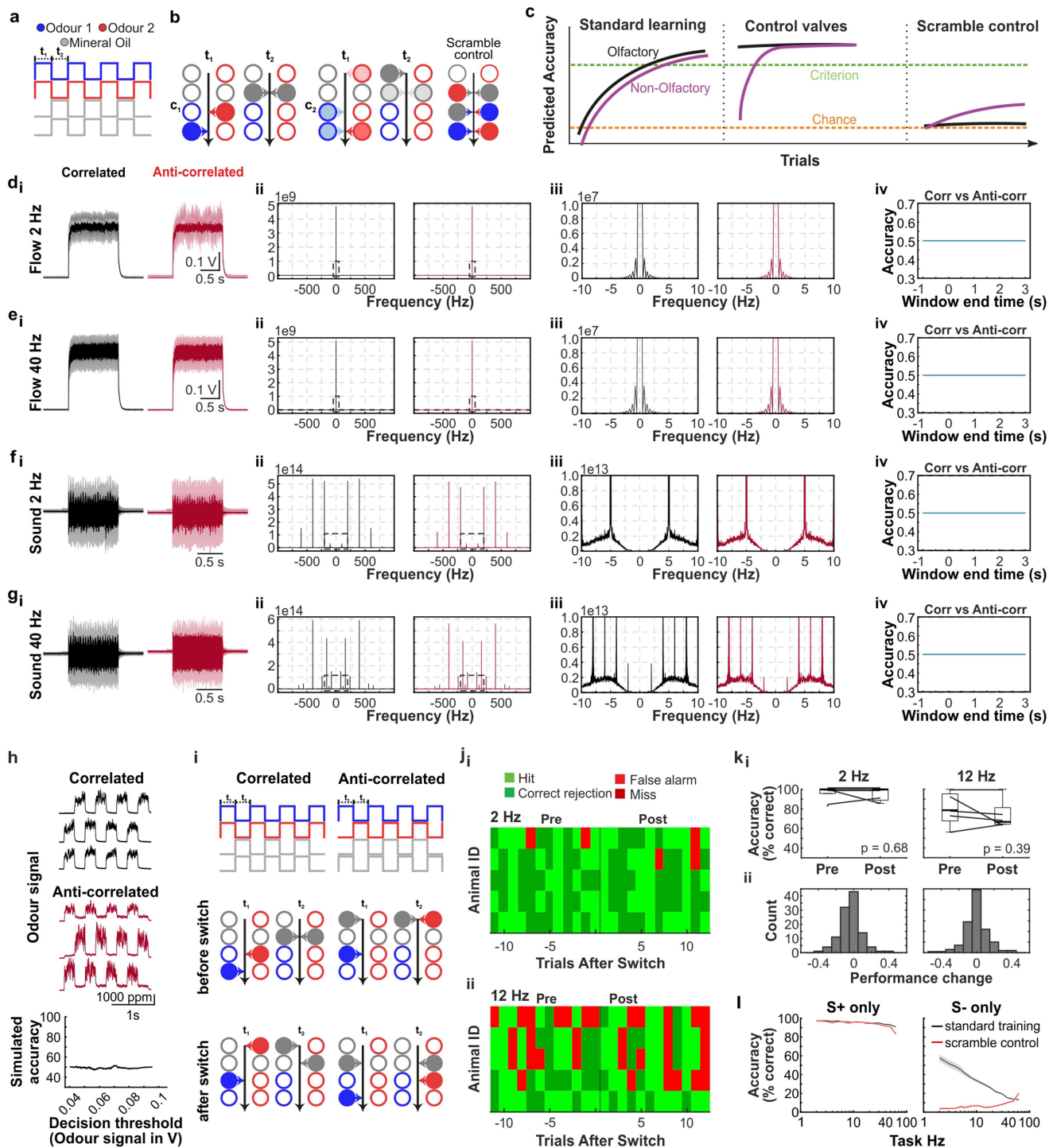
a

Odour 1 Odour 2

2Hz

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5

20Hz

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2

Time (s)

b

2Hz

20Hz

Odour conc. (a.u.)

0 0.2 0.4 0.6 0.8 1.0

Time (s)

c

Performance

2 vs. 20Hz   No Odour   Recovery   Phase Switch   Equal Odours

Trials

d

Performance

2 vs. 20Hz   4 vs. 20Hz   6 vs. 20Hz   8 vs. 20Hz   10 vs. 20Hz   2 vs. 20Hz

Trials

e i

Odour 1 Odour 2 Blank

ii

0.1 s   0.1 s

f

% correct

p < 0.05

Corr vs uncorr   scramble

g

% correct

p < 0.05

ACP vs CIN   scramble

Extended Data Fig. 3 | See next page for caption.

Article

Extended Data Fig. 3 | Frequency discrimination experiments. a, Frequency discrimination stimuli are produced by alternating presentation of two odours to generate a desired odour change frequency. During odour delivery, valves are not held open but rather are randomly opened and closed over time to produce slight variation in odour amplitude for each pulse. This means that odour concentration cannot be used as a cue to learn the task and odour switching frequency is the primary stimulus signal. Furthermore, valve clicking is randomized to minimize any acoustic cues. **b**, Replacing one odour channel with blank, un-odourized air and recording the frequency stimuli with a PID reveals that the desired odour pulse frequency is being produced. **c**, Mice readily learn to discriminate 2 versus 20 Hz pulse frequency stimuli in a go/no-go task. Replacing the odours with blank channels results in chance-level performance (no odour), which recovers when odours are replaced (recovery), showing that mice were probably discriminating the odour switching frequency rather than any extraneous cues such as valve noise. The order of odour presentation in the stimuli had no effect on behaviour as when it was shifted (phase switch) no decrease in performance was observed. Additionally, performance was dependent on the alternation between different odours, as when the experiment was repeated with the same odours in each channel (equal odours) performance was at chance level. **d**, To determine the perceptual limit of frequency discrimination, the floor frequency used in the task over successive experiments was increased such that the difference in frequency between the stimuli progressively narrowed. Overall performance decreased as the difference in frequency grew smaller, reaching near-chance level with a frequency difference of 10 Hz (10 versus 20 Hz). Switching back to the original discrimination (2 versus 20 Hz) recovered performance quickly, showing that the drop in discrimination ability was truly due to the frequency difference rather than general deterioration of performance over time. e_i, Example uncorrelated stimuli. Combinations of odour 1 (red) and odour 2 (blue) valves are opened with temporal offsets and randomized pulse timing resulting in a correlation of 0 (Methods). Blank (black) valves are used to keep total airflow constant throughout the stimulus. e_{ii}, Higher magnification of the area shaded in grey. **f**, Animals show similar average accuracy as shown in Fig. 2k when probed to discriminate correlated from uncorrelated odour pulses at 10 Hz ($n = 19$ mice, mean ± s.e.m. of average accuracy = $0.6506 ± 0.0016$; after scrambling stimulus identity: $0.4997 ± 0.0032$; $P = 0.0175$, unpaired two-sided t-test). **g**, Animals show similar average accuracy when discriminating the correlation structure of a different odour pair (acetophenone versus cineol) at 10 Hz ($n = 19$ mice, mean ± s.e.m. of average accuracy = $0.6558 ± 0.0026$; after scrambling stimulus identity: $0.5165 ± 0.0048$; $P = 0.0129$, unpaired two-sided t-test). Grey dots mark average performance of individual animals. Boxes in **f**, **g** indicate 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers (Methods).
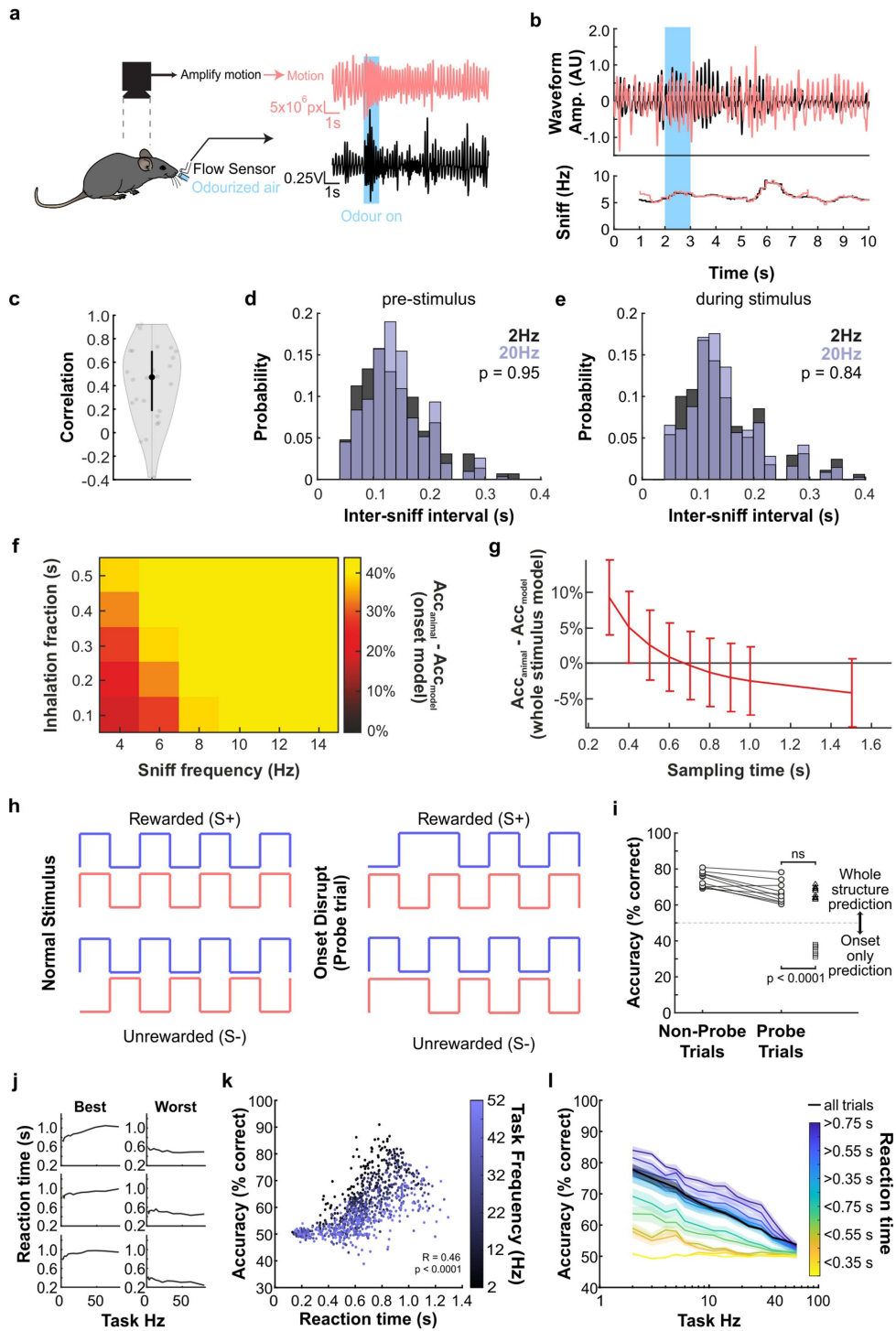
Extended Data Fig. 4 | See next page for caption.

Article

Extended Data Fig. 4 | AutonoMouse stimulus and experimental design.
a, Detailed schematic of stimulus production; odour presentation (odour 1: blue, odour 2: red) is always offset by clean air (mineral oil: grey) valves at the same flow levels, to ensure that total flow during the stimulus is constant. **b**, Schematic of the use of valve subsets to produce the desired stimulus. t_1 and t_2 represent valve openings at the corresponding time points shown in **a**. c_1 (left) and c_2 (middle) represent two possible configurations that could be used to produce the same resulting stimulus at the two time points. Opacity in the colours represents total concentration contribution to the resulting stimulus at the time point. For example, to produce the dual odour pulse at t_1, configuration c_1 can be used where odour 1 (blue) is delivered from one valve and odour 2 (red) from another valve. During t_2 two valves contribute clean air. Alternatively, configuration c_2 can be used in which during t_1 odour 1 (blue) is generated by 50% opening of two valves, with odour 2 (red) produced by 70%–30% opening of two other valves. Right, scramble control: valve maps (represented by arrow colour) are maintained compared to the training condition but odour vial positions are scrambled resulting in odour stimuli that are uninformative about reward association while maintaining any non-odour cue such as putative sound or flow contributions. **c**, Predicted accuracy for animals in the case that they use solely olfactory temporal correlations (black) and in the case that they use extraneous non-olfactory cues or non-intended olfactory cues (for example, contaminations, clicking noises) (violet). Note that when switching stimulus preparations to a new set of valves (as in Fig. 2i and **i**–**k**), such non-intended cues would not provide any information about stimulus–reward association, so accuracy would transiently drop back to chance. **d**$_i$, Average flow recordings (mean ± s.d.) of 2 Hz correlated (black, $n = 75$) and anti-correlated (red, $n = 70$) trials taken from the AutonoMouse odour port. **d**$_{ii}$, Fourier transform of the flow plots from **d**$_i$, showing the power of the signal over a range of 1 kHz. **d**$_{iii}$, An expanded view over the range of 10 Hz indicated by the dotted box in **d**$_{ii}$. **d**$_{iv}$, Mean accuracy of a series of linear classifiers trained on an increasing window of the integrated signal starting from 1 s before trial shown in **d**$_i$. Classifiers were tested on two withheld trials, one correlated and one anti-correlated, and repeated 100 times. **e**, As in **d** but for 40 Hz trials ($n = 69$ correlated and $n = 72$ anti-correlated). **f**$_i$, Average audio recording trace (mean ± s.d.) of 2 Hz stimuli using a microphone placed in close proximity to the AutonoMouse odour port. **f**$_{ii}$, **f**$_{iii}$, Fourier transforms of the audio signal from **f**$_i$. Note, although there are notable peaks at specific frequencies, these are present in both correlated and anti-correlated trials.

f$_{iv}$, Accuracy of a series of linear classifiers as shown in **d** but using the modulus of the audio signal. **g**, As in **f** but for 40 Hz trials. Note, whereas the sound profile and the Fourier transforms are different between 2 and 40 Hz, there is no difference detectable between correlated and anti-correlated trials. **h**, Example traces of odour signal (ethyl butyrate, isoamyl acetate, PID recorded) during correlated (top) and anti-correlated trials (middle). Simulated maximum accuracy based on differences in mean odour signal (bottom). Simulated accuracy was calculated as the fraction of trials correctly identified as correlated or anti-correlated based on a decision threshold set at some level between the minimum and maximum mean signal. Simulated accuracy was calculated for multiple decision thresholds, increasing the decision threshold from minimum odour signal to maximum odour signal in steps of 1/5,000th of the range between minimum and maximum. **i**, Detailed schematic of correlated (top left) and anti-correlated (top right) stimulus production before (middle) and after (bottom) switching valves. For the switch control, a set of previously unused odour valves is introduced to rule out potential bias towards a specific valve combination when performing the odour correlation discrimination task. **j**, Trial map of five representative animals during 2 Hz (**j**$_i$) and 12 Hz (**j**$_{ii}$) correlation discrimination tasks before and after the introduction of control valves ($n = 12$ trials before and 12 trials after new valve introduction, which is indicated by black vertical dotted line). Each row corresponds to an animal, each column represents a trial. Light green: hit, dark green: correct rejection, light red: false alarm, dark red: miss. **k**$_i$, Boxplots of mean accuracy for animals ($n = 5$ mice) pre- and post-control for 2 Hz (left) and 12 Hz (right). Box indicates 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers; Methods. P values derived from unpaired t-tests. **k**$_{ii}$, Summary histograms of performance change for all animals during all 'valve switch' control tests (Methods), indicating that discrimination accuracy was based on intended olfactory cues. The five animals with the best performance before the valve switch or bottle change (and thus the largest potential to drop in performance) were analysed. **l**, Discrimination accuracy ($n = 33$ animals, mean ± s.e.m.) for rewarded S+ (left) and unrewarded S– (right) trials when odours were presented using standard training valve configurations (black) and scrambled valve identity (red), data from Fig. 2k. Note that frequencies above 40 Hz were presented predominantly in the last block of the training schedule and reduced licking in the control group (decreased S+ performance and increased S– performance) might be due to decreased motivation at that point.
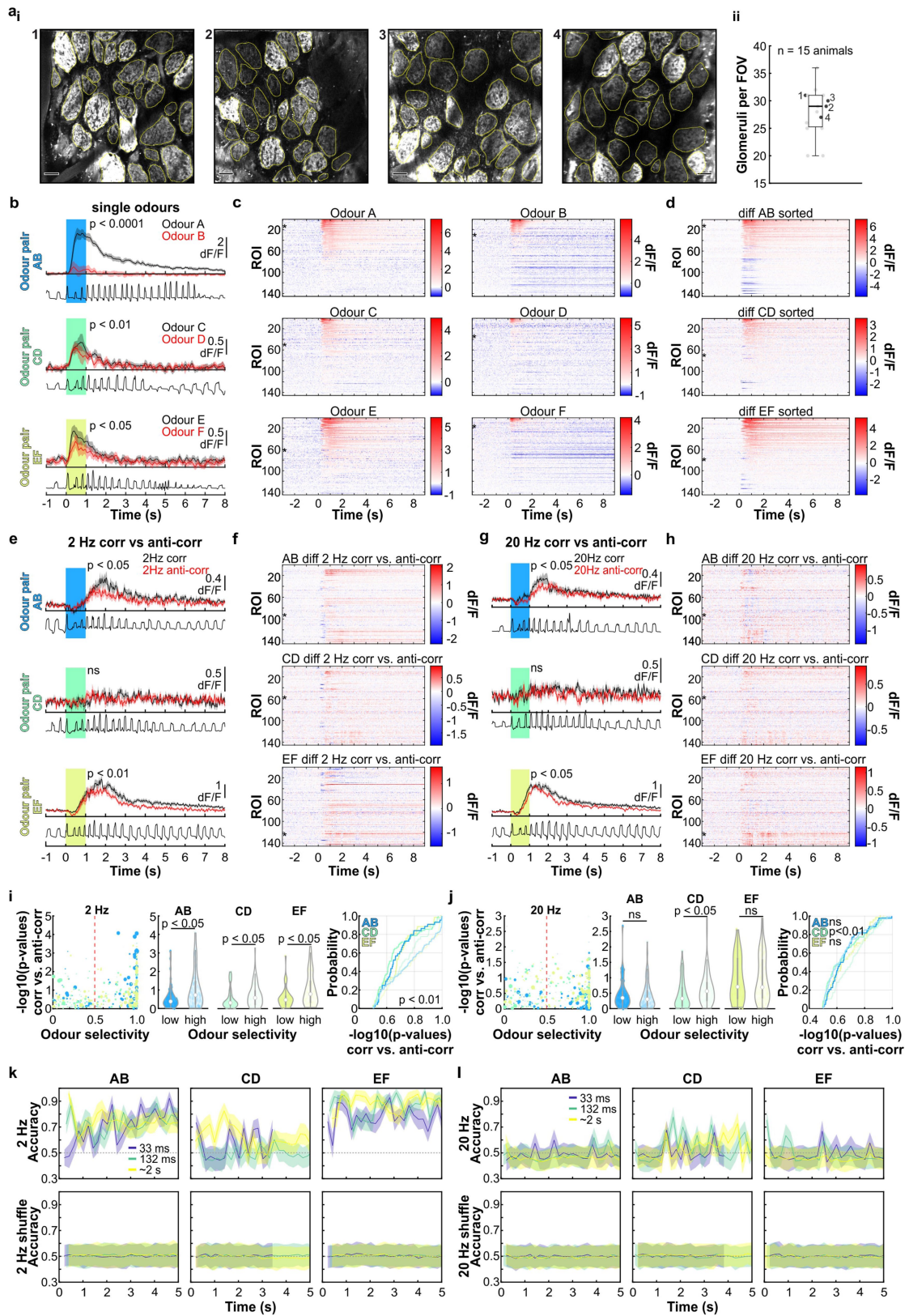
Extended Data Fig. 5 | See next page for caption.

Article

Extended Data Fig. 5 | Respiration recordings, stimulus onset model and reaction time for correlation discrimination experiments. a, An overhead camera was used to image a head-fixed mouse during a sequence of odour presentations. Simultaneously, a flow sensor was placed close to one nostril to monitor respiration to establish the validity of motion imaging-based respiration recording. Phase-based motion amplification was used to magnify motion on the animal's flank to capture body movements associated with respiration. Right, example of simultaneous respiration measurement with motion imaging (red) and flow sensor (black; Methods and Supplementary Video 2). **b**, Three further example trials with respiration rate extracted from motion imaging (red) and simultaneous flow sensor recording (black). Below, instantaneous sniff frequencies calculated from either sensor were tightly correlated. **c**, Correlation between respiration traces extracted from motion imaging and respiration captured by flow sensor ($n = 26$ trials, 10 s duration each). Violin plot shows the median as a black dot and the first and third quartiles by the bounds of the black bar. **d**, **e**, Probability distributions of inter-sniff intervals for odour presentations (isoamyl acetate versus ethyl butyrate, 2 Hz and 20 Hz) for freely moving animals in AutonoMouse before stimulus onset (**d**) and during 2 s odour stimulation (**e**; $n = 605$ sniffs for 2 Hz and $n = 668$ sniffs for 20 Hz, two-sample Kolmogorov–Smirnov test). **f**, Heat map of accuracy difference between a model in which animals rely on onset information only (Methods) and actual animal accuracies across a range of sniff frequencies and inhalation fractions ($n = 10$ mice). No matter the assumed sniff frequency and inhalation frequency, the 'onset model' deviates substantially from the accuracy measured in the behavioural experiments (**h**, **i**). **g**, Difference between a model in which animals use the entire stimulus structure (Methods) and actual behavioural accuracies across different stimulus sampling times ($n = 10$ repeats, mean ± s.d.). The 'whole stimulus' model accurately describes animal behaviour, indicating that mice do not base a decision about the correlation structure of a stimulus predominantly on the onset. Note the different scales in **f** and **g**. **h**, Schematic of experimental stimuli in which the first stimulus pulse was disrupted when presented on probe trials. Top, normal stimulus design; bottom, 'onset disrupt' stimuli, in which the first pulse in a correlated stimulus is disrupted to be anti-correlated; and vice versa for an anti-correlated stimulus. **i**, Animals were trained on standard (non-probe) correlation discrimination stimuli ($f = 10$ Hz) but onset disrupt (probe) stimuli were presented randomly on probe trials with a 1/10 probability. Accuracy was only slightly degraded on probe trials (mean ± s.d. of accuracy for non-probe trials 75.8 ± 4.4%; for probe trials 67.8 ± 6.1%; $P = 0.001$, paired two-sided t-test, $n = 9$ mice) but did not drop below chance ($P = 7.3 \times 10^{-6}$, paired t-test). Notably, accuracy on probe trials was consistent with whole-structure prediction (70.3 ± 3.5%, $P = 0.13$, paired t-test of comparison to probe trials) and differed significantly from the accuracy of onset-only prediction (41.6 ± 1.5%; $P = 1.02 \times 10^{-6}$, paired t-test of comparison to probe trials). **j**, Mean reaction time (time from stimulus onset to first lick in S+ trials) plotted as a function of stimulus pulse frequency for the three animals with the best (left) and the worst (right) global accuracy (mean accuracy across all trials). Better-performing animals tend to increase their reaction time as stimulus pulse frequency increases. **k**, Scatter plot of mean accuracy versus mean reaction time for each animal and stimulus pulse frequency condition (averaged over blocks of 100 trials). Points are colour-coded according to stimulus pulse frequency. Accuracy was significantly positively correlated with reaction time, suggesting that mice that sampled a greater portion of the stimulus made more accurate decisions about its correlation structure (Pearson correlation coefficient $R = 0.49$, $P < 1.1 \times 10^{-112}$). **l**, Accuracy (mean ± s.e.m.) is plotted as in Fig. 2k, but only trial blocks with reaction times above or below a certain threshold (colour code) are included in the analysis. Where only longer reaction times are considered, global performance is higher than the case in which only shorter reaction times are included, again suggesting that longer stimulus sampling improves discrimination of odour correlation structure across all stimulus pulse frequencies.
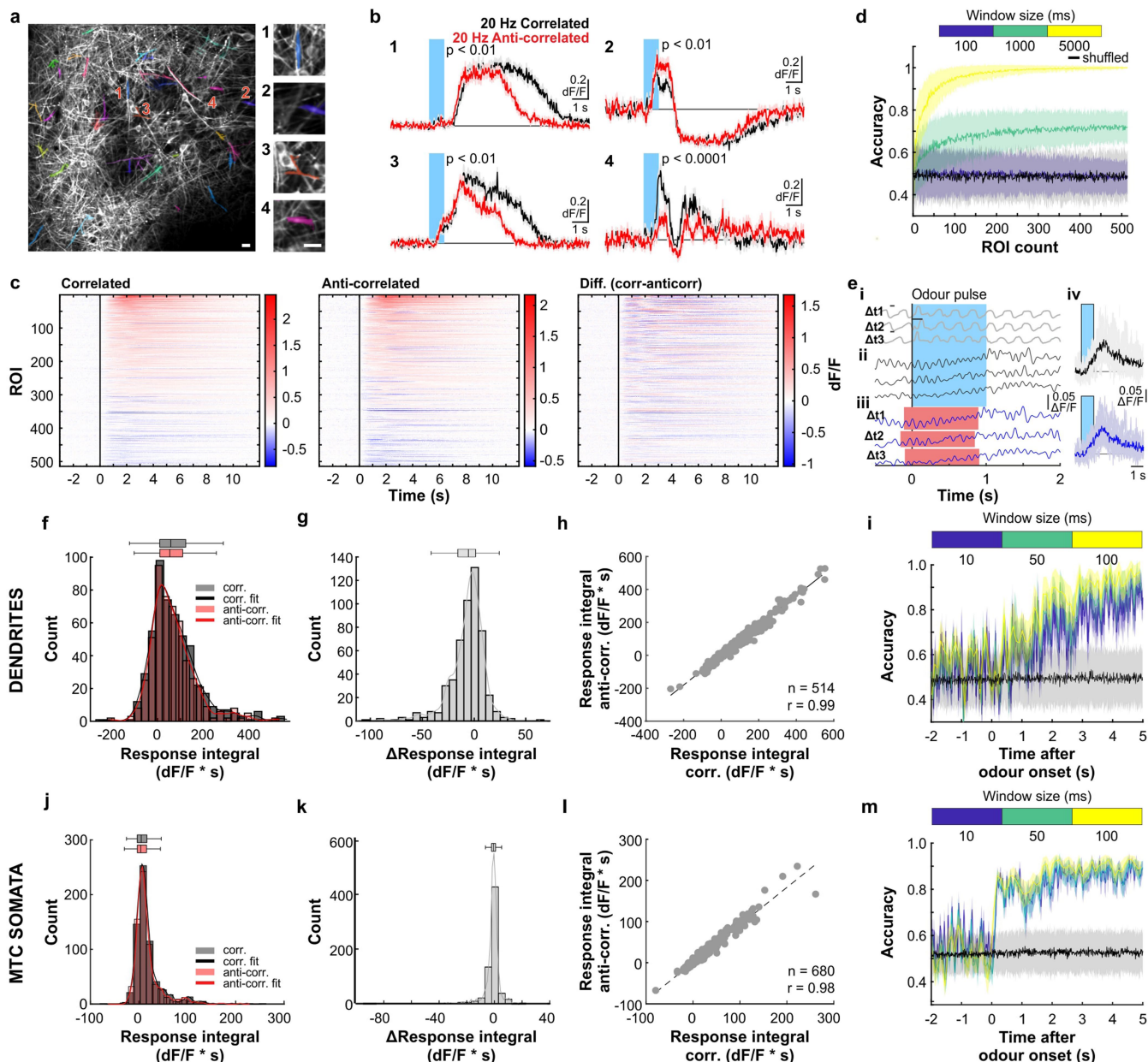
Extended Data Fig. 6 | See next page for caption.

Article

Extended Data Fig. 6 | OSN imaging in response to correlated versus anti-correlated odour stimulation. a_i, Four example fields of view (FOVs) recorded from the dorsal olfactory bulb of individual mice. a_{ii}, Number of individual glomeruli per FOV in all experimental mice ($n = 15$). The number of individually delineated glomeruli ranges from 20 to 36 with an average of 28 glomeruli per FOV. Labelled data points (1–4) correspond to FOVs shown in a_i. Scale bars, 50 μm. Centre line is median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered as outliers; Methods. **b,** Example glomerulus response from *OMP-Cre:Ai95(RCL-GCaMP6f)-D* mice to presentation of individual odours plotted pairwise (AB, CD, EF; mean ± s.e.m. of 6 trials). Stimulation period (1 s) is indicated by vertical bar (blue, green and yellow). Bottom, typical example respiration trace. P values derived from unpaired two-sided t-tests comparing 2-s integrated responses between paired odours. **c,** Averaged calcium transients from all glomeruli ($n = 145$ from 5 individual animals) in response to individual odours, plotted as colour maps sorted by response magnitude. **d,** Difference between glomerulus responses to individual odours plotted pairwise as colour maps. Glomeruli are sorted by average magnitude of response difference. **e,** Example glomerulus response to presentation of correlated versus anti-correlated odour pairs fluctuating at 2 Hz (mean ± s.e.m. of 12 trials). Bottom, typical example respiration trace. P values derived from unpaired two-sided t-tests comparing 2-s integrated responses of individual trials to correlated and anti-correlated odour stimulation. **f,** Difference between glomerulus responses to 2 Hz correlated and anti-correlated odours
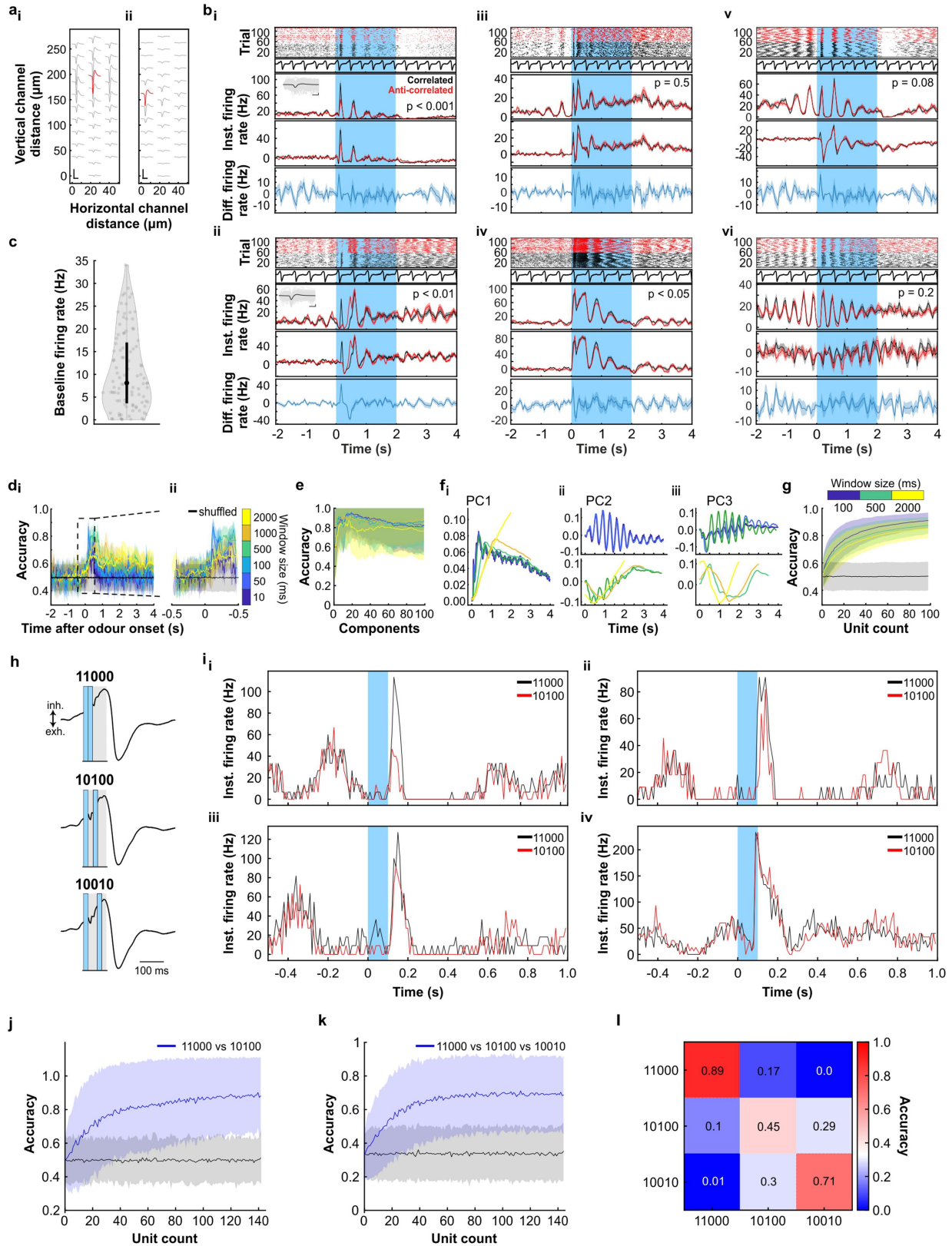
as colour maps, sorted as in **d**. **g, h,** As in **e, f** but for 20 Hz correlated versus anti-correlated stimuli. Example glomerulus from **b, e, g** indicated with an asterisk in colour maps in **c, d, f, h. i,** Left, P values derived from comparing trials of the summed 2-s response to correlated versus anti-correlated odour stimulation at 2 Hz (unpaired two-sided t-tests) for three odour pairs (colour-coded) as a function of glomerulus selectivity to individual odours ($n = 145$ glomeruli). Selectivity is calculated as the difference between the absolute response to single odours scaled by the summed absolute response. A threshold is set at 0.5 to define glomeruli as low- or high-selective. Dot size represents magnitude of the summed response. Middle, comparison of P values between low- and high-selective glomeruli ($P < 0.05$, unpaired two-sided t-test). Violin plots show the median as a white dot and the first and third quartiles by the bounds of the grey bar. Right, cumulative distribution function of P values for low- and high-selective glomeruli ($P < 0.01$ for all pairwise comparisons, two-sample Kolmogorov–Smirnov test). **j,** As in **i** but for 20 Hz ($n = 145$ glomeruli). **k,** Top, mean ± s.d. of classifier accuracy over 100 repetitions when trained on all responsive glomeruli ($n = 145$ available, from 5 individual animals, Methods) to discriminate 2 Hz correlated versus anti-correlated stimuli, trained separately for each of the three odour pairs and within sliding windows of different widths (colours); x-coordinates indicate latest extent of each window. Bottom, same as top row but with labels shuffled as control. **l,** As in **k** for 20 Hz correlated versus anti-correlated odours. Some data points in **k, l** are absent because not all time points had responsive ROIs for every window size (Methods).

Extended Data Fig. 7 | Odour correlation structure is encoded in dendrites of olfactory bulb output neurons. a, GCaMP6f fluorescence from mitral and tufted cells and their dendrites recorded in the dorsal portion of the olfactory bulb of a *Tbet-cre: Ai95(RCL-GCaMP6f)-D* mouse (maximum projection of 8,000 frames). Dendritic ROIs are superimposed in colour. Four dendritic segments (1–4) are shown in higher magnification; scale bars, 20 μm. **b**, Four example calcium traces extracted from dendritic segments shown in **a** that show differential response kinetics to correlated (black) and anti-correlated (red) stimulation (mean ± s.e.m. of 24 trials, f = 20 Hz). In total, 24% of dendritic segments showed significantly different integral responses (0–5 s after odour onset, $P < 0.01$, unpaired two-sided t-test; 121/514) to the two stimuli. **c**, Average calcium transients as colour maps for correlated (left) and anti-correlated stimuli (middle) and the difference between the two (right) of all analysed dendritic segments (n = 514, from 6 individual animals). **d**, Classifier accuracy over an increasing number of dendritic ROIs trained on several response windows (colour-coded) to discriminate correlated versus anti-correlated stimuli at 20 Hz (n = up to 514, mean ± s.d. from 6 individual animals, black: shuffle control). **e**, Method of aligning calcium traces to first inhalation after odour stimulus onset. **e_i**, Representative respiration traces recorded using a flow sensor placed in front of the nostril contralateral to the imaging window.
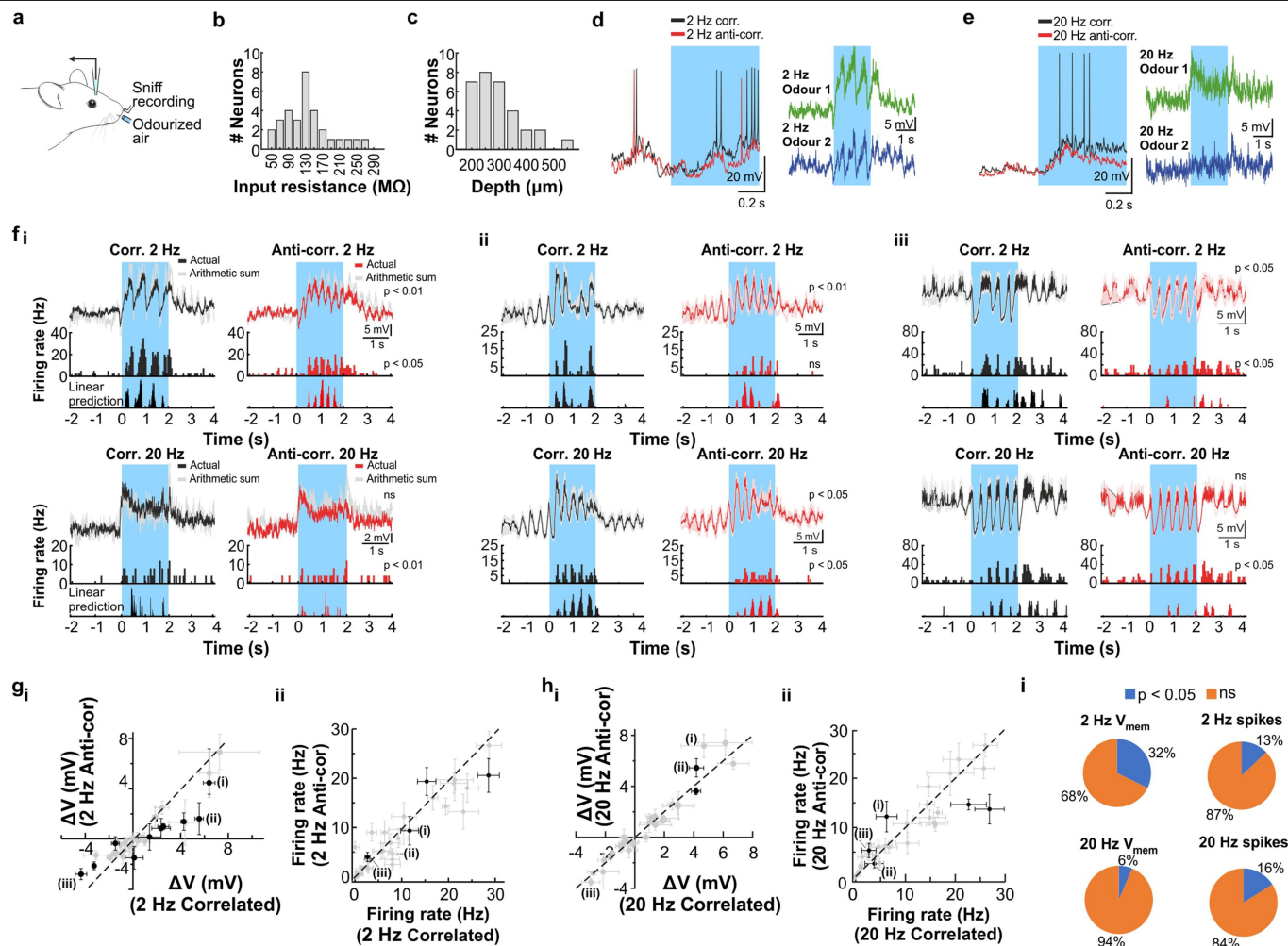
The first inhalation peaks were detected and the time (Δt) to the first inhalation after odour onset was calculated for each trial individually. **e_ii**, Representative calcium transients in response to a single odour presentation (here: 20 Hz correlated). **e_iii**, Transients are shifted according to Δt. **e_iv**, Individual calcium transients (faint colours, 24 trials) in response to 20 Hz correlated odour presentations with the average calcium signal (thick traces) superimposed. Top, before aligning to first inhalation after odour onset; bottom, after alignment. Blue bar represents the odour presentation phase (approximate for the aligned data). **f**, Distribution of odour response integrals from all recorded ROIs (n = 514) for correlated (grey) and anti-correlated (red) stimulation. Box indicates 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers; Methods. **g**, Histogram of the difference between correlated and anti-correlated odour responses. Box plot as in **f**. **h**, Comparison of correlated and anti-correlated odour responses of all dendritic ROIs (f = 20 Hz, n = 514 dendrites). **i**, Classifier accuracy when trained on all dendritic ROIs recorded with a sliding window of different durations starting 2 s before odour onset (colour-coded; black, shuffle control; n = 514 from 6 individual animals; mean ± s.d., 100 repetitions). **j–m**, As in **f–i** but for projection neuron somata (f = 20 Hz, n = 680 cells; Fig. 3).

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Projection neuron unit recordings in response to correlated versus anti-correlated stimulation and short odour pulse combinations. a, Data from unit recordings as in Fig. 3h–k. Average waveforms across all channels of two isolated units shown in **b**$_{i,ii}$. Each waveform represents the average waveform for the unit on a specific channel. Red waveform indicates the channel with the largest average waveform for the unit. Scale bars, 100 μV (vertically) and 1 ms (horizontally). **b**, Additional example single unit odour responses to correlated (black) and anti-correlated (red) stimuli shown as raster plots (top) and PSTHs (mean ± s.e.m. of 64 trials for each condition) of spike times before (second from top) and after baseline subtraction (second from bottom), and the differential PSTHs for correlated and anti-correlated stimuli (bottom, blue). Average spike waveforms shown as insets in **b**$_{i,ii}$. Duration of odour presentation (2 s) is indicated in light blue. *P* values derived from a two-sided Mann–Whitney *U* test comparing the spike time distributions of correlated and anti-correlated trials during the 4 s after odour onset. **c**, Average baseline firing rate for all units (*n* = 97 from 6 individual animals). Baseline firing rates were calculated from 4 s to 0 s before odour onset for each of the 1,312 trials presented during all recordings. Violin plot shows the median as a black dot and the first and third quartiles by the bounds of the black bar. **d**$_i$, Classifier accuracy when trained on all baseline-subtracted units in response to 20 Hz correlated versus anti-correlated stimulation (*n* = 97 units, mean ± s.d. from 6 individual animals) with a sliding window of different durations (colour-coded; black, shuffle control; 100 repetitions) starting at 2 s before odour onset. Time along the *x*-axis represents the end time of the window. **d**$_{ii}$, Time period between −0.5 and 0.5 s from odour onset shown at higher magnification (*n* = 97 units, mean ± s.d. from 6 individual animals). **e**, To take the entire temporal structure of responses into account we performed a PCA on the temporal evolution of the firing rate responses (Methods). Shown here is the accuracy for linear SVM classifiers (mean ± s.d.) trained on increasing numbers of PCs. Classifiers were trained on all but two trials (one correlated, one anti-correlated). Training and testing were repeated 1,000 times. The colour code represents the same window sizes as in **d**. **f**, The first (**f**$_i$), second (**f**$_{ii}$), and third (**f**$_{iii}$) PCs found from PCA for different rolling window sizes (colour code as in **d**). In the second and third PCs, the windows have been split as to better compare the similarities in PCs for different window
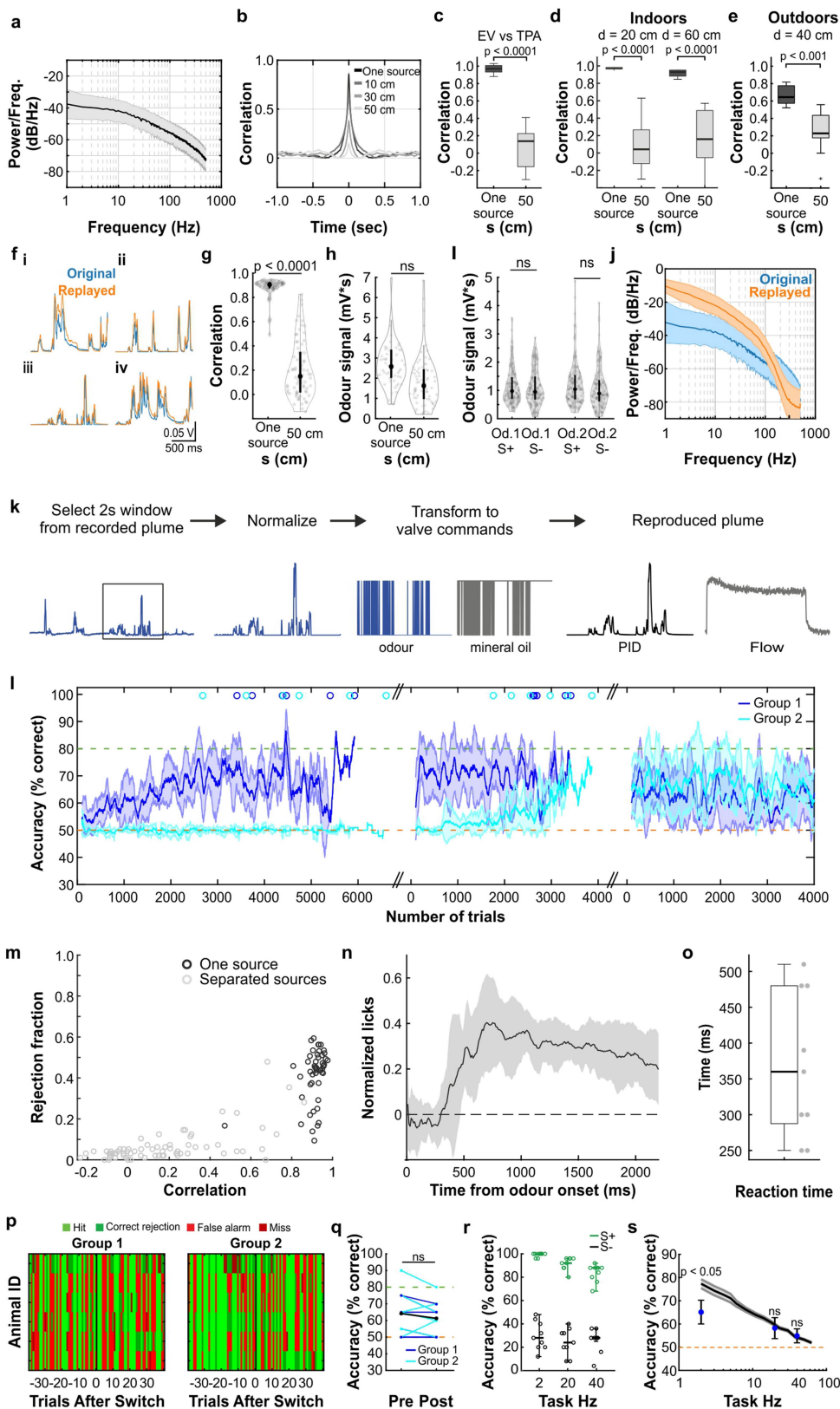
sizes. **g**, Average classifier accuracy of a set of classifiers trained on the PC weights of increasing number of units. Classifiers were trained on all but two trials (one correlated, one anti-correlated). The number of PCs used for each window was selected by the peak accuracies in **e** (colour-coded; *n* = up to 97 units from 6 individual animals; mean ± s.d. of 1,000 classifier repetitions). **h**, Schematic of odour pulse stimulus timings in relation to the respiration cycle. Three combinations were presented, with each trial 120 ms in length. For example, 11000 (top) consisted of a 40-ms odour pulse (light blue) followed by 80 ms of blank odourless air (grey). All trials were triggered at the onset of inhalation. **i**, PSTHs from four example units (**i**$_{i–iv}$) showing their average firing rate before, during, and after odour presentation (light blue vertical bar). Responses are to either 11000 trial (black) or 10100 odour presentation (red). The instantaneous firing rate was calculated by summing the number of detected spikes in 10-ms windows and multiplying the value by 100 to get Hz. **j**, Accuracy of linear classifiers as a function of the number of units available for training or testing (mean ± s.d. of *n* = up to 145 units from 8 individual anaesthetized animals). Each classifier was trained on the summed spike count of the available units in a window of 500 ms starting at odour onset. The classifiers were trained on all but two trials (one 11000 and one 10100 trial) and the number of repeats between animals varied between 11 and 30. To account for this and to minimize the variability of the training set, trial number was bootstrapped to 1,000 repeats. This was achieved by randomly selecting a repetition for each unit independently. The test set was isolated from the responses before bootstrapping and thus was not seen by the classifier until it was tested on it. Each classification was repeated 500 times with a different selection of units, and a different test set. The shuffled control (black) was accomplished by shuffling the training labels during each iteration of the classifier without shuffling test labels. **k**, As in **j** but classifying all three odour pulse combinations shown in **h**. **l**, Confusion matrix showing the fractions that each trial type was classified as (*n* = 145 units from 8 individual animals). True labels are shown on the *x*-axis and labels predicted by the classifier on the *y*-axis. Accuracies correspond to maximum unit count shown in **c**, **d**. The classifiers can readily separate between trials containing a single 40-ms odour pulse. Accuracy is lower when distinguishing between an intermission of 20 or 40 ms but remains above chance (chance = 0.33).

Extended Data Fig. 9 | Whole-cell recordings of projection neurons in response to correlated versus anti-correlated odour stimulation.
a, Schematic of the whole-cell patch-clamp recording approach.
b, **c**, Distributions of input resistance (**b**) and recording depth (**c**) as measured from all recorded projection neurons ($n = 31$). **d**, Left, example recordings from single cells with consecutive presentations of correlated (black) and anti-correlated (red) odour stimulus at 2 Hz. Duration of odour presentation (2 s) is indicated in light blue. Right, baseline-subtracted and spike-clipped subthreshold voltage response from a single cell to odour 1 (green) and odour 2 (blue) for 2 Hz. **e**, As in **d** but for 20 Hz odour stimulation. **f**, Voltage response from three example cells for correlated (black) and anti-correlated (red) odour stimuli for 2 Hz (top) and 20 Hz (bottom). The cell shown in **f$_i$** corresponds to the cell shown in **d**, **e**. The grey overlaid traces correspond to the arithmetic sum estimated from the response to individual odours. Bottom, linear prediction histogram calculated by thresholding the arithmetic sum of the subthreshold responses to the individual odours. Differences here suggest that correlation can be calculated at the single-cell level if the two individual odours engage overlapping OSN populations. P values are derived from a paired two-sided t-test of the membrane potential and the firing rate in the first 500 ms after odour onset. **g**, **h**, Average change in voltage (**g$_i$**) and in instantaneous spike frequency (**g$_{ii}$**) in the first 500 ms after odour onset from baseline membrane potential for 2 Hz (**g**) and 20 Hz (**h**) correlated versus anti-correlated odour presentation. Each marker corresponds to a single cell; error bars represent s.e.m. Data points in black represent cells where $P < 0.05$ between correlated and anti-correlated conditions. P values are derived from a paired t-test of the membrane potential and the firing rate in the first 500 ms after odour onset. Indicators **i**, **ii** and **iii** represent cells shown in **f**. **i**, Pie charts depicting the proportions of cells showing significant difference as described above (blue) in subthreshold membrane potential (left) and spike frequency (right) for all 2 Hz (top) and 20 Hz (bottom) cells. P values are derived from a paired t-test of the membrane potential and the firing rate in the first 500 ms after odour onset.

Extended Data Fig. 10 | See next page for caption.

Article

Extended Data Fig. 10 | Odour plume generation and additional analysis of source separation experiments. a, Power spectrum of all recorded odour plumes (mean ± s.d. of log power, $n = 132$ plumes). **b**, Cross-correlation of all recordings at different lateral separation distances. **c**, Correlation coefficients over all recordings for odours from the same source and for odour sources separated by 50 cm in a controlled laboratory environment with complex airflow (indoors; ethyl valerate (EV) versus tripropylamine (TPA); $n = 25$ for same source, $n = 27$ for sources separated by 50 cm; $P < 0.0001$, unpaired two-sided t-test). Box indicates 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers; Methods. **d**, As in Fig. 4b (for α-terpinene and ethyl butyrate) but for radial distances to the PID of 20 cm and 60 cm ($P < 0.0001$, unpaired two-sided t-test). **e**, As in **d** but measured outdoors ($n = 7$ for same source, 10 for sources separated by 50 cm; $P < 0.001$, unpaired t-test; indoors versus outdoors, one source: $P = 0.0060$, $s = 50$ cm: $P = 0.0632$, unpaired two-sided t-test). **f**, Example plume structures originating from the same source or separated sources as recorded with a PID (blue) and replayed with the multi-channel high bandwidth odour delivery device (orange). **g**, Correlation coefficients over all recordings of replayed plumes for one source ($n = 53$ plumes) and for sources separated by 50 cm from each other ($n = 74$ plumes; $P = 2.27 \times 10^{-41}$, unpaired two-sided t-test). **h**, Odour signals integrated over 2 s for all recordings of replayed plumes for one source ($n = 53$ plumes) and for sources separated by 50 cm ($n = 74$ plumes; $P = 0.75$, unpaired two-sided t-test). **i**, Odour plume signals integrated over 2 s for rewarded and unrewarded trials ($n = 150$ trials each; odour 1: $P = 0.4739$, odour 2: $P = 0.0923$, unpaired two-sided t-test). **j**, Overlaid power spectra (mean ± s.d. of log power) of all plumes ($n = 127$ plumes) recorded in complex, natural airflow conditions (blue) and replayed plumes (orange). **k**, Schematic of plume reproduction. First, a 2-s window is selected from the PID recording, starting around the middle of the trace and such that odour is present during the first 500 ms. Second, the trace is normalized between 0 and 1. Third, the trace is converted into a series of binary opening and closing commands directly related to the value of the normalized signal. A value of 1 translates to a continuous opening, and a value of 0 translates to continuously closed. This series of commands is relayed to an odour valve and an inverted version of the commands is relayed to a mineral oil valve to generate a compensatory airflow. The resulting output resembles the original plume, as measured with a PID, and there is constant airflow throughout the trial, as measured with a flow meter. The same procedure is then applied to the accompanying odour, to create both plumes needed for each trial. **l**, Group learning curves (mean ± s.d.) for the two groups of animals trained on the virtual source separation task, but on different sets of valves. Group 1 ($n = 6$ mice, blue) were trained on the task from the start, whereas group 2 ($n = 6$ mice, cyan) were first exposed to a scrambled version of the task and were later transferred to the same plumes as group 1. This served as a control that the cue required for learning is indeed olfactory information contained in the odour plumes. For the third stage of learning, the plumes were refined to ensure odour was always present in the first 500 ms of the trial and performance stabilized for the two groups. Mice progressed through these learning stages as a group, based on time elapsed from the beginning of training. Therefore, some mice performed more trials than others. The last trial performed by a mouse in each phase is represented by a colour-coded circle above the plot. Accuracy is calculated over a 100-trial sliding window. **m**, Rejection fraction (fraction of trials the mouse abstained from licking) calculated for each plume pair plotted in relation to the correlation between the two odour traces in that plume pair. Animals are trained to lick (expected low rejection fraction) for source-separated trials (low correlation) and abstain from licking (high rejection fraction) for one-source trials (high correlation). **n**, Difference in lick rates in response to source-separation training trials ($n = 9$ mice, mean ± s.d.), calculated for each mouse as lick rate (licks per 100 ms) in response to S+ trials minus the lick rate in response to S− trials, normalized to averaged lick rate for all trials across the corresponding time period. **o**, Reaction times for each mouse, calculated as the time point when the difference in lick rate for each mouse crossed a threshold (mean + 3 s.d. over the baseline, defined as the first 200 ms of the trace, when odour was not present). Box indicates 25th–75th percentiles, thick line is median, whiskers are most extreme data points not considered outliers; Methods. **p**, Trial map of all animals during virtual source separation tasks before and after introduction of control valves similar to Extended Data Fig. 4 ($n = 40$ trials before and 40 trials after new valve introduction, which is indicated by black vertical line). Each row corresponds to an animal, each column represents a trial. Light green, hit; dark green, correct rejection; light red, false alarm; dark red, miss. **q**, Mean performance of animals ($n = 11$ mice) that reached performance criterion during training during before and after control. **r**, Discrimination accuracy split by stimulus valence (green, S+; black, S−) for odour correlation fluctuation frequencies 2, 20 and 40 Hz (Fig. 4e; $n = 9$ mice, data are mean ± s.d., unpaired two-sided t-test). **s**, Group performance for the square pulse probe trials at different frequencies, in animals trained on the source separation task (blue dots, $n = 9$ mice, data are mean ± s.d.), compared to group performance where animals were trained on correlated and anti-correlated square pulse trains (from Fig. 2k, black line and s.e.m. band, $n = 33$ mice; 2 Hz: $P = 0.0018$, 20 Hz: $P = 0.19$, 40 Hz: $P = 0.94$, unpaired two-sided t-test). Violin plots in **g**–**i** show the median as a black dot and the first and third quartiles by the bounds of the black bar.

nature research

Corresponding author(s):   Andreas T Schaefer

Last updated by author(s): 3/14/2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Odour transmission data was collected with commercially available Spike 2 software (Version 8.10). Behavioral data was collected with custom open-source Python (3.6) code that controlled the automated behavior system: github.com/RoboDoig/autonomouse-control, with Qt5 as the framework for GUI generation. Other custom Python modules used to run this software are also availale at github.com/RoboDoig. 2-photon data was acquired with commercially available Scientifica software (SciScan 1.3). Extracellular unit recordings data was collected using the Open Ephys system (Version 0.5.3.1) and analysed using Kilosort 2. Experiment control and odour stimulation in 2-photon and electrophysiology experiments was orchestrated with custom open-source Python software: github.com/RoboDoig/PulseBoy and http://github.com/warnerwarner/PulseBoy Sound recordings were acquired via a Focusrite Scarlett 18i8 USB audio interface and saved using Audacity 2.4.2. Code and related data for the OSN model can be found at https://github.com/stootoon/crick-osn-model-release.
Data analysis	All data analysis was performed with Python (3.6 and 3.8), ImageJ (1.52), suite2p, Igor Pro 6 or MATLAB (2017b-2020a). Code related to the glomerular classifier analysis is available at https://github.com/stootoon/crick-osn-decoding-release.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

> Data related to the OSN model is available at https://github.com/stootoon/crick-osn-model-release.
> Data related to the glomerular classifier analysis is available at https://github.com/stootoon/crick-osn-decoding-release.
> The remaining data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](#)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Maximum sample sizes for each cohort were determined by the availability of simultaneously weaned mice to allow for large groups of male mice to be housed simultaneously for long periods of time without significant aggression within the group. Minimum sample sizes for each cohort were determined by requiring 3+ mice be in each subgroup within a cohort. Three separate experimental cohorts of mice were used for AutonoMouse experiments (correlation discrimination: cohort 1: n = 14 mice, cohort 2: n = 25 mice; cohort 3: plume discrimination: n = 24 mice). For physiological experiments a total of 49 mice were used (glomerular imaging: n = 9 mice, Mitra I/Tufted cell imaging: 9 mice, extracellular unit: n = 6 mice and whole-cell patch clamp recordings n = 25 mice). We found that due to the consistency in the structure of odour representations apparent across individuals, the number of animals sufficed for each experimental approach.
Data exclusions	For behavioral experiments, in the case that animals did not complete the pre-training phase (did not learn to gain sufficient daily water through self-initiated tasks) they were removed from the experiment. This was a pre-determined condition for experimental animal removal. For correlation discrimination experiments, one animal did not successfully pass the pre-training (1/39). For plume discrimination experiments, the second half of the cohort (n = 12/24 mice) with reversed valence (one source (S+, unrewarded), from separated sources (S-, rewarded) were not included in the analysis or carried forward to probe trials as they did not pass the performance criterion within the given timeframe. For imaging experiments, exclusion criteria were pre-established to select for experiments where fields of view could be imaged continuously for at least 2 hours with minimal drift and motion artifacts and where odour-evoked activity could be detected over the course of the entire imaging session. For extracellular recordings, units were classified as 'good' if they displayed a well defined inter-spike interval, a stable waveform, and a minimally varying baseline firing rate over at least half of the recording. Units were not selected by their odour responsiveness. For whole-cell patch recordings, cells with series resistance <25 MOhm were used for further analysis. No other data was excluded from our analyses.
Replication	For the odour transmission data, results were gathered from 3 environmental scenarios over 3 separate days and experiments. The principle result was confirmed across all these experiments and was reproducible from day-to-day. In the behavioral experiments, many of the key results (correlation detection, psychophysical threshold for correlation detection) were initially reported in one cohort of mice and then confirmed in an entirely separate cohort containing more animals and more detailed controls. Evidence for olfactory bulb input and output responses encoding odour correlation structure or paired pulses was found in 49 animals (imaging: n = 18 mice, extracellular recordings: n = 6 mice, whole-cell patch recordings: n = 25 mice) across 49 independent experiments.
Randomization	In all behavioural experiments, animals were first trained on a simple odour discrimination go/no-go task. Based on the performance levels in this task, animals were randomly assigned to different test subgroups until performance levels between the subgroups were statistically indistinguishable by a 1-way ANOVA.
Blinding	Due to the group selection method, experimenters could not be completely blinded to group allocation, as different experimental parameters had to be manually assigned to different animals based on their subgroup, in particular for test vs. control (scramble) mice. However, group allocation was fixed at the start of the experiment based on performance in a simple go/no-go task, and investigators did not move mice between groups after this initial choice. Also, since trial allocation was done entirely in software and the only distinguishing feature between mice was their RFID chip code (which could only be viewed in software), there was no possibility of investigators handling or otherwise treating mice differently due to their group selection. Therefore, although investigators could read-out the identity of the mice in the groups, it is very unlikely that this had any effect on subgroup differences.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	Human research participants
☒	Clinical data
☒	Dual use research of concern

Methods

n/a	Involved in the study
☒	ChIP-seq
☒	Flow cytometry
☒	MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	All mice used for behavioural experiments were 6-8 week old C57/Bl6 males at experiment initiation, and were used in the experiment for a maximum of 78 weeks. All mice used for in vivo calcium imaging experiments were 12-20 week old Tbet-cre (Haddad et al., 2013) or OMP-cre (Ishii et al., 2003) mice crossed with a GCaMP6f reporter line (Otazu et al., 2015) of either sex. All mice used for extracellular and whole-cell recordings were 5-8 week old C57 /Bl6 males. Mice were housed up to 5 per cage in a 12/12h light/dark cycle with food and water provided ad libitum.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples.
Ethics oversight	All animal procedures performed in this study were approved by the UK government (Home Office) and by the Institutional Animal Welfare Ethical Review Panel.

Note that full information on the approval of the study protocol must also be provided in the manuscript.