PLOS Computational Biology

RESEARCH ARTICLE

Shapley fields reveal mosaic chemotopic organization in the mouse olfactory bulb across diverse chemical feature sets

Nikola Milićević[1], Shawn D. Burton[2¤], Matt Wachowiak[2], Vladimir Itskov[3*]

1 Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, Texas, United States of America, **2** Department of Neurobiology, University of Utah School of Medicine, Salt Lake City, Utah, United States of America, **3** Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania, United States of America

¤ Current address: Department of Biological Sciences, Lehigh University, Bethlehem, Pennsylvania, United States of America
* vladimir.itskov@psu.edu

Check for updates

OPEN ACCESS

Citation: Milićević N, Burton SD, Wachowiak M, Itskov V (2026) Shapley fields reveal mosaic chemotopic organization in the mouse olfactory bulb across diverse chemical feature sets. PLoS Comput Biol 22(11): e1014917. https://doi.org/10.1371/journal.pcbi.1014917

Editor: Rich Pang, Neurotaxis, LLC, UNITED STATES OF AMERICA

Received: March 13, 2026

Accepted: October 21, 2026

Published: November 4, 2026

Copyright: © 2026 Milićević et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The data and code used in this study are publicly available at https://github.com/milicevic-nikola/Shapley-fields-analysis.

Funding: This work was supported by the NSF Next Generation Networks for Neuroscience Program, award 2014217 (M.W. and V.I.),

Abstract

Neural representations in the olfactory bulb (OB) remain poorly understood compared to well-established coding principles in other sensory systems. This difficulty arises because a natural coordinate system for olfactory stimulus space has not yet been identified, complicating the characterization of neural receptive fields analogous to those in other sensory modalities. Moreover, the extent to which these tuning properties are spatially organized within the OB remains unclear. To advance our understanding of the olfactory coding, we introduce an interpretable method of Shapley fields, as an olfactory analog of retinotopic receptive fields. Shapley fields are spatial distributions of chemical feature importance for the tuning of OB glomeruli. We used this tool to investigate chemotopy in the OB with diverse sets of chemical features using widefield epifluorescence recordings of the mouse dorsal OB in response to stimuli across a wide range of the chemical space. We found that Shapley fields reveal a weak chemotopic organization of the chemical feature sensitivity of dorsal OB glomeruli. This organization was consistent across animals and mostly agreed across very different chemical feature sets: (i) the expert-curated PubChem database features and (ii) features derived from a Graph Neural Network trained on human olfactory perceptual tasks. This chemotopy involved at least three chemical feature dimensions, rather than one or two dominant properties. Moreover, we found that the principal components of the Shapley fields corresponded to single commonly accepted chemical classification groups. This implies that these chemical groups could have been identified purely from the mouse olfactory responses. Our findings suggest that Shapley fields may serve as a chemical feature-agnostic method for investigating olfactory perception.

NIH grant F32MH115448 (S.D.B.), NSF grants 1555919 (M.W.) and 1555925 (V.I.), and NIH NINDS grant NS109979 (M.W.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

The sense of smell remains poorly understood compared with other sensory systems in the mammalian brain. This is because there is little understanding of how specific chemical features drive olfactory perception and how these features are spatially organized within the olfactory bulb. We introduce a novel method for understanding the coding properties of olfactory processing in the mammalian brain. Using neural activity in the mouse olfactory bulb, we characterized the non-random organization of chemical sensitivity across the surface of the olfactory bulb. We also found that the chemical sensitivity of mouse olfactory glomeruli, small structures that receive input from smell receptor neurons, was well described by AI-generated chemical features derived from human perceptual data. Rather than forming a simple map dominated by one or two chemical properties, this organization involved at least three effective chemical-feature dimensions. The spatial pattern was weak and mosaic-like: glomeruli with related sensitivities tended to occur near one another, but their responses overlapped rather than forming sharply separated regions. Our results and methods provide a feature-agnostic framework for studying olfactory coding in the mammalian brain.

Introduction

Primary sensory areas of the mammalian brain, such as the visual and the auditory systems, are known to encode characteristic features of the stimuli, such as the wavelength of light and frequency of sound. Furthermore, encoding of such stimulus features is spatially organized within topographic maps across primary and downstream cortical areas of visual, auditory, gustatory, and other sensory systems [1–3]. In contrast to these sensory systems, however, how physicochemical features of olfactory stimuli are encoded within the olfactory system, including whether they are topographically organized within the olfactory bulb (OB) – the primary olfactory sensory area in the mammalian brain – remains poorly understood.

Axons of the olfactory sensory neurons (OSNs) expressing the same olfactory receptor (OR) gene converge to deliver sensory input from the periphery to stereotypically positioned glomeruli on the OB surface. Due to the large number of OR genes, the dimensionality of olfactory stimulus space is potentially quite large. However, dimensionality reduction has been found to account for odorant similarity in chemical, neuronal and perceptual spaces in humans [4,5], mice [6] and across species [7–9]. These results indicate that olfaction follows a combinatorial logic rather than a simple labeled-line code, similar to receptive field patterns in lower-dimensional sensory systems.

Several studies indicate that glomeruli in mammalian OBs tuned to similar combinations of odorant chemical features are also located in close proximity on the OB and form chemical-feature clusters [10,11]. For example, in the rat OB, a homologous

series of odorant molecules can evoke distinct but spatially clustered responses that change systematically with carbon chain length [12–15]. Moreover, glomerular responses to certain odorant functional groups and hydrocarbon substructures also follow a systematic organization in the rodent OB [16–21]. These findings provide evidence for a topographic mapping of chemical features – or "chemotopy" – in glomeruli of the OB.

Several studies have cast doubt on the validity of the chemotopic organization, however. It was found in [6,22,23] that commonly used physicochemical features were not systematically represented by glomerular activity in the mouse OB. Instead, [6] argued that there is no chemotopic representation of chemical features in either fine or broad scales, while there is still a significant correlation between pairwise tuning similarity and distances between glomeruli. In [10,22,23] the authors argued for a functional "mosaic" organization of glomerular activity.

It is natural to define weak chemotopy as the property by which glomeruli, tuned to chemical features in "close proximity" in chemical feature space, are also in a spatial proximity on the OB. We call this chemotopy "weak" in reference to the imprecise and overlapping nature of the topographic organization with respect to the chemical features. Here, chemotopy refers to a spatially non-random mosaic of overlapping chemical-feature sensitivities, without implying a continuous map of chemical space onto the OB. This is in contrast to retinotopy in the primary visual cortex, where the visual receptive fields are sharply localized and organized into a continuous map of visual space. The analogy to retinotopic receptive fields is therefore methodological rather than a claim of equivalent topographic organization. However, the nature of chemotopy in the OB seems to greatly depend on the choice of chemical features, because different choices of chemical features significantly change the distances in the poorly understood chemical space. This dependence on the considered features may explain the inconsistencies of the previous studies of chemotopy.

Here, we set out to characterize chemotopy and its relationship to the choice of the chemical features that are being represented in the neural responses. To this end we developed a novel method of Shapley fields, as an olfactory analog of the retinotopic receptive fields. Shapley fields are distributions of chemical feature importance for the tuning of OB glomeruli. They are derived as an aggregate from Shapley values, a machine learning tool that allocates credit to features that explain model predictions [24], originally introduced in the context of game theory [25]. Another reason to use Shapley fields is that this tool automates feature selection and thus allows analysis of olfactory receptive properties using any significantly diverse chemical feature set. We used this tool to characterize chemotopy relative to diverse sets of chemical features using epifluorescence recordings of the mouse dorsal OB in response to stimuli across a wide range of odorants.

Our findings are consistent with a weak chemotopic organization of odor representations across OB glomeruli. We also established that our analyses do not depend on different choices of chemical feature sets: (i) the expert-curated PubChem chemical feature set and (ii) an artificial chemical feature set produced by a neural network, trained on human perception and molecular graphs. Both feature sets captured similar receptive properties in the mouse OB.

Results

Odorant-evoked responses were imaged with widefield epifluorescence across the dorsal surface of the OB in anesthetized mice, using artificial inhalation and a tracheotomy to directly deliver isoflurane anesthesia to ensure consistent odorant sampling (Figs 1A and S1B) and to minimize the possibility that isoflurane entered the nasal cavity, where it could perturb odorant responses by activating and potentially adapting isoflurane-sensitive OSNs [26], see Methods and [10].

Data were collected from 4 OBs (three mice), using 59 odorants that included a range of structural features covering a substantial portion of a previously studied physicochemical space of odorants [27] (S1A Fig). Responses to a given odorant across the four OBs were consistent in terms of the number and overall spatial distribution of activity (S1B Fig), as expected from the stereotyped nature of odorant receptor-to glomerulus mapping and consistent with prior studies showing stereotyped functional response patterns across OB glomeruli [10,23,28,29]. Additionally, all the ROIs from all OBs are illustrated in S1C Fig.

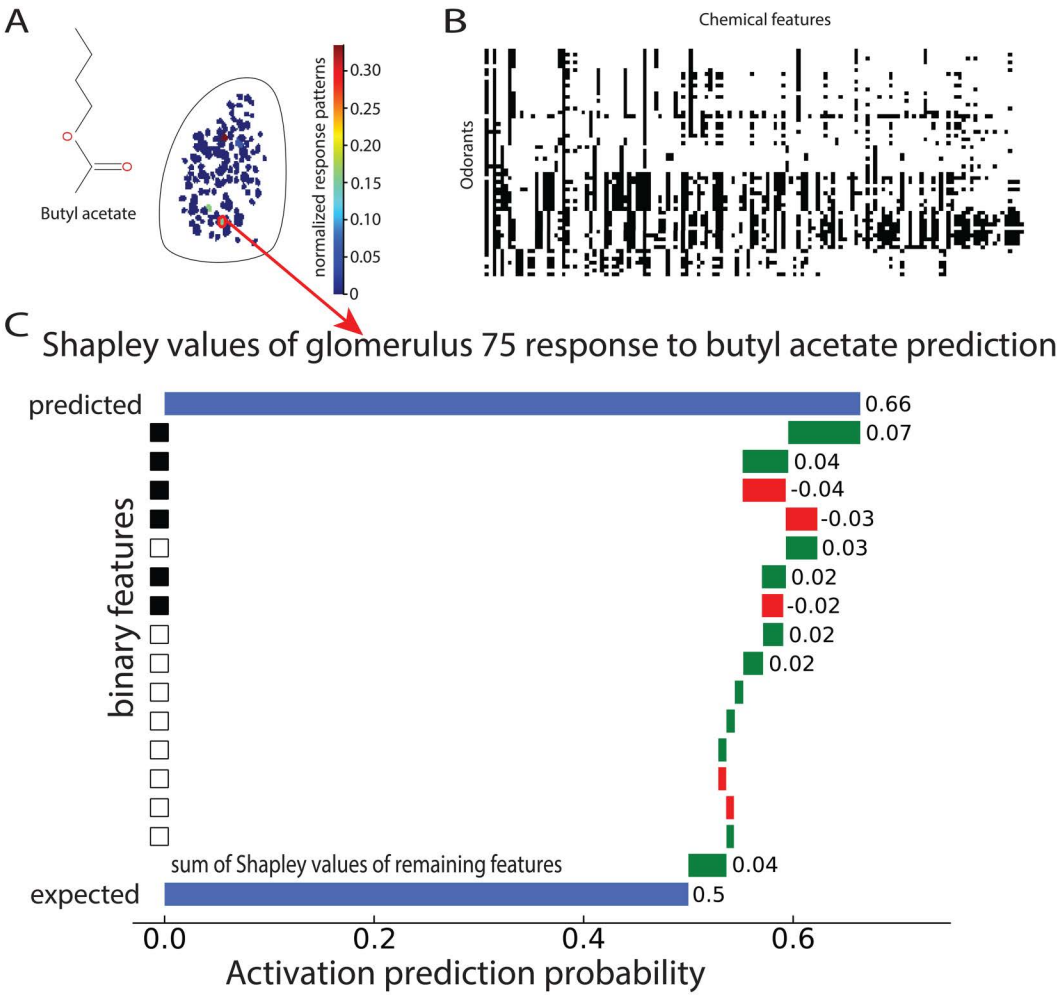


Fig 1. Predicting glomerular activation in response to odorants. A. Glomerular activation. An example of glomerular activation pattern in response to butyl acetate. B. PubChem chemical features. Binary chemical features of the presented odorants, from the PubChem database (Methods). Black indicates presence of a chemical feature, while white indicates absence. C. Explanation of chemical feature contribution towards predicting activation of a glomerulus to a particular odorant (butyl acetate). The contribution of each feature towards model prediction is calculated with Shapley values (Methods). Positive Shapley values (green) show contribution towards predicting activation; negative Shapley values (red) show contribution towards predicting no activation. The black squares on the *y*-axis indicate presence of a particular chemical feature in the odorant, the white squares indicate absence. The bottom blue line is the expected prediction; probability of activation the model would predict if no chemical features were given to the model. In this example, all Shapley values of features together with the model expected prediction sum to give the 66% probability of activation (top blue bar). This is represented in the waterfall plot; each feature's Shapley value contributes to the model prediction starting from 0.5 and moving the prediction to the left (if Shapley value is negative) or the right (if Shapley value is positive). Only the 15 features with the largest absolute Shapley values are shown in decreasing order from top to bottom. The sum of the Shapley values of the remaining features is also displayed. All values displayed are rounded to two decimal places. The boundary of the glomerulus in question is highlighted in red in panel A.

https://doi.org/10.1371/journal.pcbi.1014917.g001

To evaluate the sensitivity of glomeruli to odorant chemical features, we trained a Chemotuning model using the PubChem binary molecular fingerprint features (Fig 1B) [30]. The Chemotuning model assumes that each glomerulus activates based on a set of chemical features of a given odorant stimulus, irrespective of the glomerulus position in the OB. For each glomerulus, a random forest binary classifier was fit to predict activation of the glomerulus in response to the odorant, from the chemical features. More precisely, we trained a classifier to only detect activation or no activation. Our analysis did not consider whether a particular odorant activated a glomerulus more than another odorant (see Methods).

To understand the chemical receptive properties of glomeruli, we used Shapley values. Shapley values are a machine learning tool that enable allocation of credit to features explaining model predictions [24]. This tool is inspired by classical work in game theory [25]; our usage is similar to that in [31], and [32]. Shapley values are a general framework that unifies several other popular additive feature attribution methods such as LIME [33], DeepLIFT [34] and Layer-wise relevance propagation [35]. We used mean absolute Shapley values, similar to [24] (see Methods). For a given glomerulus-odorant pair, the magnitude of a feature's Shapley value quantifies its contribution to the predicted response, while its sign indicates whether the feature supports activation or no activation. Features that vary little across odorants generally contribute little to the prediction and therefore tend to have Shapley values near zero. Fig 1C shows the signed Shapley values for the Chemotuning model's prediction of the response of glomerulus 75 in the left OB of animal 1 to butyl acetate; the complete response profile of this glomerulus is shown in S2 Fig. The model's expected probability of activation before accounting for any chemical features is 0.5. This value does not reflect a balanced response dataset: only approximately 10% of the response-matrix entries are nonzero. Instead, it results from the balanced class weights used during model training, which penalize incorrect predictions of the minority activation class more heavily.

Shapley fields reveal a stereotypical mosaic organization of chemical receptive fields within the olfactory bulbs. To understand the organization of OB chemical tuning, we constructed Shapley fields. Shapley fields represent signed distributions of feature credit allocations for prediction of glomerular activation, mapped on the surface of the OB. Their computation is analogous to the computation of retinotopic fields (see Shapley fields section in Methods), except instead of using firing rates in response to stimuli we use the Shapley values. In other words, they are a signed projection of odorant chemical feature importance according to the Chemotuning model on the OB. Alternatively, a Shapley field of a chemical feature is the vector of Shapley values for each glomerulus. Shapley fields associated with four of the PubChem features are shown in Fig 2A.

To investigate the spatial organization of the feature importance on the surface of the OB, we computed the 20-dimensional PCA projections of Shapley fields (see Methods) and then performed a hierarchical clustering of glomeruli, based on the values of Shapley fields after the projection assigned to the glomeruli (Fig 2B). The resulting clusters revealed a mosaic organization of glomeruli across the left and the right OBs of an animal. This organization was inferred from the Shapley values of our Chemotuning model. This organization was also observed across animals and with respect to a different chemical feature set (S3A–S3C Fig).

Notably, while the Chemotuning model does not consider glomerular positional data, spatial clustering by chemical feature responsivity nevertheless emerged. This is similar to the clustering previously reported by some studies [10,11,17]. This organization proved highly statistically significant (Fig 2C), as revealed by shuffling the identity of glomeruli ($N = 1000$) and computing spatial-autocorrelation statistics, Moran's I and Geary's C (see Methods) of the original clustering and the clustering after each shuffling. The same was true for a different set of chemical features (S3D Fig). For the sake of clarity, we displayed the analysis using 8 clusters, however any choice of number of clusters between 6 and 10 was statistically significant ($p < 0.05$) with respect to autocorrelation scores of Moran's I and Geary's C (S4 Fig).

We note that our odorant panel, while designed to sample a broad range of chemical space, is relatively small (n = 59). While the mosaic organization we observe is consistent across chemical feature sets, OBs, and animals (S3 Fig), and is robust to the choice of cluster number (S4 Fig), we cannot rule out that it depends in part on the particular odorants sampled here.

The Chemotopy model is more predictive than the Chemotuning model. The Chemotuning model prediction does not take advantage of the position of glomeruli on the OB. If there is a spatial organization based on chemistry, then adding positional information as a feature for learning should improve the model's performance. To this end, we also developed the Chemotopy model. The Chemotopy model assumes that the response of a glomerulus is a function of both position and the chemical features of the odorant and predicts glomerular activation from both the chemical features of the odorant as well as the positional information of the glomerulus. The Chemotopy model is thus a single random forest model for predicting activation of all the observed glomeruli on the OB.

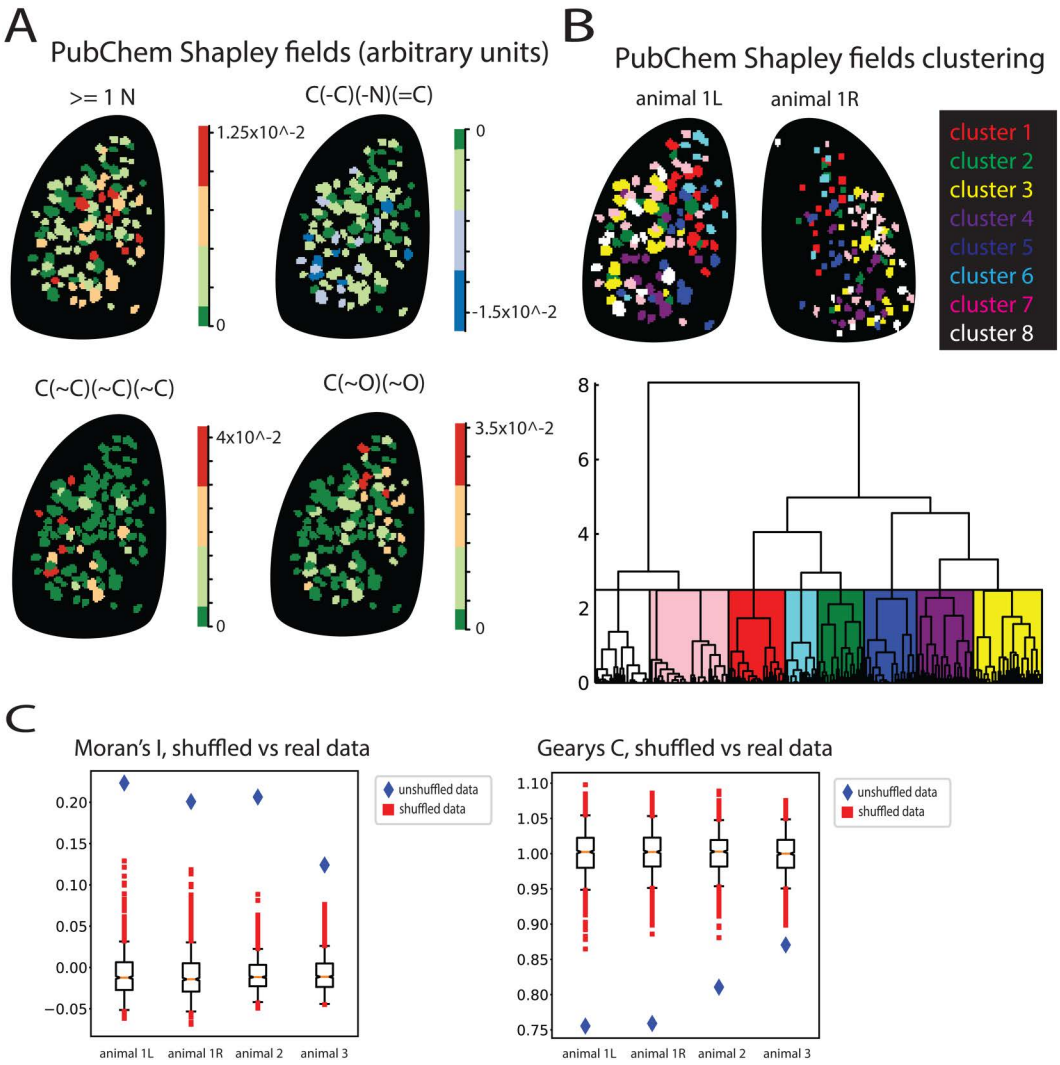


Fig 2. Mosaic chemotopic organization of feature importance on the OB. A. Shapley fields. Importance of four chemical features from the PubChem feature set is displayed on the surface of the left OB, of animal 1, via Shapley fields (Methods). The top right figure is the negative Shapley field of the feature C(-C)(-N)(=C). The rest of the figures are the positive Shapley fields of the respective features. B. Clustering of Shapley fields results in spatially correlated clusters on the surface of OB. Hierarchical clustering of glomeruli on left and right OBs of animal 1, based on the first 20 principal components of Shapley fields, for the Pubchem chemical features. Cosine distance ward linkage clustering was used on the principal components of Shapley fields. Most clusters have a "mosaic" distribution that is roughly consistent between the left and the right OBs. C. The spatial clustering inferred from Shapley fields is statistically significant. Glomeruli positions on the OB were randomly shuffled n times ($n = 1000$). After each shuffling, a glomerulus was assigned to its original cluster from Panel B. The Moran's I and Gearys C clustering scores (Methods), across all 4 datasets show a highly significant difference ($p = 0$) between the clusters from the unshuffled data (blue) and the randomly shuffled ones. Red squares indicate shuffled values falling below $Q1 - 0.75 \times IQR$ or above $Q3 + 0.75 \times IQR$ (IQR = interquartile range).

https://doi.org/10.1371/journal.pcbi.1014917.g002

We compared the predictions of the Chemotuning and the Chemotopy models, as well as two different baseline models. Here we considered the baseline of predicting that the glomerulus would not be activated (because glomerular activation is sparse in the dataset) as well as the peer prediction model (see below). In order to have a relatively small number of model parameters, we used the principal components of the PubChem features ("PCA-PubChem"), while ordering them according to their Shapley values (see Shapley rankings in Methods). The Chemotopy model outperformed

the Chemotuning model when we used the Shapley ranking of the PCA-PubChem features. The Shapley rankings are approximately stable with respect to the randomness inherent in our random forest models (Methods). In particular, S5 Fig shows that the relative rankings are stable with respect to the Kendall Tau correspondence and S6 Fig shows that the identities of the top 10 features are approximately consistent as well.

To provide more context to this comparison, we additionally considered a peer prediction model as the other baseline. Our peer prediction model predicts a given glomerulus's response to an odorant using the responses of the other simultaneously recorded glomeruli to that same odorant, rather than using any chemical features of the odorant itself (see Methods). Because it draws on the full pattern of correlated glomerular activity rather than a necessarily incomplete set of chemical descriptors, its performance provides an approximate empirical upper bound on how well any chemistry-based model (Chemotuning or Chemotopy) could be expected to perform on this dataset.

Our peer prediction model is inspired by [36–39], where it has been observed that one could predict spiking of individual neurons based on the activity of other neurons. In our data, the peer prediction model provided high prediction accuracy (Fig 3) because the olfactory representations are likely redundant for the stimuli that we considered, as should be expected from combinatorial coding of olfactory stimuli in the OB [40–43]. We observed that approximately 70 – 80% of the difference between the peer prediction model's quality of prediction and the baseline is accounted for by the chemical feature sets that we used. Note that our odorant sample size is rather small ($n = 59$) as compared to the number of recorded glomeruli in a single OB ($m = 113,108,128,132$) and that most chemical classes are represented by five or fewer odorants (S10 Fig). Given this, we cannot rule out that the gap between the chemistry based models and the peer prediction model partly reflects limitations of the current odorant panel rather than an intrinsic ceiling on chemistry-based prediction.

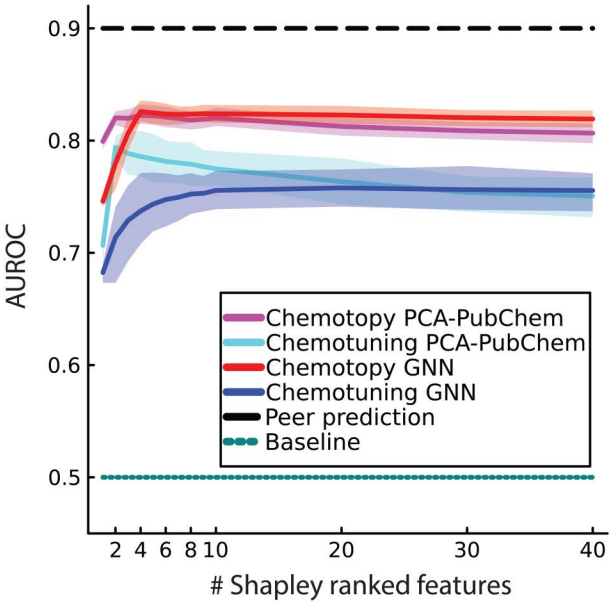


Fig 3. Chemotopy model outperforms chemotuning model. Three different models: chemotuning, chemotopy and the peer prediction are compared vs baseline, across chemical feature sets. The features in a chemical feature set were ranked according to their global Shapley values, largest global Shapley value first (Methods). The features were added, one at a time according to their Shapley ranks, to the two models of glomerular activation prediction from chemical features; chemotopy and chemotuning (Methods). The mean cross-validated AUROC score is plotted along with its 95% confidence interval (shaded regions), $N = 200$, against the number of Shapley ranked chemical features used. The mean cross-validated peer prediction model (black dashed line) AUROC score is also shown as well as the baseline (teal) having value of 0.5 which is the AUROC score of a random binary classifier. The 95% confidence interval for the peer prediction was too narrow to be perceptible on the plot (0.8970 – 0.9026).

https://doi.org/10.1371/journal.pcbi.1014917.g003

Note that the analysis in Fig 3 was for the response data on the left OB of animal 1, however the same observations held for all the other OBs in our data (S7 Fig).

Given the previously discussed discrepancies in reported results, which may arise from the differing chemical feature sets used, we sought to explore whether the superior performance of the Chemotopy versus Chemotuning models was influenced by the specific choice of chemical features. To investigate this, we also did our analysis using a second set of chemical features. We generated these features from a graph neural network (GNN) trained on human derived odor labels of approximately 6000 odorant molecules from the GoodScents company [44], similar to how [45–47] learned odorant representations. We used the same GNN architecture as was reported in [46] (see Methods). The Chemotopy model significantly outperformed the Chemotuning model, with the accuracy gap widening as more Shapley-ranked features were included (Fig 3). For the OB shown in Fig 3, the GNN Chemotopy model reached its maximum AUROC with four Shapley-ranked features, whereas the GNN Chemotuning model showed no comparably clear optimum. Across OBs, the AUROC-optimal number of GNN features varied from 4 to 13, as described below and in S1 Table. Here, Shapley-based feature ranking provided a feature-selection procedure that improved predictive performance by retaining only the most informative features.

The AUROC for the Chemotopy PCA-PubChem, Chemotopy GNN, and Chemotuning PCA-PubChem models plateaued or decreased as additional Shapley-ranked features were included beyond the best-performing subset (Fig 3). This indicates diminishing returns from additional features and, where performance declined, overfitting at higher input dimensionality. Because the number of selected GNN features does not directly measure the number of independent dimensions, we computed the effective dimension of the Shapley-ranked GNN features at the feature count that maximized the Chemotopy-model AUROC for each OB (S1 Table, also see Methods). The resulting effective dimensions were 3.233, 4.733, 4.828, and 5.332 for OBs 1L, 1R, 2R, and 3R, respectively. Thus, across all four OBs, the AUROC-optimal GNN Chemotopy models had an effective dimensionality of at least three ($3.2 \leq d_{\text{eff}} \leq 5.3$). This provides strong evidence that, within the GNN representation, the predictive structure captured by the AUROC-optimal Chemotopy models cannot be accounted for by only one or two effective chemical-feature dimensions. For the Chemotopy PCA-PubChem model, the first 3–6 Shapley-ranked principal components (depending on the hemibulb) similarly captured most of the predictive signal (S7 Fig), whereas additional components increased model complexity without improving cross-validated performance.

We also found that GNN features did not outperform the PCA-PubChem feature set in predicting glomerular activation (Fig 3). PCA-PubChem had greater predictive power when only a small number of Shapley-ranked features were included; however, as more features were added, GNN features matched PCA-PubChem performance in both the Chemotopy and Chemotuning models. This result is consistent with [47], who found that GNN features performed comparably to Mordred chemical features when predicting glomerular responses in the mouse olfactory bulb.

Chemical features derived from human perception yield predictions and mosaic chemotopic organization that is similar to those derived from the PubChem chemical features. How do the chemical receptive properties of the glomeruli depend on the choice of the chemical feature sets? To answer this question, we first computed the Shapley fields of GNN features (Fig 4A). The first principal component of the GNN Shapley fields was highly correlated to the glomerular carboxylic acid sensitivity vector, which counts how many acids each glomerulus responded to (between 0 and 5 in the stimulus set). Tuning to acids was also correlated to several principal components of the PubChem Shapley fields, albeit to a lesser extent (Fig 4B). Note that these are the Shapley fields of PubChem features, and not PCA-PubChem features. This was also true for the OBs where the response to amines was discarded from the analysis (Methods, S8A and S8B Fig).

Pearson correlation analysis also revealed significant correlations among the first ten principal components of both PubChem and GNN Shapley fields (Figs 4C and S8C). To validate the apparent correlations between multiple principal components of GNN and PubChem Shapley fields in Fig 4C, we computed the chordal distances between subspaces generated by up to $k = 20$ principal components of each chemical feature set (see Methods). The chordal distances

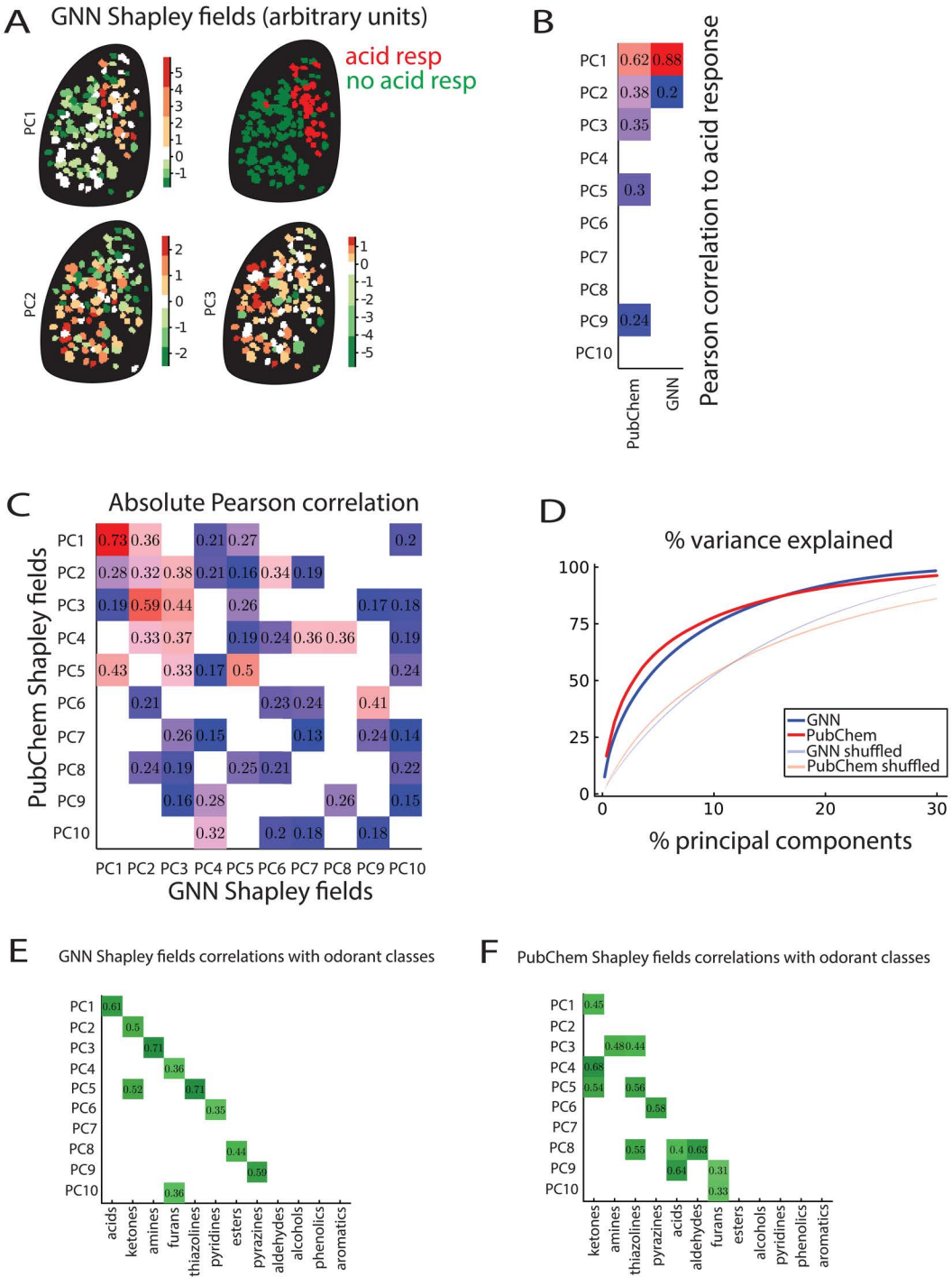


Fig 4. Shapley fields reveal an underlying "universal" chemical organization on the OB. A. Principal components of Shapley fields. Importance of GNN features is projected on the OB via Shapley fields (Methods). First 3 principal components of Shapley fields on the left OB of animal 1, along with a classification of glomeruli into acid responsive (red) and acid non-responsive (green) are displayed. Visually the first principal component of GNN Shapley fields corresponds to acid tuning separation on the OB. B. Interpretation of principal components of Shapley fields. Absolute Pearson correlation values between first 10 principal components of PubChem and GNN Shapley fields with the acid sensitivity vector. The acid sensitivity vector is a count of the number of acids that activate a given glomerulus. Only significant correlations ($p < 0.05$) are shown. GNN Shapley fields PC1 is highly correlated with responsiveness to acids ($r = 0.88$). All the principal components of the GNN Shapley fields beyond PC2 show no significant correlation with acid

sensitivity vector. In contrast, several PubChem Shapley fields principal components show correlation to the acid sensitivity vector to some degree, albeit none as high as the PC1 of GNN Shapley fields. C. The Shapley fields of PubChem features and GNN features are highly correlated. The absolute values of Pearson correlations of GNN's and PubChem's first 10 principal components of the respective Shapley fields are displayed. Only significant correlations ($p < 0.05$) are shown. There are significant degrees of correlation among the 10 principal components of Pubchem and GNN Shapley fields. D. Variance explained by principal components. We find that the first 20 principal components of PubChem Shapley fields account for about 70% of the total variance in Shapley fields (red curve). In contrast, the first 20 principal components of GNN Shapley fields account for about 50% (blue curve). The most probable reason for this difference is the simple difference in the number of Shapley fields; there are 280 PubChem Shapley fields and 512 GNN Shapley fields. The percentage of variance explained is statistically significant when compared to the shuffled Shapley fields principal components, from panel D, for both PubChem (red shaded region) and GNN (blue shaded region). E. Principal components of Shapley fields are correlated with sensitivities of ≤ **2** odorant classes. Principal components of GNN Shapley fields were projected onto the linear subspace spanned by odorant classes sensitivity vectors on the OB (Methods). The projection of principal components of Shapley fields can be written as a linear combination of different odorant classes sensitivity vectors. Statistical analysis involving the shuffling of odorant identities ($N = 10000$) determined which of the coefficients in the projection are statistically significant ($p < 0.05$), see Methods. Only the absolute values of statistically significant coefficients, for each projection, are shown in the table. F. Principal components of PubChem fields are correlated with sensitivities of different odorant classes on the OB. Same analysis as in panel E, but for the PubChem feature set.

https://doi.org/10.1371/journal.pcbi.1014917.g004

between the subspaces of different chemical features are significantly lower than expected from random controls (S9 Fig). Moreover, we found that 20 principal components of the Shapley fields captured 70% of variance for the PubChem feature set and 50% of variance for the GNN feature set respectively (Figs 4D and S8D). This suggests that the Shapley fields represent approximately the same chemical sensitivity irrespective of very different choices of the chemical features that we used: (i) a collection of the chemical properties, and (ii) features derived from human perception and molecular graphs. It is thus likely that Shapley fields capture roughly the same information about receptive properties of the OB, from any diverse enough set of chemical features.

Principal components of Shapley fields recover major chemical classes. The $n = 59$ odorants that we used can be organized into 12 chemical classes (+1 that is the "mixed" class), as defined by major substructural motifs (see S10 Fig for the class assignments). To understand how the Shapley fields relate to this common chemical classification of odorants, we extended our earlier analysis (Fig 4A and 4B) to examine how the principal components of Shapley fields correlated with glomerular responses to all other odorant classes, by considering the odorant class sensitivity vectors. An odorant class sensitivity vector represents the sensitivity of all the recorded glomeruli to a given odorant class (see Methods). We excluded the mixed odorant class sensitivity vector from this analysis as it only contained one odorant, methyl anthranilate. Surprisingly, we found that the individual principal components of GNN Shapley fields are significantly correlated (see Methods) with only one (8 out of 10 first PCs) or two (one out of 10 first PCs) odorant class sensitivity vectors ($p < 0.05$) (Fig 4E). This is not an expected property of the Shapley fields, as GNN features were derived from a graph neural network trained on human perception, combined with molecular graph, yet the first ten principal components of the Shapley fields on the mouse OB often corresponded to a single "common chemical class" used in the set of stimuli. The same analyses applied to the PubChem Shapley fields proved less related to individual odorant chemical classes (Fig 4F), with 4 out of 10 first PCs correlated with just one chemical class.

The above analysis identifies which individual odorant classes contribute significantly to each principal component. A complementary question is how well the spatial pattern of each component can be represented by the odorant-class sensitivity vectors considered together. To assess this, we projected each of the first 10 principal components onto the subspace spanned by all odorant-class sensitivity vectors and calculated the Pearson correlation, together with its associated p-value, between the component and its projection (S2 Table). For the PubChem Shapley fields, these correlations were consistently moderate to strong across all 10 PCs ($0.50 \leq r \leq 0.87$), indicating broad alignment with the class-defined subspace. For the GNN Shapley fields, this alignment was less uniform: PC1 was exceptionally strongly correlated with its projection ($r = 0.93$, $p = 1.0 \times 10^{-96}$), whereas several later PCs had correlations below 0.5. Thus, the two analyses capture different aspects of chemical-class structure. GNN PCs more often correspond selectively to one or two individual classes, whereas PubChem PCs show more consistent alignment with the class-defined subspace as a whole.

Note that the PubChem features show an approximate clustering of the odorant classes used in our experiment, whereas there is barely any clustering by odorant classes in the GNN feature space (S11 Fig). Surprisingly, the GNN Shapley fields showed more pronounced one-to-one correspondence between principal components and chemical classes as compared to the PubChem Shapley fields, despite the lack of significant clustering of the GNN chemical features.

While there was no significant overperformance of the GNN-based prediction over the PubChem feature-based predictions (Fig 3), it seems that the principal components of the GNN Shapley fields are more consistent with the commonly used chemical classification. This raises the possibility that the human descriptors like "spicy" or "fruity" that were used in training the GNN model, together with the mouse OR responses and the molecular graphs used in the GNN, effectively yield the common classification of the odorant molecules (S10 Fig). We speculate that this correspondence reflects a degree of conservation between mouse and human olfactory representations, despite the extensive evolutionary divergence between their olfactory systems ($\sim$ 1,100 functional olfactory receptor genes in the mouse versus $\sim$ 320 in humans). Previous work provides some support for this interpretation. [8] identified a dominant response dimension shared across odor-response datasets from several species: Haddad et al. re-analyzed 12 previously published odor-response datasets from seven species to determine whether the dominant dimensions of population activity were conserved across datasets. They found that the first principal component of neural responses often reflected a common odor axis associated with human pleasantness. More directly, [47] reported that a related GNN feature space, learned from molecular graphs and human perceptual labels, performed comparably to generic chemical representations in predicting mouse OB activity and better in predicting piriform-cortex activity. These findings suggest that representations derived from human olfactory perception can capture structure relevant to mouse olfactory responses, although their relative advantage may differ across processing stages. Our Fig 3 is consistent with this broader interpretation: GNN features did not significantly outperform PubChem features in the mouse OB, but their comparable performance indicates that the human-trained representation remains informative about mouse OB responses.

Discussion

Motivated by investigating odorant chemical feature tuning in the OB, we introduced an application of Shapley fields to explore the sensitivity of glomeruli to chemical features as a function of their position on the dorsal OB surface. The Shapley fields serve as an olfactory analogue of receptive fields in vision and other sensory areas. While chemical response spectra define receptive fields in the odorant space, they do not correspond to receptive fields in the chemical feature space, which Shapley fields aim to characterize as a more nuanced representation of glomerular sensitivity. This analogy is methodological rather than topographic: Shapley fields characterize feature sensitivity across neural tissue, but do not imply that a low-dimensional chemical coordinate system is continuously mapped onto the two-dimensional OB surface.

Using two different computational approaches, we confirmed that the chemical receptive properties of glomeruli in the OB are consistent with a weak chemotopy in the organization of odorant representations across the OB surface. This agrees with earlier studies in [10,12–21] and at least partially addresses the concerns raised in later works [6,22,23].

The analyses in Figs 3 and S7, and S1 Table suggest that the observed chemotopy can *not* be explained by only one or two dominant chemical dimensions. At the feature count that maximized AUROC for each OB, the GNN Chemotopy models had effective dimensionalities between 3.2 and 5.3. The PCA-PubChem Chemotopy models showed a similar pattern, with the first 3–6 Shapley-ranked principal components capturing most of the predictive signal, depending on the hemibulb. Thus, within the present data and modeling framework, attaining optimal Chemotopy performance required an effective dimensionality of at least three. Accordingly, the observed organization is unlikely to be explained solely by one or two known chemical properties, such as the preferential sensitivity of Class-I domain glomeruli to carboxylic acids [48] or the previously described spatial progression of responses to carbon-chain length within aliphatic acid and aldehyde response domains [18,49]. These properties may contribute to the observed organization, but they do not account for its full multidimensional predictive structure.

We therefore interpret these results as evidence for a weak, mosaic chemotopy: chemical-feature sensitivities are spatially non-random, but they overlap and do not form a single continuous map. The combination of significant spatial organization and an effective dimensionality of at least three places the observed organization between complete spatial randomness and a simple one- or two-dimensional chemotopic gradient.

Our analysis characterizes this intermediate, mosaic organization but does not determine how it arises. In particular, the present data cannot establish whether the organization reflects an optimization balancing wiring length against continuity in chemical-feature representation. Testing this possibility would require additional anatomical and developmental information. Shapley fields nevertheless provide quantitative spatial distributions of chemical-feature sensitivity that could be compared with mechanistic models of olfactory map formation in future work.

To investigate how our conclusions are influenced by inherently arbitrary choices of the chemical feature sets, we repeated our analyses using an artificially constructed chemical feature set that combined molecular graph information with human perceptual data, using methods developed in [45–47]. We found that the outperformance of the Chemotopy vs. Chemotuning model does not depend on the choice of chemical feature sets, and that both chemical feature sets describe the receptive properties of the OB glomeruli with comparable accuracy. This is consistent with [47], who similarly found that a random forest model trained on GNN features performed comparably to the same model trained on Mordred chemical features [50], when predicting glomerular responses in the mouse olfactory bulb data set of [22]. In contrast, when activity in a mouse *piriform cortex* was used in [47], it was found that a model using GNN features outperformed the same model using Mordred features, suggesting that the GNN features become more predictive than the standard chemical feature sets in the later stages of the mouse olfactory processing.

Note that our findings are based on roughly $\sim 10\%$ of the total number of distinct olfactory glomeruli in the mouse, and these glomeruli are located on the surface of the dorsal OB. It has been previously established in [51], that the mouse dorsal OB may exhibit somewhat different spatio–temporal dynamics, than ventro-lateral regions, however we hypothesize that qualitatively similar results should hold in the other OB regions with respect to non-random chemotopic properties.

Furthermore, the Shapley fields derived from the two feature sets considered were strongly correlated, suggesting that the wide array of chemical features in PubChem and chemical features derived from human perception cover similar properties of mouse olfactory perception, as represented by the neural activity in the dorsal OB. This suggests that if one chooses a sufficiently diverse set of chemical features, the corresponding Shapley fields would capture the same receptive properties of the OB glomeruli and thus may provide a feature set-agnostic tool for investigating olfactory perception.

We also found that each principal component of the Shapley fields often significantly correlates with a single commonly accepted chemical class used in the presented stimuli. This suggests that these chemical classes could have been "discovered" based on the mouse olfactory responses.

Our findings show that Shapley fields recover shared and interpretable chemical organization from substantially different feature sets, providing a feature-agnostic framework for studying olfactory coding.

Materials and methods

Ethics statement

All procedures were performed following the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of Utah Institutional Animal Care and Use Committee. (IACUC protocols 19–06007, 19–06008).

Data collection: Data were collected from 2 male and 1 female compound heterozygous crosses of OMP-IRES-Cre (Jackson Laboratory stock #006668) and Rosa26-CAG-lox-STOP-lox-GCaMP6f (RCL-GCaMP6f; Ai95D; Jackson Laboratory stock #024105) mouse lines, aged $16-18$ weeks. Prior to data collection mice were housed up to 5 per cage on a 12-hour light/dark cycle with food and water available ad libitum. Odorant responses were collected with epifluorescence imaging in isoflurane-anesthetized mice outfitted with a double tracheotomy to decouple respiration and anesthetic

delivery from odorant presentation and sampling, using procedures identical to those described in [52,53]. The respiratory tracheal tube extended beyond the junction with a surrounding vacuum line, and vacuum flow was maintained above the oxygen/isoflurane carrier flow to minimize leakage of isoflurane into the air surrounding the animal [53]. Although this procedure minimizes direct nasal isoflurane exposure, anesthesia and artificial inhalation could still influence response sensitivity or magnitude. However, Burton et al. found that several functionally identified glomeruli responded to diagnostic odorants at similar picomolar concentrations in awake and isoflurane-anesthetized mice, suggesting that high sensitivity and narrow selectivity can persist across these conditions [10].

Olfactometry: Odorants were obtained from Sigma-Aldrich, TCI America, Bedoukian Research, or ICN Biomedicals and delivered using a custom olfactometer that allowed mixing of odorant vapor in free space prior and computer-controlled timing and choice of odorant identity [52]. Odorant delivery parameters were identical to those used in [10] except that the end-stage eductor was omitted. Odorants were pre-diluted in caprylic/capric medium chain triglyceride oil (C3465, Spectrum Chemical Mfg. Corp.) as described in [10] using dilutions ranging from $1:10-1:50$ in order to achieve estimated delivered concentrations ranging from $1-1000$ nM (median: 100 nM, interquartile range: $20-600$ nM).

Data processing: Odorant response matrices (odorant x glomerulus) were generated using the analysis pipeline described in [10]. Briefly, averaged ΔF response maps were generated from $4-6$ repeated presentations of each odorant, after which regions of interest (ROIs) were selected manually using custom visualization software written in MATLAB (https://github.com/WachowiakLab). ROIs representing presumed individual glomeruli from the dorsal OB were segmented manually, with the segmenter blind to odorant identity [10]. Note that ROIs were drawn either following the rough contours of activation or with stereotyped-sized squares roughly centered within the contours of activation. After segmentation, response matrices were generated using the mean ΔF value across all pixels in an ROI, for all 59 tested odorants.

Response matrices derived from epifluorescence data required further segmentation to minimize effects from signal spread due to scattered or out-of-focus light arising from nearby glomeruli. We verified that responses were accurately attributed to a given ROI by visual inspection. The inspection procedure consisted of sorting odorant response maps in descending order of their ΔF response for a given ROI, and then setting to zero any entries in the response matrix that did not correspond to clear ΔF signal foci that were centered within the ROI of interest. Signals attributable to responses in adjacent ROIs were easily distinguished with this approach. Due to the sparse nature of odorant responses, the inspection algorithm required visual inspection of, typically, 5–20 response maps per ROI. Inspection was performed blind to odorant identity. In rare cases where odorants evoked clear activity that was not easily segmented into glomerular foci (e.g., octanal response maps), no ROI was chosen and these signals were not included for analysis.

Responses to 7 of the amine-class odorants were included *only* for two of the four OBs (1 mouse), due to concern about decomposition and the effective delivered concentrations of these compounds for the remaining two animals.

Random Forests: The random forest binary classifier models were trained using the scikit-learn python library [54]. The hyperparameters we were optimizing for were n_estimators (the number of trees in the forest), max_features (the number of features to consider when looking for the best split) and max_samples (the number of samples to draw from data to train each base estimator). Every other hyperparameter was chosen to be the default one given by the random forest classifier constructor, except the class_weight hyperparameter. Since we are dealing with heavily imbalanced datasets, the class_weight was set to "balanced_subsample." Model performances were measured with the AUROC (Area Under Receiver Operating Characteristic curve). The cross-validation details are described in the relevant model descriptions.

Chemotuning model: The Chemotuning model predicts glomerular activation from the chemical features of the odorants alone. More specifically, given a glomerulus-odorant pair (g,o), a random forest binary classifier predicts activation of g to o based on the chemical features of o. Thus, the Chemotuning model is a collection of random forest models predicting activation, one for each glomerulus on the OB. The hyperparameters for the RF that were optimized with respect to running time and performance were: n_estimators = 128, max_features = "sqrt," max_samples = 0.5, class_weight = "balanced_subsample."

Chemotopy model: The Chemotopy model predicts glomerular activation from the chemical features of the odorants and position information about the glomeruli. More specifically, given a glomerulus-odorant pair (g,o), a random forest binary classifier predicts activation of g to o based on the chemical features of o, and the position of g. Thus, the Chemotopy model is a single random forest model for predicting activation of all the observed glomeruli. To compare with the Chemotuning model (see above) we added an additional feature, which is a random integer identity for each glomerulus. That is, we wanted to make sure the models are trained on the feature space of the same dimension and then compared. The prediction is cross validated across 10 folds and the AUROC metric is used to evaluate performance. The hyperparameters for the random forest that were optimized with respect to running time and performance were as follows: n_estimators = 256, max_features = 1.0, max_samples = 0.7, class_weight = "balanced_subsample."

Peer prediction model: The peer prediction model predicts glomerular activation from the activation of its peers (the other glomeruli on the OB). More specifically, given a glomerulus-odorant pair (g,o), a random forest binary classifier predicts activation of g to o using as features the activations of all glomeruli except g, to o. Thus, the peer prediction model is a collection of random forest models predicting activation, one for each glomerulus on the OB. The model is cross validated in a leave-one-out fashion and the AUROC metric is used to measure performance. The hyperparameters for the random forest were optimized with respect to running time and performance and were the same as the ones for the Chemotuning model.

Position of glomeruli: Each glomerulus was approximated by a minimal rectangle that encloses the region of interest (ROI) associated to the glomerulus. The center of the rectangle was taken to be the position of the glomerulus. For each OB, the x and y coordinates of the glomeruli were normalized to both have mean 0 and variance 1.

The PubChem Substructure Fingerprint: The PubChem System generates a binary substructure fingerprint for chemical structures [30]. When considering unique and nonzero PubChem fingerprints for our datasets we ended up with 140 features for the 59 odorants dataset and 132 features for the 52 odorants dataset.

PCA PubChem features: PCA dimensionality reduction was performed on the PubChem Substructure Fingerprint binary features with 59 and 52 odorants. Both were projected to 40 dimensions. The first 40 principal components accounted for around 90% of the variance in both datasets.

GNN features: We used a dataset of 5955 small molecules that were made publicly available, after pre-processing, by the authors in [46]. Each molecule comes with a set of human perceptual odor labels. Originally these molecules, with their odor labels, were obtained from The Good Scents Company (TGSC) Database [44], and Leffingwell Database [55]. The 160 odor labels were grouped into 21 categories and used for training a multitask binary classification Graph Neural Network (GNN) [56,57]. We trained a multitask binary classification GNN using the DeepChem library [58] and the same network architecture as the one specified in [46]. Furthermore, we cross validated our GNN model across 5 folds and evaluated its performance with the AUPRC (Area Under Precision-Recall Curve) score. Our mean AUPRC score of approximately 0.40 agreed with the reported AUPRC score in [46]. We then trained a network by using the whole human perceptual dataset of 5955 odorants. Inspired by the idea of transfer learning [59], we produced learned GNN features for our smaller dataset of odorants. That is, we used the learned weights in the network to produce human-derived learned features for our odorants, by considering their representations in the penultimate layer of the network, before the multitask softmax classification layer. We call these the GNN features. There were 256 GNN features in total. All the odorants in our datasets were present in the human perceptual dataset, except for 7; 6 of the 7 amines and methyl salicylate.

Shapley values: Shapley values are a game theoretic approach to explain the outputs of machine learning models and we used the SHAP (SHapley Additive exPlanations) python library to compute them [24]. Shapley values calculate the optimal credit allocation to features which locally explains the model outputs (Fig 1C). This is done by treating features as players in a cooperative game, and the model output as the game prize and then calculating the classic Shapley values from game theory [25]. For an arbitrary model, computing Shapley values has exponential time complexity and in practice Shapley values are only approximated. Instead, we used shap.TreeExplainer which uses dynamic programming to exactly compute Shapley values for tree-based models, in polynomial time [60].

Shapley rankings: Because Shapley values are associated with a given model, and the Chemotuning model is in fact a collection of individual models, one for each glomerulus, we ranked the chemical features using the Chemotopy model. Let $F = \{f_i\}_{i=1}^N$ be any of the sets of chemical features that were described above. For a given OB, we trained a Chemotopy random forest model, M, on every glomerulus-odorant pair (g,o) (no test/train split, no cross-validation) that predicted glomerular activation based on the position of g and chemical features of o. Let f_i denote a chemical feature in F, and $\varphi_i(g,o)$ denote the Shapley value of f_i, which gives the contribution of $f_i(o)$ towards the prediction $M(g,o)$. The global Shapley value of f_i is the mean absolute value of Shapley values of f_i, over all observations (g,o), that is

$$\Phi_i = \frac{\sum_{g,o} |\varphi_i(g,o)|}{|G||O|},$$

where $|G|$ is the number of glomeruli on the OB, and $|O|$ is the number of odorants. We thus have a collection $\{\Phi_i\}_{i=1}^N$ of non-negative numbers that we rank in decreasing order. We call these ranks the Shapley rankings of the set of chemical features F. Note that we did not use a cross-validated training here, because we observed that the ranks didn't change significantly, and the cross-validation procedure used for all the analyses in Fig 3 already provided a check against overfitting.

Due to the randomness inherent to random forest models, a natural concern is the stability of these ranks. However, we found the Shapley ranks are for the most part stable with respect to the random choices of random forest models (S5 and S6 Figs). The Shapley ranks of PubChem features were more stable than the ranks of GNN features. This is most likely due to GNN being a larger feature set (256 dimensions) as opposed to the PubChem features (140 or 132 dimensions) and hence there was more possibility of ranks changing during random forest training. The stability analysis was done by considering the Kendall rank correlation coefficients [61] of Shapley rankings from different random seeds.

Shapley fields: Given an OB, and a set of chemical features $F = \{f_i\}_{i=1}^N$, for each glomerulus g we train a Chemotuning model, T_g, using the chemical features F (no train/test split, no cross-validation). Let $\varphi_i(g,o)$ be the Shapley value of f_i, which gives the contribution of $f_i(o)$ towards the prediction $T_g(o)$. Let O_g^+, O_g^- be the collections of odorants for which $\varphi_i(g,o) \geq 0$, $\varphi_i(g,o) < 0$, respectively. Let

$$\Phi_i^+(g) = \frac{1}{|O_g^+|} \sum_{o \in O_g^+} \varphi_i(g,o),\ \Phi_i^-(g) = \frac{1}{|O_g^-|} \sum_{o \in O_g^-} \varphi_i(g,o),$$

where $|O_g^+|$, $|O_g^-|$ are the numbers of odorants in O_g^+ and O_g^-, respectively. This construction yields, for each feature f_i, two distributions on the OB which we call the positive and negative Shapley fields associated to f_i. Intuitively, the Shapley fields represent the distribution of feature importance on the OB, very much analogous to receptive fields; one can also think of them as signed measures on the OB. We could have done a cross-validated training and made sure we add positive and negative Shapley values appropriately from different testing folds, and then define positive and negative Shapley fields. However, we observed that the values of the Shapley fields don't change significantly and thus the additional computation time, which is not insignificant, would yield diminishing returns.

Hierarchical clustering of Shapley fields: We concatenate the positive and negative Shapley fields, calculated from a fixed set of chemical features, from different OBs. Given the concatenated Shapley fields $\Phi_i^+(g), \Phi_i^-(g)$, we perform a principal component dimensionality reduction and project to the first 20 principal components, which accounted for 70% (PubChem) or 50% (GNN) of the variance, see Figs 4D and S8D. After the projection, we perform hierarchical ward linkage clustering with the cosine distance. The resulting dendrogram (Figs 2B and S3B) and a height parameter are used to assign glomeruli to clusters.

Spatial autocorrelation metrics: We quantified spatial autocorrelation using *Moran's I* and *Geary's C*, two complementary statistics that characterize global spatial clustering and local spatial heterogeneity, respectively [62,63]. Moran's I has been advocated as a statistically rigorous measure of spatial coherence in neural imaging data [64].

Given a particular clustering of N glomeruli, we denote by W_{ij} the $N \times N$ binary spatial weights matrix with zeroes on the diagonal, where for any $i \neq j$, $W_{ij} = 1$ if the i-th and j-th glomeruli belong to the same cluster. Let $X \in \mathbb{R}^{2\times N}$ be the matrix whose i-th column x_i is the 2D position of the i-th glomerulus, let $\bar{x} \in \mathbb{R}^{2\times 1}$ be the column vector obtained by averaging each row of X and define $\bar{x}_i = x_i - \bar{x}$ for $i = 1, ..., N$. Let $S = \sum_{i,j} W_{ij}$ be the sum of all entries of W. Then the Moran's I and Geary's C spatial autocorrelations are defined as follows:

$$I = \frac{N}{S}\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\bar{x}_i^T\bar{x}_j}{\sum_{i=1}^{N}\bar{x}_i^T\bar{x}_i}, \quad C = \frac{N-1}{2S}\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}(x_i-x_j)^T(x_i-x_j)}{\sum_{i=1}^{N}\bar{x}_i^T\bar{x}_i}.$$

Effective dimension: Effective dimension of the feature space was computed to account for possible correlations among the non-linear GNN features. Given N features, the effective dimension was computed as $d_{\text{eff}} = D_{PR}$, where D_{PR} is the participation ratio dimension [65], computed as follows. Let R denote the $N \times N$ correlation matrix of the chemical features across the observed odorants, and λ_i denote the N eigenvalues of R. Then the participation ratio dimension,

$$D_{PR} = \frac{\left(\sum_{i=1}^{N}\lambda_i\right)^2}{\sum_{i=1}^{N}\lambda_i^2},$$

estimates the number of degrees of freedom explained by the variance in the dataset. Note that $1 \leq D_{\text{PR}} \leq N$, interpolating between two extremes of having only one independent feature, and having N independent features.

Odorant class sensitivity vectors: For each odorant class in our dataset, except "mixed," (12 odorant classes in total) we define its odorant class sensitivity vector. It is a vector on the OB, meaning its coordinates are indexed by the glomeruli on the OB, and it counts to how many odorants from the given odorant class a given glomerulus was sensitive to. For example, if the odorant class in question is the acids, then the acid sensitivity vector counts for each glomerulus on the OB, to how many acids the glomerulus responded to. Since we used 5 acids in our experiments, this vector assigns values between 0 and 5 to each glomerulus.

Correlations of Shapley fields principal components to odorant class sensitivity vectors: A principal component of Shapley fields from a chemical feature set is represented by a vector whose indices correspond to the glomeruli on the OB. This vector therefore lives in an n-dimensional space, where n is the number of glomeruli. The 12 odorant classes sensitivity vectors are also vectors in this n-dimensional space, and they span a 12-dimensional linear subspace. We measured the correlation between a principal component of Shapley fields and its projection to this linear subspace. The projection is a linear combination of odorant class tuning vectors with some weights. We calculated the statistical significance of these weights ($p < 0.05$) by randomly permutating the odorant identities that define the odorant class sensitivity vectors ($N = 10000$ trials). The statistically significant coefficients gave us the odorant classes' tuning vectors the principal component of Shapley fields is correlated to.

Distances between linear subspaces of Shapley fields: To compare Shapley fields induced by different chemical feature sets, we compared the linear subspaces spanned by k principal components they generate, where $1 \leq k \leq 20$. Here, each principal component is a high-dimensional vector whose coordinates are its values on each of the glomeruli (see Fig 4A). These k-dimensional subspaces can be represented as a matrix whose columns are the principal components. Given Shapley fields from different chemical feature sets we can then compare distances between these resulting linear subspaces, as we vary the PCA dimension k we are projecting to. The distances between these subspaces are calculated via principal angles, described below.

Principal angles between subspaces: Angles between vectors in an n-dimensional space can be determined from the cosine-dot-product formula, i.e., if θ is the angle between vectors x and y we have $\cos\theta = \frac{x\cdot y}{\|x\|\|y\|}$. Treating vectors as

the one-dimensional subspaces they span; this angle represents the angle between these one-dimensional subspaces. This idea extends to higher dimensions, and it is possible to define so-called principal angles between linear subspaces of any dimension. Specifically, given real-valued ($n \times k$) matrices A and B, each one with pairwise orthogonal column vectors of length 1, one can find k principal angles $\theta_1 \leq \theta_2 \leq ... \leq \theta_k$ between the linear subspaces that are spanned by the column vectors of A and B. In practice the principal angles can be computed from the singular values of the matrix A^TB [66]. That is given σ_i, the i-th singular value of A^TB, we define $\cos\theta_i = \sigma_i$.

Chordal distance between subspaces: Given linear subspaces V and W of dimension p, let $\theta_1 \leq \theta_2 \leq ... \leq \theta_p$ be the principal angles between V and W. The chordal distance between V and W is the quantity $d_c(V, W) = (\sum_{i=1}^{p} \sin^2\theta_i)^{1/2}$, [67]. If we write A and B for matrices of column vectors of orthonormal basis vectors that generate V and W, respectively, by definition $\cos^2\theta_i = 1 - \sin^2\theta_i = \sigma_i^2$, where σ_i is the i-th singular value of $C = A^TB$. In particular, $\sum_i \sigma_i^2 = \text{trace}(C^TC)$. From this it follows that $d_c(V, W) = (p - \text{trace}(C^TC))^{1/2}$.

Supporting information

S1 Fig. Consistency of odorant responses across OBs. A. Odorant panel is representative of odorant space in general. Distribution of 59 odorants used in the analysis, plotted in the first two principal components of a physicochemical descriptor matrix. Grey points, 2587 odorants. Red points, tested odorants. B. Odorant responses across OBs. Odorant response maps from all four olfactory bulbs used in the analysis (3 mice), showing responses to 3 odorants. Concentration values indicate estimated delivered odorant concentration. C. The ROIs used in the study. The ROIs used in our analysis for all OBs.
(PDF)

S2 Fig. Odorant response profiles of glomeruli in the left OB of animal 1. The heat map shows the response magnitude of every identified glomerulus (rows) to each of the 59 odorants (columns). Colors indicate relative response magnitude on the displayed 0–1 scale; responses judged absent during response-matrix segmentation were set to 0 (Methods). The response profile of glomerulus 75, whose Chemotuning-model prediction for butyl acetate is explained in Fig 1C, is outlined in white.
(PDF)

S3 Fig. Mosaic chemotopic organization of feature importance on the OB. A. The clustering in the space of PubChem Shapley fields results in spatially correlated clusters on the surface of OB. Hierarchical clustering of glomeruli on right OBs of animals 2 and 3, based on the first 20 principal components of Shapley fields, for the Pubchem chemical features. Cosine distance ward linkage clustering was used on the principal components of Shapley fields. A cutoff for clustering purposes was chosen as shown on the dendrogram. Most clusters have a “mosaic” distribution that is roughly consistent between the OBs. B and C. The clustering in the space of GNN Shapley fields results in spatially correlated clusters on the surface of OB. Hierarchical clustering of glomeruli on all OBs in our dataset based on the first 20 principal components of Shapley fields, for the GNN chemical features. Cosine distance ward linkage clustering was used on the principal components of Shapley fields. A cutoff for clustering purposes was chosen as shown on the dendrogram. Most clusters have a “mosaic” distribution that is roughly consistent between the OBs. D. The spatial clustering inferred from GNN Shapley fields is statistically significant. Glomeruli positions on the OB were randomly shuffled n times ($n = 1000$). After each shuffling, a glomerulus was assigned to its original cluster from Panels B and C. The Moran’s I and Gearys C clustering scores (Methods), across all 4 datasets show a highly significant difference ($p < 0.05$) between the clusters from the data (blue) and the randomly shuffled ones. Red squares indicate shuffled values falling below $Q1 - 0.75 \times IQR$ or above $Q3 + 0.75 \times IQR$ (IQR = interquartile range).
(PDF)

S4 Fig. Statistical significance of clustering of Shapley fields. A. The spatial clustering inferred from PubChem Shapley fields is statistically significant. Hierarchical clustering of glomeruli, based on the first 20 principal components of Shapley fields, for the Pubchem chemical features. Cosine distance ward linkage clustering was used on the principal components of Shapley fields. Different numbers of resulting clusters were considered (6,7,9,10 from top to bottom). Glomeruli positions on the OB were randomly shuffled n times ($n = 1000$). After each shuffling, a glomerulus was assigned to its original cluster. The Moran's I and Gearys C clustering scores (Methods), across all 4 datasets show a highly significant difference ($p < 0.05$) between the clusters from the data (blue) and the randomly shuffled ones. Red squares indicate shuffled values falling below $Q1 - 0.75 \times IQR$ or above $Q3 + 0.75 \times IQR$ (IQR = interquartile range). B. The spatial clustering inferred from GNN Shapley fields is statistically significant. Same analysis as in Panel A, but for the GNN Shapley fields.
(PDF)

S5 Fig. Stability of Shapley ranks of chemical feature sets. A. Kendall rank correlation coefficients of Shapley rankings of PubChem features. Chemotopy model (see Methods) was trained 100 times (100 different random seeds), and the Shapley rankings of PubChem features used by the model were computed (see Methods). To analyze the stability of rankings (how much they change between random seeds) we computed the Kendall rank correlation coefficients between rankings. This comparison was done for Shapley rankings of PubChem features on the same OB (red), same animal but different OBs (blue), and on different OBs in different animals (green). Each violin plot contains 10000 such rank comparisons. The high values of Kendall rank correlation coefficients (the y-axis) indicate the stability of Shapley ranks of features. B. Kendall rank correlation coefficients of Shapley rankings of GNN features. Same analysis as in A, but for the the GNN feature set. Overall there is less stability for GNN Shapley ranks than for the PubChem Shapley ranks.
(PDF)

S6 Fig. Consistency of Shapley ranks of chemical feature sets. Chemotopy model (Methods) was trained 100 times (100 different random seeds), and the Shapley rankings of PubChem features used by the model were computed (Methods). We examined which features were ranked top 10 across different random seeds and counted how many times a feature was consistently ranked in the same position (among the top 10). The y-axis on the histograms show the percentage of times a feature was in a particular rank and the bars indicate the rank; 1–10 going from left to right. The feature id is on the x-axis. This was done for every OB in the experiment. Certain feature ids being present in high percentage of times in the same top 10 rankings indicate the consistency of Shapley ranks. The OBs of animal 1 are most similar to one another and so are the right OBs of animals 2 and 3. However, across these pairs there is less consistency. This is likely due to the fact that OBs of animal 1 had the amines responses included, whereas the OBs of animals 2 and 3 did not (Methods).
(PDF)

S7 Fig. Chemotopy model outperforms chemotuning model. A. Three different models: chemotuning, chemotopy and the peer prediction are compared vs baseline, across chemical feature sets. The features in a chemical feature set were ranked according to their global Shapley values, largest global Shapley value first (Methods). The features were added, one at a time according to their Shapley ranks, to the two models of glomerular activation prediction from chemical features; chemotopy and chemotuning (Methods). The mean cross-validated AUROC score is plotted along with its 95% confidence interval (shaded regions) against the number of Shapley ranked chemical features used. The mean cross-validated peer prediction model (black dashed line) AUROC score is also shown as well as the baseline (teal) having value of 0.5 which is the AUROC score of a random binary classifier. The 95% confidence interval for the peer prediction was too narrow to be perceptible on the plot (0.8913–0.8962). The data models performances shown are for the right OB of animal 1. B. Three different models: chemotuning, chemotopy and the peer prediction are compared vs baseline, across chemical feature sets. The same as in Panel A, but for the right OB of animal 2. C. Three different models: chemotuning, chemotopy and the peer prediction are compared vs baseline, across chemical feature sets. The same as in Panel A, but for the right OB of animal 3.
(PDF)

S8 Fig. Shapley fields reveal an underlying "universal" chemical organization on the OB. A. Principal components of Shapley fields. Importance of GNN features is projected on the OB via Shapley fields (Methods). First 3 principal components of Shapley fields on the right OB of animal 2, along with a classification of glomeruli into acid responsive (red) and acid non-responsive (green) are displayed. Visually the first principal component of GNN Shapley fields corresponds to acid tuning separation on the OB. B. Interpretation of principal components of Shapley fields. Absolute Pearson correlation values between first 10 principal components of PubChem and GNN Shapley fields with the responsiveness to acids vector. The responsiveness to acids vector is a count of the number of acids that activate a given glomerulus. Only significant correlations ($p < 0.05$) are shown. GNN Shapley fields PC1 is highly correlated with responsiveness to acids ($r = 0.86$). All the principal components of the GNN Shapley fields beyond PC2 show no significant correlation with acid response. In contrast, several PubChem Shapley fields principal components show correlation to responsiveness to acids to some degree, albeit none as high as the PC1 of GNN Shapley fields. C. The Shapley fields of PubChem features and GNN features are highly correlated. The absolute values of Pearson correlations of GNN's and PubChem's first 10 principal components of the respective Shapley fields are displayed. Only significant correlations ($p < 0.05$) are shown. There are significant degrees of correlation among the 10 principal components of Pubchem and GNN Shapley fields. D. Variance explained by principal components. We find that the first 20 principal components of PubChem Shapley fields account for about 69% of the total variance in Shapley fields (red curve). In contrast, the first 20 principal components of GNN Shapley fields account for about 47% (blue curve). The most probable reason for this difference is the simple difference in the number of Shapley fields; there are 264 PubChem Shapley fields and 512 GNN Shapley fields. The percentage of variance explained is statistically significant when compared to the shuffled Shapley fields principal components, from panel D, for both PubChem (red shaded region) and GNN (blue shaded region).
(PDF)

S9 Fig. Human-derived chemical features and molecular fingerprints share chemical organization on the OB. A. Distances between linear subspaces of PCs of GNN and PubChem Shapley fields. Principal angles and the resulting chordal distance between GNN and PubChem Shapley field principal component subspaces (Methods) were computed. We only considered the first 20 principal components of both GNN and PubChem Shapley fields and computed the principal angles of the linear subspaces they span, starting with PC1s and adding the other PCs one at a time. The resulting distance vs dimension curve is shown in red. We shuffled the activation patterns on both OBs of animal 1 ($N = 1000$), and recomputed both GNN and PubChem Shapley fields along with their chordal distances of subspaces generated by the first 20 principal components. The resulting curve (blue) with its 95% confidence interval shows that the real chordal distance between GNN and PubChem Shapley fields is statistically significant. B. Distances between linear subspaces of PCs of GNN and PubChem Shapley fields. Same analysis as in Panel A but for right OBs of animals 2 and 3.
(PDF)

S10 Fig. Odorant names and odorant classes used in experiments. Responses to the amine-class odorants were included only for two of the four OBs (1 mouse), as explained in Methods.
(PDF)

S11 Fig. PCA projections of chemical feature sets. A. PCA projections of the PubChem feature set. We reduced the dimension of the PubChem feature space of the 59 odorants used in our experiments to 2 dimensions, using PCA. Each odorant is represented as a point in the 2 dimensional space, with a color representing its odorant class. B. PCA projections of the GNN feature set. Same as in Panel A, but for the GNN feature set.
(PDF)

S1 Table. Effective dimensions of GNN feature subsets. For each OB and each subset containing the top N Shapley-ranked GNN features ($N = 2, \ldots, 13$), we report the mean cross-validated AUROC of the Chemotopy model

(Figs 3 and S7) and the effective dimension, d_{eff}, of the corresponding feature subset, calculated using the participation ratio (Methods). For each OB, red values identify the maximum AUROC and the effective dimension at the corresponding feature count. All values are rounded to three decimal places. At the AUROC-maximizing feature counts, d_{eff} ranged from 3.233 to 5.332 across OBs, providing evidence that these optimal GNN feature subsets are not reducible to only one or two linearly independent dimensions within the GNN representation.
(PDF)

S2 Table. Correlations of Shapley fields to odorant classes sensitivity vectors. Principal components of Shapley fields were projected onto the linear subspace spanned by odorant classes sensitivity vectors on the OB (Methods). The Pearson correlations between the principal components and their respective projections are shown in the table, along with the p-value of the correlation.
(PDF)

Acknowledgments

The authors would like to thank Michael Schmuker for helpful discussions and suggestions.

Author contributions

Conceptualization: Nikola Milićević, Matt Wachowiak, Vladimir Itskov.

Data curation: Nikola Milićević, Shawn D. Burton, Matt Wachowiak, Vladimir Itskov.

Formal analysis: Nikola Milićević, Vladimir Itskov.

Funding acquisition: Matt Wachowiak, Vladimir Itskov.

Investigation: Shawn D. Burton, Matt Wachowiak, Vladimir Itskov.

Methodology: Nikola Milićević, Shawn D. Burton, Vladimir Itskov.

Project administration: Matt Wachowiak, Vladimir Itskov.

Resources: Shawn D. Burton, Vladimir Itskov.

Software: Nikola Milićević.

Supervision: Vladimir Itskov.

Validation: Nikola Milićević, Vladimir Itskov.

Visualization: Nikola Milićević, Vladimir Itskov.

Writing – original draft: Nikola Milićević, Matt Wachowiak, Vladimir Itskov.

Writing – review & editing: Nikola Milićević, Vladimir Itskov.

References

1. Chen X, Gabitto M, Peng Y, Ryba NJP, Zuker CS. A gustotopic map of taste qualities in the mammalian brain. Science. 2011;333(6047):1262–6. https://doi.org/10.1126/science.1204076 PMID: 21885776
2. Hubel DH, Wiesel TN. Ferrier lecture - functional architecture of macaque monkey visual cortex. Proc R Soc London Ser B Biol Sci. 1977;198(1130):1–59.
3. Kandler K, Clause A, Noh J. Tonotopic reorganization of developing auditory brainstem circuits. Nat Neurosci. 2009;12(6):711–7. https://doi.org/10.1038/nn.2332 PMID: 19471270
4. Koulakov AA, Kolterman BE, Enikolopov AG, Rinberg D. In search of the structure of human olfactory space. Front Syst Neurosci. 2011;5:65. https://doi.org/10.3389/fnsys.2011.00065 PMID: 21954378
5. Keller A, Vosshall LB. Olfactory perception of chemically diverse molecules. BMC Neurosci. 2016;17(1):55. https://doi.org/10.1186/s12868-016-0287-2 PMID: 27502425

6. Ma L, Qiu Q, Gradwohl S, Scott A, Yu EQ, Alexander R, et al. Distributed representation of chemical features and tunotopic organization of glomeruli in the mouse olfactory bulb. Proc Natl Acad Sci U S A. 2012;109(14):5481–6. https://doi.org/10.1073/pnas.1117491109 PMID: 22431605
7. Haddad R, Khan R, Takahashi YK, Mori K, Harel D, Sobel N. A metric for odorant comparison. Nat Methods. 2008;5(5):425–9. https://doi.org/10.1038/nmeth.1197 PMID: 18376403
8. Haddad R, Weiss T, Khan R, Nadler B, Mandairon N, Bensafi M, et al. Global features of neural activity in the olfactory system form a parallel code that predicts olfactory behavior and perception. J Neurosci. 2010;30(27):9017–26. https://doi.org/10.1523/JNEUROSCI.0398-10.2010 PMID: 20610736
9. Saito H, Chi Q, Zhuang H, Matsunami H, Mainland JD. Odor coding by a Mammalian receptor repertoire. Sci Signal. 2009;2(60):ra9. https://doi.org/10.1126/scisignal.2000016 PMID: 19261596
10. Burton SD, Brown A, Eiting TP, Youngstrom IA, Rust TC, Schmuker M, et al. Mapping odorant sensitivities reveals a sparse but structured representation of olfactory chemical space by sensory input to the mouse olfactory bulb. eLife. 2022;11:e80470. https://doi.org/10.7554/eLife.80470 PMID: 35861321
11. Mori K, Takahashi YK, Igarashi KM, Yamaguchi M. Maps of odorant molecular features in the Mammalian olfactory bulb. Physiol Rev. 2006;86(2):409–33. https://doi.org/10.1152/physrev.00021.2005 PMID: 16601265
12. Johnson BA, Woo CC, Hingco EE, Pham KL, Leon M. Multidimensional chemotopic responses to n-aliphatic acid odorants in the rat olfactory bulb. J Comp Neurol. 1999;409(4):529–48. https://doi.org/10.1002/(sici)1096-9861(19990712)409:4<529::aid-cne2>3.0.co;2-n PMID: 10376738
13. Johnson BA, Leon M. Odorant molecular length: one aspect of the olfactory code. J Comp Neurol. 2000;426(2):330–8. https://doi.org/10.1002/1096-9861(20001016)426:2<330::aid-cne12>3.0.co;2-5 PMID: 10982472
14. Meister M, Bonhoeffer T. Tuning and topography in an odor map on the rat olfactory bulb. J Neurosci. 2001;21(4):1351–60. https://doi.org/10.1523/JNEUROSCI.21-04-01351.2001 PMID: 11160406
15. Rubin BD, Katz LC. Optical imaging of odorant representations in the mammalian olfactory bulb. Neuron. 1999;23(3):499–511. https://doi.org/10.1016/s0896-6273(00)80803-x PMID: 10433262
16. Johnson BA, Farahbod H, Leon M. Interactions between odorant functional group and hydrocarbon structure influence activity in glomerular response modules in the rat olfactory bulb. J Comp Neurol. 2005;483(2):205–16. https://doi.org/10.1002/cne.20409 PMID: 15678471
17. Johnson BA, Leon M. Chemotopic odorant coding in a mammalian olfactory system. J Comp Neurol. 2007;503(1):1–34. https://doi.org/10.1002/cne.21396 PMID: 17480025
18. Matsumoto H, Kobayakawa K, Kobayakawa R, Tashiro T, Mori K, Sakano H, et al. Spatial arrangement of glomerular molecular-feature clusters in the odorant-receptor class domains of the mouse olfactory bulb. J Neurophysiol. 2010;103(6):3490–500. https://doi.org/10.1152/jn.00035.2010 PMID: 20393058
19. Soelter J, Schumacher J, Spors H, Schmuker M. Computational exploration of molecular receptive fields in the olfactory bulb reveals a glomerulus-centric chemical map. Sci Rep. 2020;10(1):77. https://doi.org/10.1038/s41598-019-56863-4 PMID: 31919393
20. Takahashi YK, Kurosaki M, Hirono S, Mori K. Topographic representation of odorant molecular features in the rat olfactory bulb. J Neurophysiol. 2004;92(4):2413–27. https://doi.org/10.1152/jn.00236.2004 PMID: 15152015
21. Uchida N, Takahashi YK, Tanifuji M, Mori K. Odor maps in the mammalian olfactory bulb: domain organization and odorant structural features. Nat Neurosci. 2000;3(10):1035–43. https://doi.org/10.1038/79857 PMID: 11017177
22. Chae H, Kepple DR, Bast WG, Murthy VN, Koulakov AA, Albeanu DF. Mosaic representations of odors in the input and output layers of the mouse olfactory bulb. Nat Neurosci. 2019;22(8):1306–17. https://doi.org/10.1038/s41593-019-0442-z PMID: 31332371
23. Soucy ER, Albeanu DF, Fantana AL, Murthy VN, Meister M. Precision and diversity in an odor map on the olfactory bulb. Nat Neurosci. 2009;12(2):210–20. https://doi.org/10.1038/nn.2262 PMID: 19151709
24. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al, editors. Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc.; 2017. p. 4765–74.
25. Shapley LS. Notes on the N-person game—II: The value of an N-person game. Research Memorandum. RM-670-PR. Santa Monica (CA): RAND Corporation; 1951.
26. Vincis R, Gschwend O, Bhaukaurally K, Beroud J, Carleton A. Dense representation of natural odorants in the mouse olfactory bulb. Nat Neurosci. 2012;15(4):537–9. https://doi.org/10.1038/nn.3057 PMID: 22406552
27. Pashkovski SL, Iurilli G, Brann D, Chicharro D, Drummey K, Franks KM, et al. Structure and flexibility in cortical representations of odour space. Nature. 2020;583(7815):253–8. https://doi.org/10.1038/s41586-020-2451-1 PMID: 32612230
28. Ressler KJ, Sullivan SL, Buck LB. Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell. 1994;79(7):1245–55. https://doi.org/10.1016/0092-8674(94)90015-9 PMID: 7528109
29. Wachowiak M, Zochowski M, Cohen LB, Falk CX. The spatial representation of odors by olfactory receptor neuron input to the olfactory bulb is concentration invariant. Biol Bull. 2000;199(2):162–3. https://doi.org/10.2307/1542876 PMID: 11081714
30. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2023 update. Nucleic Acids Res. 2023;51(D1):D1373–80. https://doi.org/10.1093/nar/gkac956 PMID: 36305812

31. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. NPJ Digit Med. 2021;4(1):3. https://doi.org/10.1038/s41746-020-00372-6 PMID: 33398013
32. Rodríguez-Pérez R, Bajorath J. Interpretation of Compound Activity Predictions from Complex Machine Learning Models Using Local Approximations and Shapley Values. J Med Chem. 2020;63(16):8761–77. https://doi.org/10.1021/acs.jmedchem.9b01101 PMID: 31512867
33. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. ACM; 2016. p. 1135–44.
34. Shrikumar A, Greenside P, Kundaje A. Learning important features through propagating activation differences. In: Precup D, Teh YW, editors. Proceedings of the 34th International Conference on Machine Learning, volume 70 of Proceedings of Machine Learning Research. PMLR; 2017. p. 3145–53.
35. Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. PLoS One. 2015;10(7):e0130140. https://doi.org/10.1371/journal.pone.0130140 PMID: 26161953
36. Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsáki G. Organization of cell assemblies in the hippocampus. Nature. 2003;424(6948):552–6. https://doi.org/10.1038/nature01834 PMID: 12891358
37. Itskov V, Curto C, Harris KD. Valuations for spike train prediction. Neural Comput. 2008;20(3):644–67. https://doi.org/10.1162/neco.2007.3179 PMID: 18045025
38. McKenzie S, Huszár R, English DF, Kim K, Christensen F, Yoon E, et al. Preexisting hippocampal network dynamics constrain optogenetically induced place fields. Neuron. 2021;109(6):1040–54.e7. https://doi.org/10.1016/j.neuron.2021.01.011 PMID: 33539763
39. Luczak A, McNaughton BL, Kubo Y. Neurons learn by predicting future activity. Nat Mach Intell. 2022;4(1):62–72. https://doi.org/10.1038/s42256-021-00430-y PMID: 35814496
40. Malnic B, Hirono J, Sato T, Buck LB. Combinatorial receptor codes for odors. Cell. 1999;96(5):713–23. https://doi.org/10.1016/s0092-8674(00)80581-4 PMID: 10089886
41. Koulakov A, Gelperin A, Rinberg D. Combinatorial on/off model for olfactory coding. 2006.
42. de March CA, Titlow WB, Sengoku T, Breheny P, Matsunami H, McClintock TS. Modulation of the combinatorial code of odorant receptor response patterns in odorant mixtures. Mol Cell Neurosci. 2020;104:103469. https://doi.org/10.1016/j.mcn.2020.103469 PMID: 32061665
43. Grabe V, Sachse S. Fundamental principles of the olfactory code. Biosystems. 2018;164:94–101. https://doi.org/10.1016/j.biosystems.2017.10.010 PMID: 29054468
44. The Good Scents Company. The good scents company information system. 2023.
45. Lee BK, Mayhew EJ, Sanchez-Lengeling B, Wei JN, Qian WW, Little KA, et al. A principal odor map unifies diverse tasks in olfactory perception. Science. 2023;381(6661):999–1006. https://doi.org/10.1126/science.ade4401 PMID: 37651511
46. Achebouche R, Tromelin A, Audouze K, Taboureau O. Application of artificial intelligence to decode the relationships between smell, olfactory receptors and small molecules. Sci Rep. 2022;12(1):18817. https://doi.org/10.1038/s41598-022-23176-y PMID: 36335231
47. Qian WW, Wei JN, Sanchez-Lengeling B, Lee BK, Luo Y, Vlot M, et al. Metabolic activity organizes olfactory representations. eLife. 2023;12:e82502. https://doi.org/10.7554/eLife.82502 PMID: 37129358
48. Bozza T, Vassalli A, Fuss S, Zhang J-J, Weiland B, Pacifico R, et al. Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. Neuron. 2009;61(2):220–33. https://doi.org/10.1016/j.neuron.2008.11.010 PMID: 19186165
49. Belluscio L, Katz LC. Symmetry, stereotypy, and topography of odorant representations in mouse olfactory bulbs. J Neurosci. 2001;21(6):2113–22. https://doi.org/10.1523/JNEUROSCI.21-06-02113.2001 PMID: 11245695
50. Moriwaki H, Tian Y-S, Kawashita N, Takagi T. Mordred: a molecular descriptor calculator. J Cheminform. 2018;10(1):4. https://doi.org/10.1186/s13321-018-0258-y PMID: 29411163
51. Baker KL, Vasan G, Gumaste A, Pieribone VA, Verhagen JV. Spatiotemporal dynamics of odor responses in the lateral and dorsal olfactory bulb. PLoS Biol. 2019;17(9):e3000409. https://doi.org/10.1371/journal.pbio.3000409 PMID: 31532763
52. Burton SD, Wipfel M, Guo M, Eiting TP, Wachowiak M. A Novel Olfactometer for Efficient and Flexible Odorant Delivery. Chem Senses. 2019;44(3):173–88. https://doi.org/10.1093/chemse/bjz005 PMID: 30657873
53. Eiting TP, Wachowiak M. Artificial Inhalation Protocol in Adult Mice. Bio Protoc. 2018;8(18):e3024. https://doi.org/10.21769/BioProtoc.3024 PMID: 30417031
54. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in python. J Mach Learn Res. 2011;12(85):2825–30.
55. Leffingwell & Associates. Flavor-base, 9th edition. 2012.
56. Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. IEEE Trans Neural Netw. 2009;20(1):61–80. https://doi.org/10.1109/TNN.2008.2005605 PMID: 19068426
57. Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, et al. Graph neural networks: A review of methods and applications. AI Open. 2020;1:57–81. https://doi.org/10.1016/j.aiopen.2021.01.001

58. Ramsundar B, Eastman P, Walters P, Pande V. Deep learning for the life sciences. Sebastopol (CA): O'Reilly Media; 2019.
59. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. In: Kurková V, Manolopoulos Y, Hammer B, Iliadis LS, Maglogiannis I, editors. Artificial Neural Networks and Machine Learning—ICANN 2018, volume 11141 of Lecture Notes in Computer Science. Springer International Publishing; 2018. p. 270–9.
60. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell. 2020;2(1):56–67. https://doi.org/10.1038/s42256-019-0138-9 PMID: 32607472
61. Kendall MG. A New Measure of Rank Correlation. Biometrika. 1938;30(1/2):81–93. https://doi.org/10.2307/2332226
62. Geary RC. The Contiguity Ratio and Statistical Mapping. Inc Stat. 1954;5(3):115–46. https://doi.org/10.2307/2986645
63. Moran PAP. Notes on continuous stochastic phenomena. Biometrika. 1950;37(1–2):17–23. https://doi.org/10.1093/biomet/37.1-2.17 PMID: 15420245
64. Schmal C, Myung J, Herzel H, Bordyugov G. Moran's I quantifies spatio-temporal pattern formation in neural imaging data. Bioinformatics. 2017;33(19):3072–9. https://doi.org/10.1093/bioinformatics/btx351 PMID: 28575207
65. Recanatesi S, Bradde S, Balasubramanian V, Steinmetz NA, Shea-Brown E. A scale-dependent measure of system dimensionality. Patterns (N Y). 2022;3(8):100555. https://doi.org/10.1016/j.patter.2022.100555 PMID: 36033586
66. Björck Å, Golub GH. Numerical methods for computing angles between linear subspaces. Math Comp. 1973;27(123):579–94. https://doi.org/10.1090/s0025-5718-1973-0348991-3
67. Conway JH, Hardin RH, Sloane NJA. Packing Lines, Planes, etc.: Packings in Grassmannian Spaces. Exp Math. 1996;5(2):139–59. https://doi.org/10.1080/10586458.1996.10504585