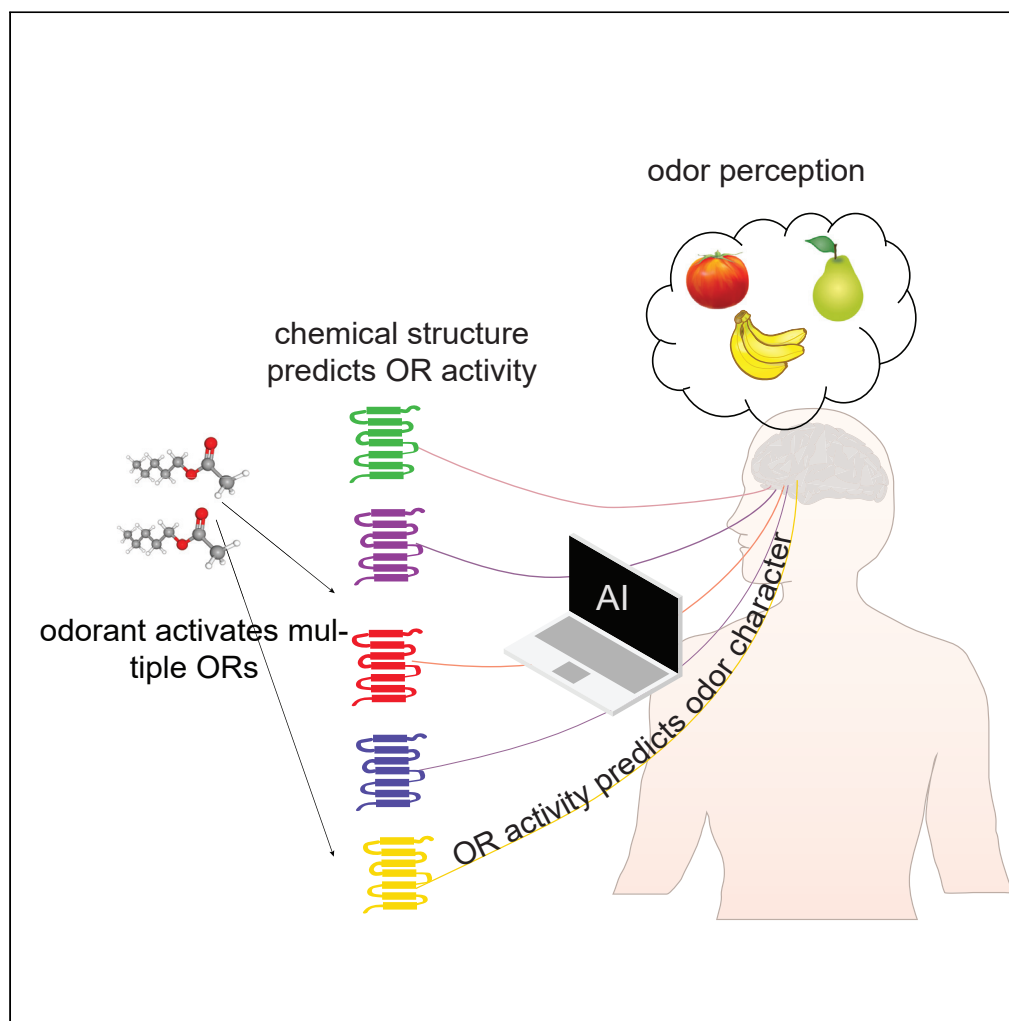


Article

Predicting Human Olfactory Perception from Activities of Odorant Receptors

Joel Kowalewski,
Anandasankar Ray

anand.ray@ucr.edu

HIGHLIGHTS

Machine learning predicted activity of 34 human ORs for ~0.5 million chemicals

Activities of human ORs could predict odor character using machine learning

Few OR activities were needed to optimize predictions of each odor percept

Behavior predictions in *Drosophila* also need few olfactory receptor activities

Kowalewski & Ray, iScience 23, 101361
August 21, 2020 © 2020 The Author(s).
<https://doi.org/10.1016/j.isci.2020.101361>

Article

Predicting Human Olfactory Perception from Activities of Odorant Receptors

Joel Kowalewski¹ and Anandasankar Ray^{1,2,3,*}

SUMMARY

Odor perception in humans is initiated by activation of odorant receptors (ORs) in the nose. However, the ORs linked to specific olfactory percepts are unknown, unlike in vision or taste where receptors are linked to perception of different colors and tastes. The large family of ORs (~400) and multiple receptors activated by an odorant present serious challenges. Here, we first use machine learning to screen ~0.5 million compounds for new ligands and identify enriched structural motifs for ligands of 34 human ORs. We next demonstrate that the activity of ORs successfully predicts many of the 146 different perceptual qualities of chemicals. Although chemical features have been used to model odor percepts, we show that biologically relevant OR activity is often superior. Interestingly, each odor percept could be predicted with very few ORs, implying they contribute more to each olfactory percept. A similar model is observed in *Drosophila* where comprehensive OR-neuron data are available.

INTRODUCTION

In humans a single odorant molecule might be described by different perceptual descriptions, influenced by culture, language, and experience (Majid and Kruspe, 2018). Such complexities suggest that, although olfactory circuitry is structurally similar across species, language or experience, which is dynamic and constantly evolving, could be a strong determinant of perceptual experience for humans. But even though the implication is that odor perception should be highly subjective, studies have shown that genetic variability in odorant receptors (ORs) contributes to odor perception. Equally, machine learning has accurately predicted perceptual descriptors of odorants from chemical features, suggesting that physicochemical properties influence perception (Debnath et al., 2019; Gutiérrez et al., 2018; Keller et al., 2017; Khan et al., 2007; Licon et al., 2019; Nozaki and Nakamoto, 2016; Sanchez-Lengeling et al., 2019). Moreover, modeling human odor perception using a large semantic similarity space has shown that accurate predictions of perceptual ratings are possible even when training and prediction are done on completely different study samples. That is, in aggregate human perceptual descriptors do not appear to be arbitrarily used and are generalizable (Gutiérrez et al., 2018).

The connection between odorant receptor activity and perception is not as well defined. It is unclear that the activity of specific ORs confers odor identity. For instance, although the human odorant receptor OR5AN1 is highly selective to musk-smelling chemicals, less selective ORs also respond to these chemicals (Ahmed et al., 2018). In simpler systems like insects, there is some evidence that activation or inhibition of certain odorant receptors is sufficient to drive behaviors from attraction and aversion to courtship, supporting the possibility of an underlying olfactory receptor code for perception (Chihara et al., 2014; Dweck et al., 2013; Kurtovic et al., 2007; MacWilliam et al., 2018; Stensmyr et al., 2012; Suh et al., 2004). Since these genetic studies are not feasible in humans, it is not yet clear how an olfactory receptor code can be generalizable, or whether it exists. It is, however, becoming increasingly plausible that there is indeed a perceptual code in humans. A few key odorant receptors have been reported for perceptual attributes other than musk (Shirasu et al., 2014) such as onion (Noe et al., 2017), general food-related volatiles (Geithe et al., 2017), and steroids (Keller et al., 2007). Sequence variation in the OR7D4 receptor has been shown to alter the perception of androstenone from a “sweaty,” unpleasant smell to one that is mildly “sweet” and pleasant (Keller et al., 2007). More recently, the specific amino acid residues of OR5AN1 that are responsible for its high selectivity to musk-smelling chemicals have also been confirmed (Ahmed et al., 2018). These studies were possible owing to three types of information: (1) perceptual responses of humans, (2) the odorant receptors that detect the chemicals from heterologous expression systems, and (3) genetic studies (Trimmer

¹Interdepartmental Neuroscience Program, University of California, Riverside, CA 92521, USA

²Department of Molecular, Cell and Systems Biology, University of California, 3401 Watkins Drive, Riverside, CA 92521, USA

³Lead Contact

*Correspondence: anand.ray@ucr.edu

<https://doi.org/10.1016/j.isci.2020.101361>



et al., 2019). Obtaining this information is not trivial for reasons that include the difficulty of receptor deorphanization and that behavioral responses are known for only a fraction of the purported volatile space, owing to low throughput data collection with human volunteers.

Although some of these limitations are not easily overcome, we reasoned that it would be of interest to leverage machine learning/artificial intelligence to better understand the ligands of odorant receptors and clarify the role of odorant receptor activity on human perceptual coding. Most prior machine learning efforts have focused on modeling odor perception according to the chemical features of odorants. Although these studies have shown promise and provide evidence for the physicochemical basis of odor perception, chemical features alone do not offer clear insight into biological coding, as this would require additional information about the olfactory receptors that odorants activate. Moreover, it is an extremely challenging task to isolate the olfactory receptors that are relevant to a percept.

Here, we tested if human odorant receptor responses from heterologous assays could be used in lieu of chemical features for modeling human odor perception and also developed models incorporating both approaches. We first created machine learning models to predict ligands for 34 human ORs. We could then use these models to evaluate how OR activity predicted perceptual descriptors. To start, we focused on hundreds of chemicals that human volunteers previously evaluated (Keller and Vosshall, 2016) and selected ORs that best predicted perceptual descriptors on a portion of training chemicals. Surprisingly, the prediction accuracy for models of only a few top scoring ORs compared favorably with large physicochemical feature models on 69 test chemicals (Keller et al., 2017), emphasizing that a small percentage of the OR pool is particularly useful for a given percept. This also suggested that specific subsets of ORs may be highly tuned to certain perceptual qualities, as implied in a prior network analysis of odorant receptors and perceptual descriptors (Bak et al., 2019).

RESULTS

Modeling OR Responses Using Chemical Features

Each odorant receptor is activated by a unique set of chemicals, and together the large olfactory receptor family can detect a vast chemical space. We compiled a database of 84 deorphanized human ORs and 54 allelic variants that have been tested with multiple odorants, altogether adding up to ~170 odorants (Adipietro et al., 2012; Braun et al., 2007; Charlier et al., 2012; Cook et al., 2009; Fujita et al., 2007; Gonzalez-Kristeller et al., 2015; Jacquier et al., 2006; Jaeger et al., 2013; Keller et al., 2007; Mainland et al., 2014; Mashukova et al., 2006; Maßberg et al., 2015; Matarazzo et al., 2005; McRae et al., 2012; Menashe et al., 2007; Neuhaus et al., 2006; Saito et al., 2009; Sanz et al., 2005; Schmiedeberg et al., 2007; Shirasu et al., 2014; Spehr et al., 2003; Topin et al., 2014). In order to generate more comprehensive odor response profiles of these ORs, we used machine learning to model structure-activity relationships. Among the 138 ORs, only 34 have a sufficient number of known ligands for machine learning models. For each of the 34 ORs, predictive chemical features were identified from the known ligands (Figure 1A). We validated the models by predicting ligands on a subset of odorants that were randomly left out of the training dataset, repeating this several times. The prediction success was high for the 34 models (average AUC = 0.88; shuffled chemical features average, AUC = 0.51, $p < 10^{-32}$) (Figures 1B, S1A, and S1B; Table S1).

The OR-ligand predictive models also gave us an opportunity to identify new ligands for the 34 ORs from a large chemical library (~450,000). In doing so, we developed a theoretical space that expands the existing data by a factor of 10 (Figure 1C). Enriched structural features were identifiable among the top predicted ligands for each OR, illustrating simple 2D features that are presumably important for activating each receptor (Figure 1D; Table S2).

Modeling Odorant Percepts from OR Responses

A key question in olfaction is how activities of ORs contribute to different perceptual qualities. Specific receptors contribute to androstenone perception (Keller et al., 2007); however, little is known about odorants commonly perceived as flavors and fragrances. One possibility is that their perception depends on a model similar to androstenone and one or few receptors contribute to perception. Alternatively, a model involving a combinatorial code of a large number of ORs is also possible, particularly since unlike androstenone, most odorants activate multiple ORs. In order to test these possibilities, we performed a series of analyses on a large dataset of human odor perception (Keller and Vosshall, 2016). Not only were a large number of chemicals tested by volunteers in this study, but computational studies have successfully demonstrated

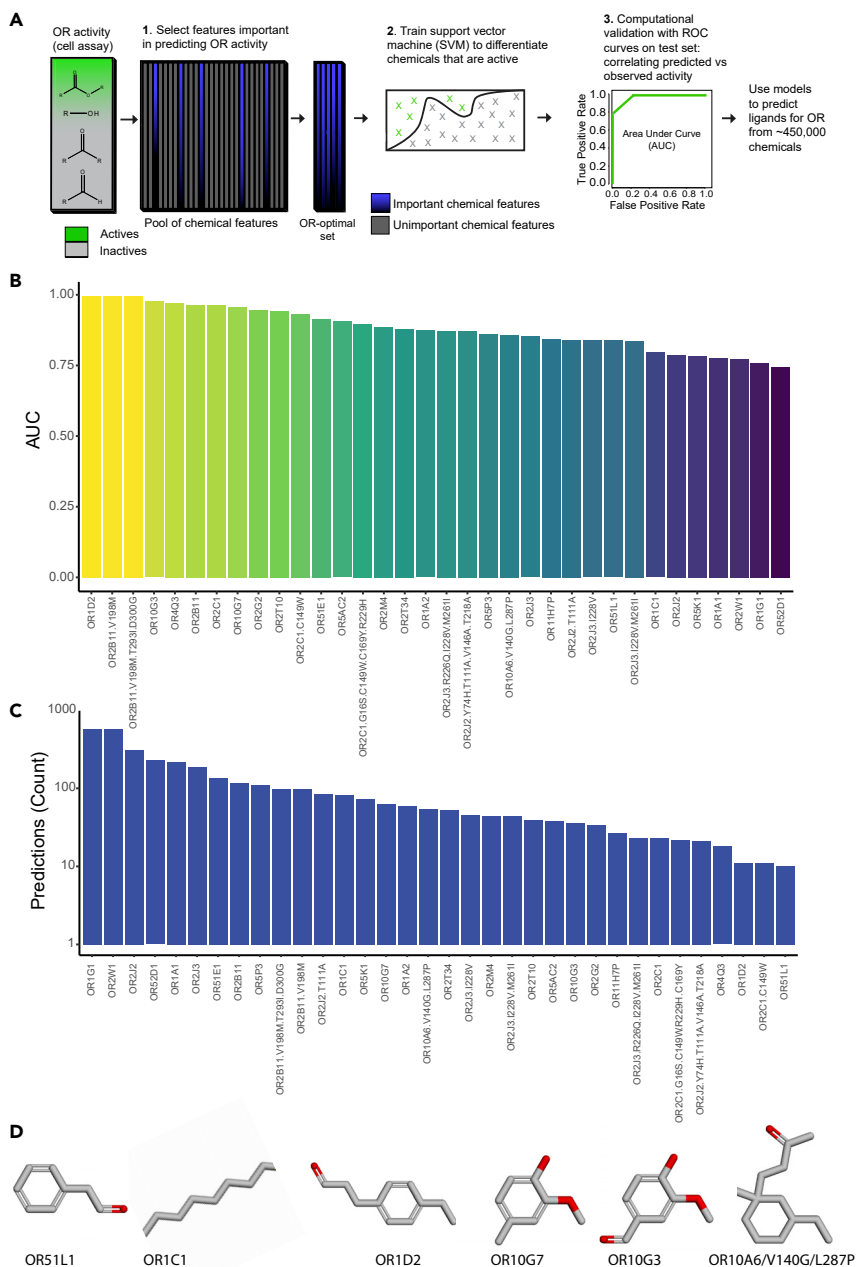


Figure 1. Features of Human Odorant Receptor Ligands Can Be Learned and New Ligands Predicted

(A) Pipeline for generating probability scores for chemicals with perceptual data. Starting with lists of ligands from heterologous assay data SVM models learn predictive physicochemical features for a subset of human ORs and OR variants with >2 ligands (34 total). These trained models in turn predict new chemicals such as those with known perceptual profiles.

(B) Average performance of 34 OR models using repeated 10-fold cross-validation.

(C) Number of ligands predicted for each of the 34 ORs in ~400,000 eMolecules library after filtering based on optimal probability score cutoffs and structural similarity to known ligands.

(D) Sample of enriched substructures among the top 10 predicted chemicals for indicated ORs. Only substructures that were non-trivial and present in at least half of the 10 highest scoring chemical ligands are shown. A comprehensive table of substructures for other receptors is provided in [Table S2](#).

structure-percept relationships (Gutiérrez et al., 2018; Keller et al., 2017; Kepple and Koulakov, 2017; Sanchez-Lengeling et al., 2019). However, several odorants used in the behavior study have not been tested for OR activities. We therefore used the OR-ligand models in the previous section to estimate activity for chemicals (Data S1), designating similar training and testing chemicals as described before (Keller et al., 2017) (407 training; 69 testing chemicals) (Figure S2A). Models containing only a few optimal ORs successfully predicted the perceptual descriptors for test chemicals (average test AUC = 0.78) (Figure S2B), particularly when compared with a similar approach based on different physicochemical feature encodings rather than ORs (Figure S2C). Lastly, because the activity on the 34 ORs was known for some chemicals in the (Keller and Vosshall, 2016) study, and it was unclear if this might affect the results, we revisited the analysis with these chemicals removed (326 train; 54 test chemicals). Test performance was not significantly reduced, compared with the earlier analysis ($p = 0.234$).

We next turned to another psychophysical study (ATLAS) that evaluated 146 perceptual descriptors for ~150 odorants. As before, most perceptual descriptors were well predicted from a small subset of ORs, despite the larger, more diverse descriptor pool in this study (Figure 2) (top 50 best performing: 10 ORs: avg. AUC = 0.84). When we compared the performance of the OR activity with the optimal chemical features, 47/146 perceptual descriptors were better predicted using the ORs. In light of this excellent performance, we further investigated ORs whose contributions to percept predictions are highest. Interestingly, only a few select ORs contributed strongly to the prediction of some perceptual descriptors (Figure 3).

In order to expand the scope and utilize activity information from the 104 ORs with few known ligands, we computed 3D similarity between chemicals in the ATLAS study and the OR ligands (Mahé et al., 2006) and identified the most likely active compound for each of the 104 ORs (materials and methods). When incorporating these additional ORs into the pipeline, predictions improved slightly for some perceptual descriptors. Among the top 50 best predicted descriptors, smaller OR models were significantly better than all 138 ORs on the test data (10 ORs AUC = 0.84; 138 OR AUC = 0.80, $t = 2.76$, $p = 0.007$), suggesting that the additional information was not often useful (Figure S3A). These 138 ORs still represent just a third of the human OR repertoire, and we anticipate our approach will help identify even better sets of ORs that are tuned to specific perceptual qualities as more human ORs get deorphanized.

Modeling Odorant Percepts from OR Responses and Chemical Features

Because many previous efforts have focused on predicting odor perception with chemical features (Keller et al., 2017), we tested if adding ORs could improve the predictions. We selected OR6P1, an OR ranked highly for “Cinnamon,” as a test case and added it to 34 optimal chemical features. Interestingly, we found a notable increase in predictive success on test chemicals (mean AUC chemical features: 0.77, mean AUC chemical features + OR6P1 = 0.81) (Figure S4A).

To determine if ORs could improve predictive models in an unbiased manner across the 146 perceptual descriptors, we combined the odor response information of the 138 ORs and the chemical features, selecting a small subset of important ORs and chemical features to create machine learning models (Figure 4A). We found that removing the top-ranked ORs and replacing them with those of lesser importance negatively impacted predictions for some descriptors (Figure 4B). If we permuted the activities of the optimal or top-ranked ORs for a given descriptor, the overall test performance significantly dropped ($p < 10^{-7}$), with 82% of descriptors better predicted with non-permuted ORs (Table S3). Collectively, these results indicate that specific ORs appear to contribute more than others and perceptual predictions are generally improved by including ORs (Table S4).

In order to visualize relationships among the perceptual descriptors based on predictive ORs and chemical features, we next performed a cluster analysis. When examining the clustering based only on perceptual ratings of chemicals (Figure 5A), we found the top five predictive ORs grouped the perceptual descriptors similarly (Figure 5B). Notably, randomly selecting five ORs failed to produce any meaningful groups or clusters of perceptual descriptors (Figure 5C). Combining the most predictive ORs and chemical features improved the clustering of perceptual descriptors (Figure 5D). Overall, the descriptors that were best clustered in Figure 5A (silhouette width > 0.3) matched completely or partially with Figures 5B and 5D, with the exception of “Fishy” and “Kippery.” This indicates that relationships among perceptual descriptors in the ATLAS training set are somewhat preserved in OR activity or chemical feature models, even when only a small amount of chemical or information is included in each model.

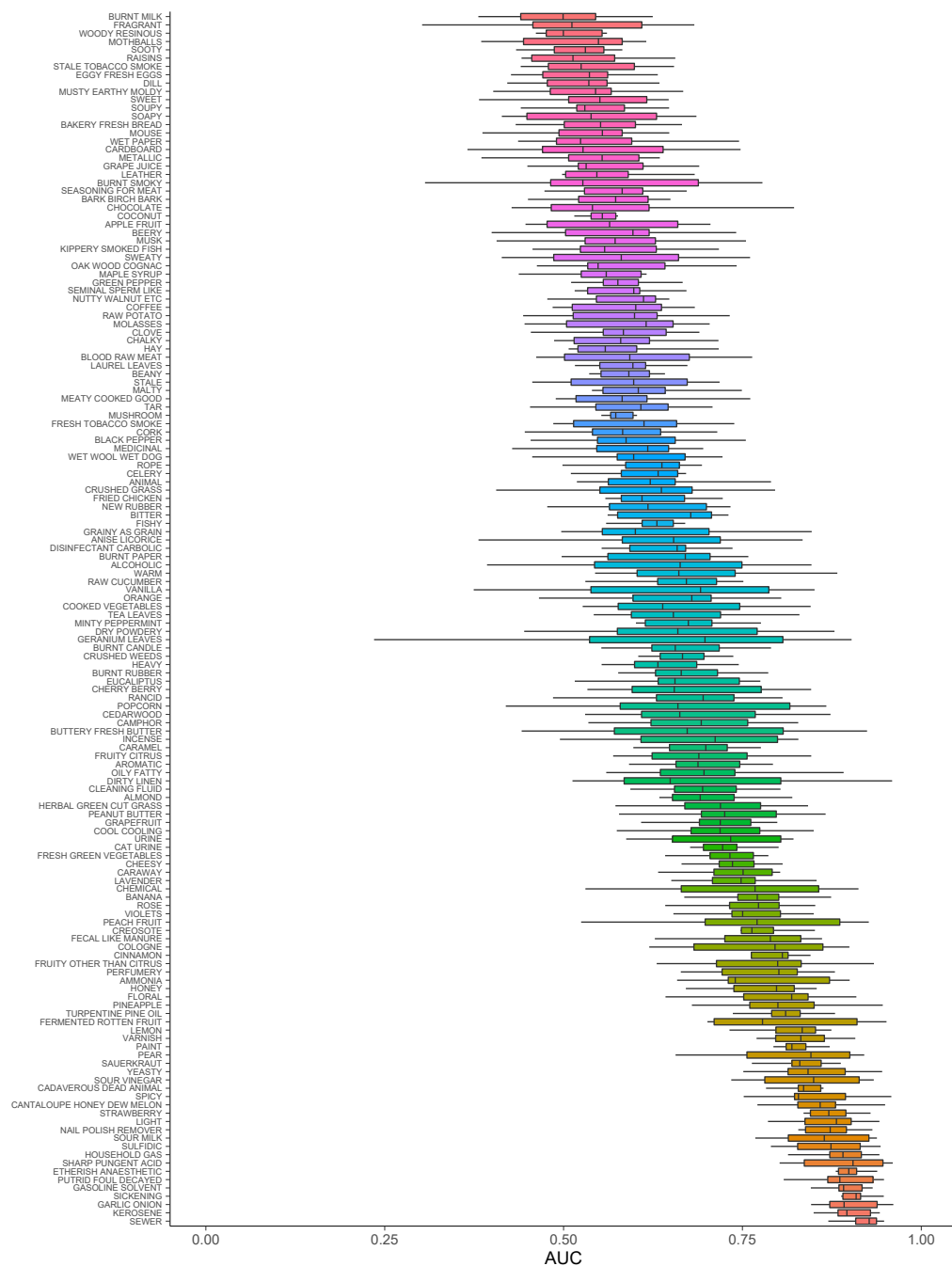


Figure 2. OR Activity Can Model Diverse Olfactory Percepts in Human Studies

Performance of RBF SVM models trained with 10 ORs for ATLAS study data. The top 10% usage chemicals are predicted for all 146 perceptual descriptors in the study. Successful classification of these chemicals is reported as the mean Area Under the Curve (AUC) over repeated 10-fold cross-validation (10-fold repeated 5 times; 50 folds total). To limit biased validation, the procedure was run twice, setting aside different test chemicals, determining important OR subsets to predict the descriptors with these chemicals excluded, then ensuring that the cross-validated AUC comprised 60% completely hidden chemicals. The variability in the plot is the standard deviation over these two distinct runs. High variability may arise as the top 10% usage is computed from the training data. SVM, Support Vector Machine; RBF, Radial Basis Function; additional algorithm details in methods.

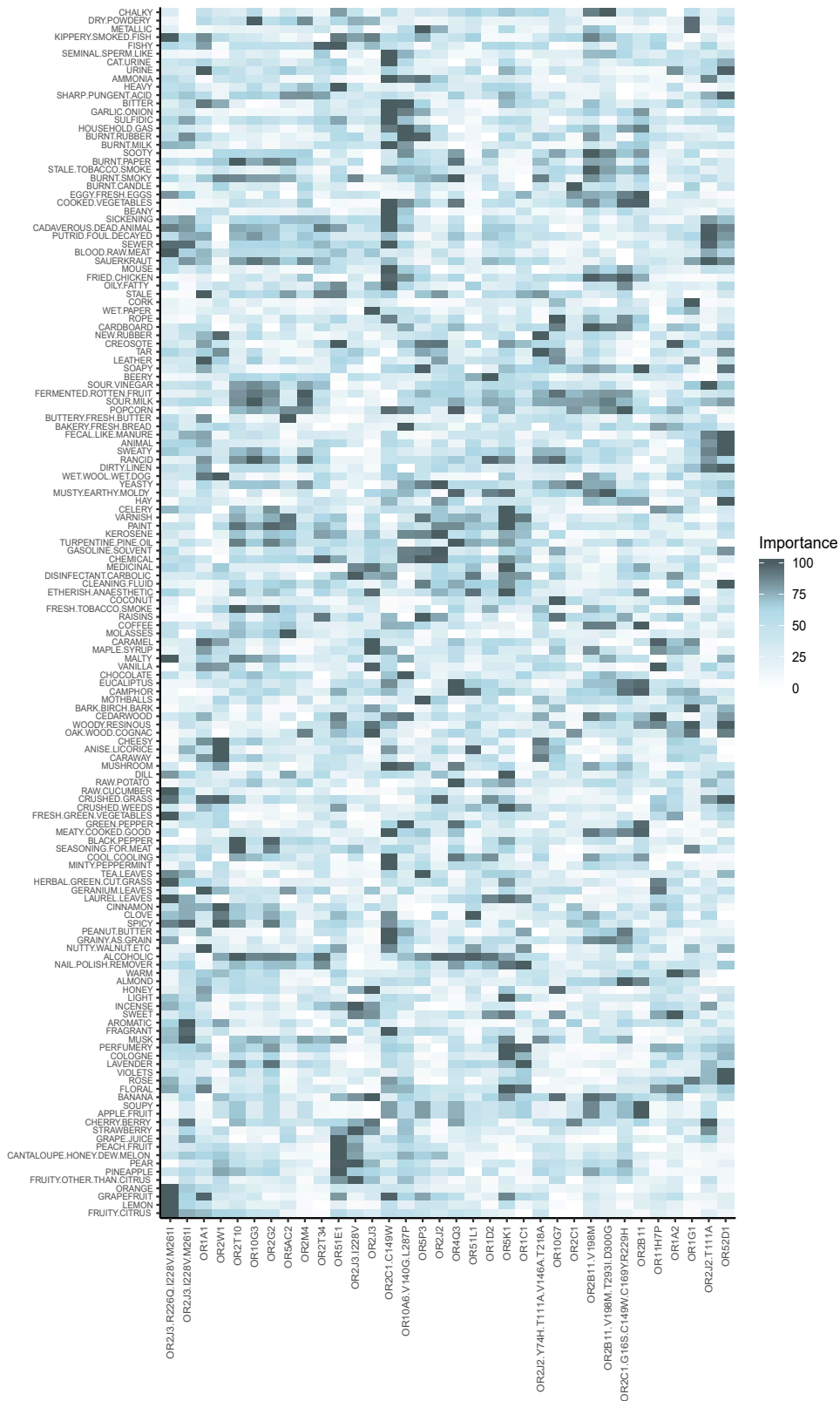


Figure 3. Contribution of ORs to Perceptual Models

Importance of individual ORs for machine learning models of each of the 146 ATLAS perceptual descriptors. The heatmap is generated by fitting models for each OR separately and scaling relative to maximum AUC (100). Importance is shown with the most important ORs in blue. Labels for the perceptual descriptors (yaxis) and ORs (xaxis) are arranged relative to similar importance values.

Modeling with *In Vivo* OR Response Data from *Drosophila*

One of the interesting observations we have is that only a few ORs are picked and are sufficient to create predictive models of odor perception. However, the perceptual descriptor—to—OR mapping we analyzed here represents data from only ~20% of the human OR repertoire and one possibility is that when more ORs are available to pick from, a larger number will be selected computationally as optimal. In order to understand the contribution of specific olfactory receptors to behavior in a system where a large fraction of odorant receptors have been deorphanized, we turned to the *Drosophila melanogaster* model system. *In vivo* odor-response spectra are known for several odorants for the majority of odorant receptors (Ors) and olfactory receptor neurons (ORNs) in the adults, as well as the behavioral valence (attraction versus aversion) to these odorants (MacWilliam et al., 2018; Hallem and Carlson, 2006).

We adapted our approach to predict behavioral valence of flies (Figure 6A), and we could do so with significant success using a small number of important chemical features and electrophysiologically measured responses from sensory neurons. Similar to what we observed with human ORs, a subset of the *in vivo* *Drosophila* Or activities was favored for odor valence predictions, beyond collections of numerous chemical features (Figure 6B). Evaluating the best valence predictors for test chemicals from a combined set of Or/ORN activities and chemical features indicated that the Or/ORNs significantly contributed to odor valence predictions, consistent with the *in vitro* human data ($R^2 = 0.66$; Shuffle ORs + Chemical Features: 0.51 , $p = 0.007$) (Figure 6C). These results also suggested that a small number of *Drosophila* Or/ORN activities is highly predictive on the same set of test chemicals. Interestingly, additional Ors/ORNs failed to improve predictions (Or/ORN subset: $R^2 = 0.53$; all other ORs: $R^2 = 0.40$, $p = 0.015$) (Table S5). Although this type of analysis remains to be done in humans, the results from flies suggest that even when a more comprehensive receptor or neuron array is added, only a small subset of the available receptors appears information rich as far as behavioral predictions are concerned (Figure 6D).

DISCUSSION

Although previous machine learning pipelines have found some success using chemical features, selecting the optimal feature sets for predictions of perception is not well defined. We found that human odorant responses from heterologous assays could be used with comparable and sometimes better predictive success. In part, the result is anticipated by the fact that each OR is presumably selective to very specific physicochemical features themselves. Both the human perceptual descriptor and fly valence predictions suggest that a substantive portion of odor identity arises early in the processing stream, at the olfactory receptors, based on high predictive success rates (~76%–91%). It is likely that the remaining portion depends on experience-dependent modulation, supporting a downstream model with reliance on distributed neuronal networks for human perceptual coding. Our findings support a “primacy model” that holds that a small number of distinct and overlapping olfactory receptor activity profiles encode odor identity (Wilson et al., 2017). Although increasing concentration activates more receptors, the highest sensitivity receptors start responding first as an animal approaches an odor source and presumably continue to convey the identity. Such a model is consistent with the findings reported here and by others (Weiss et al., 2012) because it appears that only a few ORs contribute to a perceptual descriptor and it is therefore also tractable to predict how a chemical smells from specific physicochemical properties.

Nevertheless, it is unclear how information arising early in the olfactory pathway is preserved along the complex circuits and can in fact lead to generalizable perceptual features. The spatial organization of the olfactory receptor neurons and glomeruli are for one not well preserved in the piriform cortex. Unlike the retinotopic and tonotopic patterning observed in the visual and auditory cortices, representing spatio-temporal properties of visual and auditory stimuli as they are processed at sensory neurons, piriform activity appears randomly distributed, without a clear mapping of physicochemical features (Stettler and Axel, 2009). A combination of computational models and calcium imaging has, however, shown piriform circuits, although they are qualitatively different, can support perceptual invariance amid changes in concentration and across different odorants (Roland et al., 2017; Schaffer et al., 2018). Similarly, neural tracing

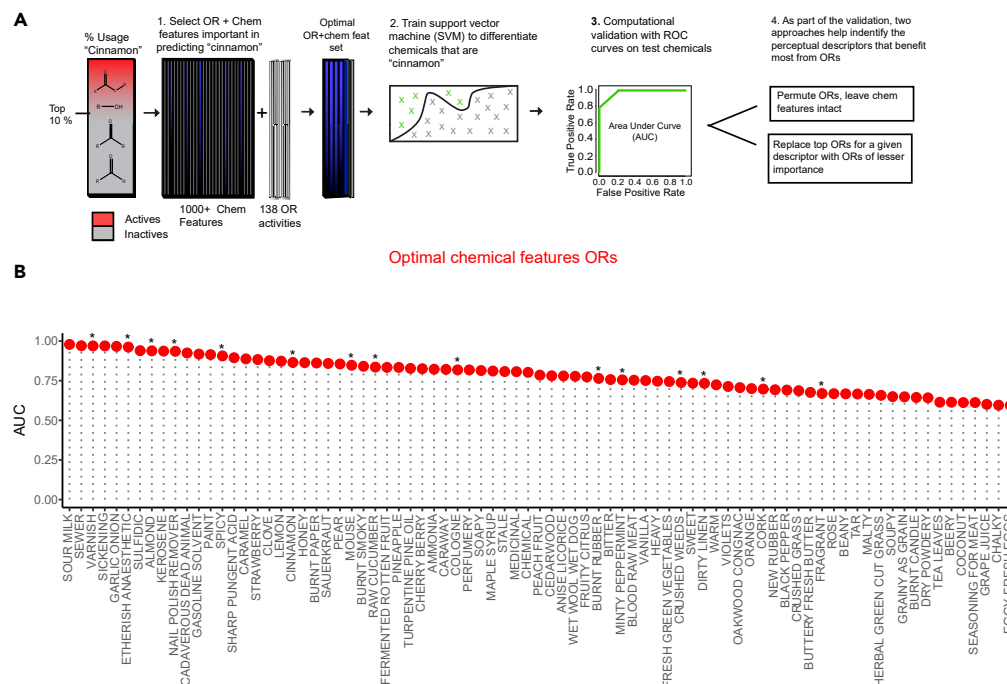


Figure 4. Few Odorant Receptors Are Needed to Predict Perceptual Descriptors

(A) Schematic of the approach to selecting a small number of important chemical features and ORs, followed by model-fitting. Two methods, including replacing top-ranking ORs with those of lesser importance and permuting (shuffling) the OR activities, help identify perceptual descriptors where ORs contribute relative to chemical features. To standardize the analysis, the training and validation are as outlined in Figure 2.

(B) Combined chemical feature-OR models predict the top 10% usage of ATLAS perceptual descriptors. The (*) symbol signifies a notable decrease in performance occurred if the ORs were replaced with ones of lesser importance (one-tailed independent samples t test, $p \leq 0.05$). For the comparison with permuted or shuffled OR activities, other metrics, and benchmarking relative to chemical features, see Tables S3 and S4.

experiments in mice support that, although olfactory circuitry differs from other sensory modalities, odor-related information is represented along equally structured neuroanatomical pathways, as in the piriform output projecting to the orbitofrontal cortex (Chen et al., 2014).

One possibility is that only one or few receptors of the many that detect an odorant actually convey percept. The evolutionary landscape should accordingly be coupled to biologically relevant or frequently encountered features of the chemical space, as has been implied by characterizations of receptors highly tuned for musk and onion-related compounds (Ahmed et al., 2018; Noe et al., 2017), in addition to the highly conserved trace amine-associated receptors (TAARs) and their importance in modulating behavioral output in mice (Dewan et al., 2018). In our analyses, the OR specialized for musk was not a top candidate for musk predictions but contributed strongly to predictions of "sweaty." Since methods for selecting and ranking ORs depend on characteristics of the available data, interpretations should be cautious, acknowledging that the human OR data are sparse and the participants and chemical sets from the ATLAS and Keller studies are not exhaustive. Yet from these same considerations the positive results achieved are unexpected, especially when compared with predictions of odor perception using chemical features.

Odorant receptors (ORs) are also expressed in non-olfactory tissues. Ligands for certain ORs have been shown to modify the function and proliferation of multiple cell types. Although the precise mechanisms are not well defined, ORs represent promising therapeutic targets. Ligands for ORs such as OR51E1, OR10G7, and OR1D2, which were included in this study, are candidate treatments for conditions ranging from prostate cancer and chronic obstructive pulmonary disease (COPD) to atopic dermatitis (Kalbe et al., 2016; Maßberg et al., 2016; Tham et al., 2019). We therefore anticipate that the predictions and the analysis of known and candidate OR ligands from this study will also have value in non-olfactory studies.

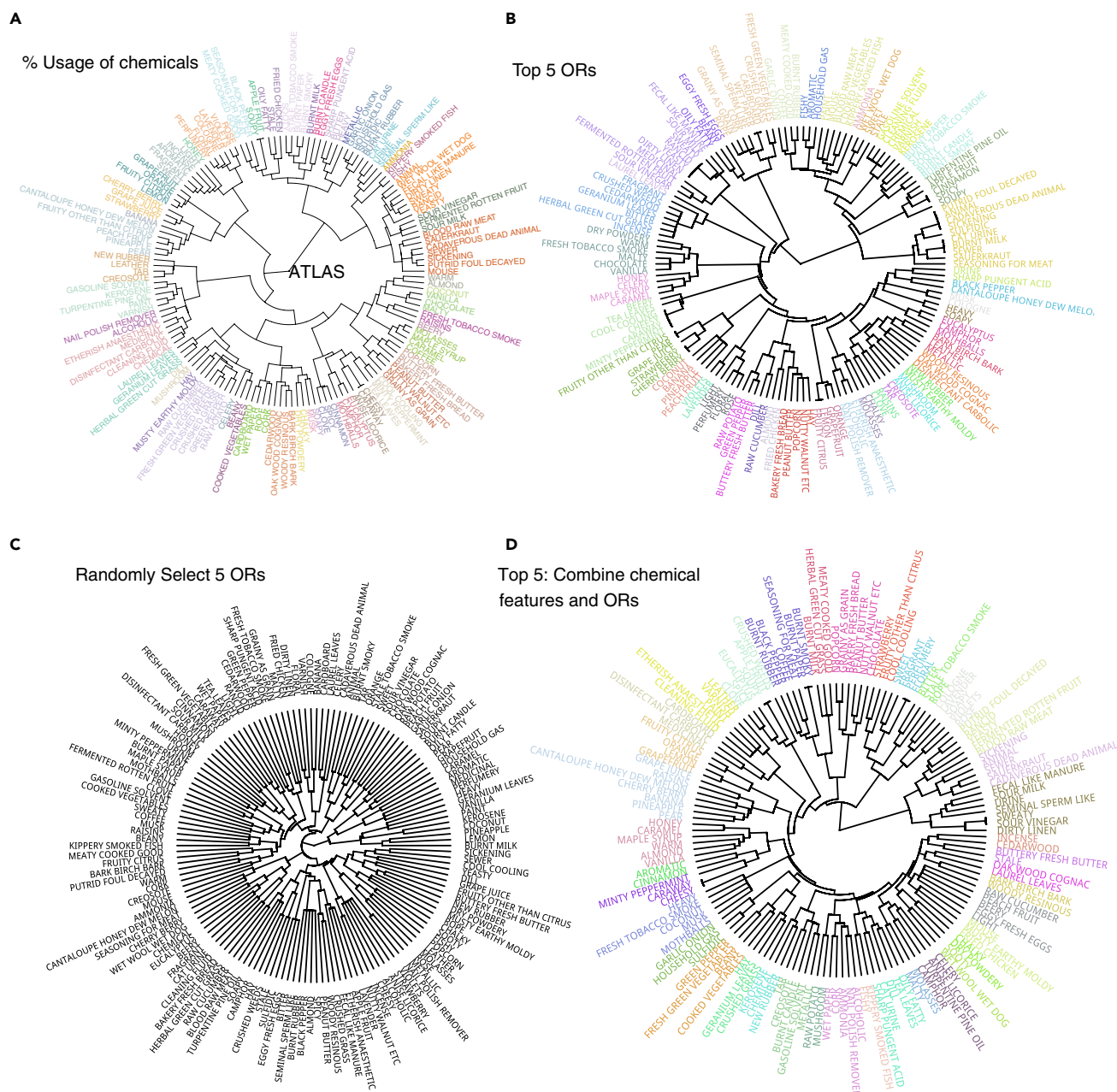


Figure 5. A Few Key ORs or Chemical Features Sensibly Cluster the Perceptual Descriptors

(A) Dendrogram representation of the Euclidean distances among perceptual descriptors based on overlap of perceptual response data (% Usage) from chemicals in the ATLAS study.

(B) Dendrogram from the top five ORs picked per perceptual descriptor.

(C) Dendrogram created from five randomly chosen ORs per perceptual descriptor.

(D) Dendrogram from the five best overall predictors including OR and chemical features per perceptual descriptor. Clustering is hierarchical and based on Euclidean distance (A) or the Jaccard distance (B–D). Cluster number (colored branches) inferred from gap statistic across bootstrap samples.

Limitations of the Study

The computational approach presented in the study is restricted by training sets from previously deorphanized human odorant receptors (OR) determined by *in vitro* assays. Only a small fraction of the human OR family has been deorphanized *in vitro*, thereby limiting the identification of the optimally predictive ORs in this study. Moreover, the number of chemicals with well-defined perceptual profiles determined behaviorally is small relative to the space of chemicals that are likely to have odorant properties. Since the

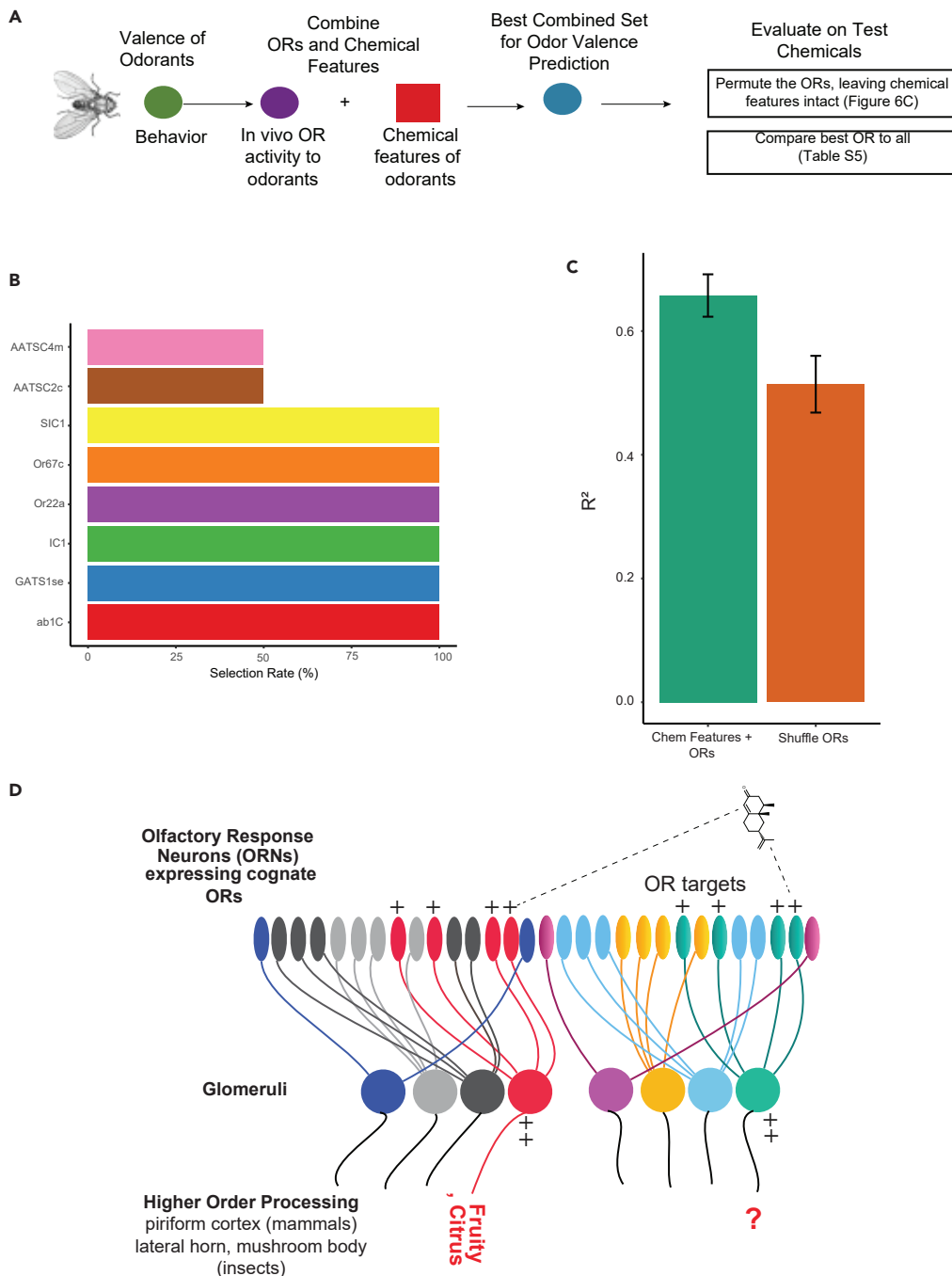


Figure 6. Few Odorant Receptor Activities in *Drosophila* Are Highly Predictive of Valence

(A) Schematic for applying machine learning to identify optimal predictors of odor valence in *Drosophila* from *in vivo* neural activity and chemical features. The best combined model is evaluated on test chemicals. OR contributions to *Drosophila* odor valence are assessed by shuffling the OR activities in the combined model as well as comparing the best OR versus all (Table S5).

(B) Selecting chemical features and *in vivo* OR activities that optimally predict odor valence. Recursive feature elimination (RFE) is run twice to accomplish this. Selection in the top 10 over these runs is plotted as a percent. Additional details on selecting optimal models in methods.

(C) The best combined model is evaluated on test chemicals, with and without the OR activities shuffled. Error bars are the SEM.

(D) Generic model displaying a many-to-one mapping between ORNs and glomeruli. Although there are >1 responding units (ORs), information that confers perceptual character is restricted to a smaller subset of the input.

computational approach we outlined depends on the size and complexity of OR and perceptual datasets, our results should be interpreted alongside these limitations.

Resource Availability

Lead Contact

Anandasankar Ray, anand.ray@ucr.edu

Materials Availability

No new materials were generated in this study.

Data and Code Availability

Data used in the analyses is publicly available from the references cited, and the data generated in this manuscript are supplied as supplemental files. Any additional data associated with this manuscript is also available in other formats on request from the corresponding author and lead contact. Upon request, the lead contact will make available any previously unreported custom computer code or algorithm used to generate results that are reported in the paper.

METHODS

All methods can be found in the accompanying [Transparent Methods](#) supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101361>.

ACKNOWLEDGMENTS

Partially funded by University of California Riverside internal HATCH funding and partially funded by NSF grant 1664172.

AUTHOR CONTRIBUTIONS

J.K. and A.R. conceived the study. J.K. planned and executed the analysis. J.K. and A.R. wrote the manuscript.

DECLARATION OF INTERESTS

J.K. and A.R. are listed as inventors on patent applications filed by the university. A.R. is the founder of a startup company called Sensorygen that is involved in finding new flavors and insect repellents.

Received: January 9, 2020

Revised: May 31, 2020

Accepted: July 9, 2020

Published: August 21, 2020

REFERENCES

- Adipietro, K.A., Mainland, J.D., and Matsunami, H. (2012). Functional evolution of mammalian odorant receptors. *PLoS Genet.* *8*, e1002821.
- Ahmed, L., Zhang, Y., Block, E., Buehl, M., Corr, M.J., Cormanich, R.A., Gundala, S., Matsunami, H., O'Hagan, D., Ozbil, M., et al. (2018). Molecular mechanism of activation of human musk receptors OR5AN1 and OR1A1 by (R)-muscone and diverse other musk-smelling compounds. *Proc. Natl. Acad. Sci. U S A* *115*, E3950–E3958.
- Bak, J.H., Jang, S.J., and Hyeon, C. (2019). Modular structure of human olfactory receptor codes reflects the bases of odor perception. *BioRxiv*. <https://doi.org/10.1101/525287>.
- Braun, T., Volland, P., Kunz, L., Prinz, C., and Gratzl, M. (2007). Enterochromaffin cells of the human gut: sensors for spices and odorants. *Gastroenterology* *132*, 1890–1901.
- Charlier, L., Topin, J., Ronin, C., Kim, S.K., Goddard, W.A., Efremov, R., and Golebiowski, J. (2012). How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case. *Cell. Mol. Life Sci.* *69*, 4205–4213.
- Chen, C.-F.F., Zou, D.-J., Altomare, C.G., Xu, L., Greer, C.A., and Firestein, S.J. (2014). Nonsensory target-dependent organization of piriform cortex. *Proc. Natl. Acad. Sci. U S A* *111*, 16931–16936.
- Chihara, T., Kitabayashi, A., Morimoto, M., Takeuchi, K.ichi, Masuyama, K., Tonoki, A., Davis, R.L., Wang, J.W., and Miura, M. (2014). Caspase inhibition in select olfactory neurons restores innate attraction behavior in aged *Drosophila*. *PLoS Genet.* *10*, e1004437.
- Cook, B.L., Steuerwald, D., Kaiser, L., Graveland-Bikker, J., Vanberghem, M., Berke, A.P., Herlihy, K., Pick, H., Vogel, H., and Zhang, S. (2009). Large-scale production and study of a synthetic G protein-coupled receptor: human olfactory receptor 17-4. *Proc. Natl. Acad. Sci. U S A* *106*, 11925–11930.
- Debnath, T., Prasetyawan, D., and Nakamoto, T. (2019). Prediction of Odor Descriptor Group of

Essential Oils from Mass Spectra Using Machine Learning (Institute of Electrical and Electronics Engineers (IEEE)), pp. 1–3.

Dewan, A., Cichy, A., Zhang, J., Miguel, K., Feinstein, P., Rinberg, D., and Bozza, T. (2018). Single olfactory receptors set odor detection thresholds. *Nat. Commun.* 9, 2887.

Dweck, H.K.M., Ebrahim, S.A.M., Kromann, S., Bown, D., Hillbur, Y., Sachse, S., Hansson, B.S., and Stensmyr, M.C. (2013). Olfactory preference for egg laying on citrus substrates in *Drosophila*. *Curr. Biol.* 23, 2472–2480.

Fujita, Y., Takahashi, T., Suzuki, A., Kawashima, K., Nara, F., and Koishi, R. (2007). Deorphanization of dresden G protein-coupled receptor for an odorant receptor. *J. Recept. Signal Transduct.* 27, 323–334.

Geithe, C., Noe, F., Kreissl, J., and Krautwurst, D. (2017). The broadly tuned odorant receptor OR1A1 is highly selective for 3-methyl-2,4-nonanedione, a key food odorant in aged wines, tea, and other foods. *Chem. Senses.* 42, 181–193.

Gonzalez-Kristeller, D.C., do Nascimento, J.B.P., Galante, P.A.F., and Malnic, B. (2015). Identification of agonists for a group of human odorant receptors. *Front. Pharmacol.* 6, 35, eCollection.

Gutiérrez, E.D., Dhurandhar, A., Keller, A., Meyer, P., and Cecchi, G.A. (2018). Predicting natural language descriptions of mono-molecular odorants. *Nat. Commun.* 9, 4979.

Hallem, E.A., and Carlson, J.R. (2006). Coding of odors by a receptor repertoire. *Cell* 125, 143–160.

Jacquier, V., Pick, H., and Vogel, H. (2006). Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds. *J. Neurochem.* 97, 537–544.

Jaeger, S.R., McRae, J.F., Bava, C.M., Beresford, M.K., Hunter, D., Jia, Y., Cheang, S.L., Jin, D., Peng, M., Gamble, J.C., et al. (2013). A mendelian trait for olfactory sensitivity affects odor experience and food selection. *Curr. Biol.* 23, 1601–1605.

Kalbe, B., Knobloch, J., Schulz, V.M., Wecker, C., Schlimm, M., Scholz, P., Jansen, F., Stoelben, E., Philippou, S., Hecker, E., et al. (2016). Olfactory receptors modulate physiological processes in human airway smooth muscle cells. *Front. Physiol.* 7, 339, eCollection.

Keller, A., and Vosshall, L.B. (2016). Olfactory perception of chemically diverse molecules. *BMC Neurosci.* 17, 55.

Keller, A., Zhuang, H., Chi, Q., Vosshall, L.B., and Matsunami, H. (2007). Genetic variation in a human odorant receptor alters odour perception. *Nature* 449, 468–472.

Keller, A., Gerkin, R.C., Guan, Y., Dhurandhar, A., Turu, G., Szalai, B., Mainland, J.D., Ihara, Y., Yu, C.W., Wolfinger, R., et al. (2017). Predicting human olfactory perception from chemical features of odor molecules. *Science* 355, 820–826.

Kepple, D., and Koulakov, A. (2017). Constructing an olfactory perceptual space and predicting percepts from molecular structure. *arXiv*, arXiv:1708.05774.

Khan, R.M., Luk, C.-H., Flinker, A., Aggarwal, A., Lapid, H., Haddad, R., and Sobel, N. (2007). Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. *J. Neurosci.* 27, 10015–10023.

Kurtovic, A., Widmer, A., and Dickson, B.J. (2007). A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nature* 446, 542–546.

Licon, C.C., Bosc, G., Sabri, M., Mantel, M., Fournel, A., Bushdid, C., Golebiowski, J., Robardet, C., Plantevit, M., Kaytoue, M., et al. (2019). Chemical features mining provides new descriptive structure-odor relationships. *PLoS Comput. Biol.* 15, e1006945.

MacWilliam, D., Kowalewski, J., Kumar, A., Pontrello, C., and Ray, A. (2018). Signaling mode of the broad-spectrum conserved CO₂ receptor is one of the important determinants of odor valence in *Drosophila*. *Neuron* 97, 1153–1167.e4.

Mahé, P., Ralaivola, L., Stoven, V., and Vert, J.P. (2006). The pharmacophore kernel for virtual screening with support vector machines. *J. Chem. Inf. Model.* 46, 2003–2014.

Mainland, J.D., Keller, A., Li, Y.R., Zhou, T., Trimmer, C., Snyder, L.L., Moberly, A.H., Adipietro, K.A., Liu, W.L.L., Zhuang, H., et al. (2014). The missense of smell: Functional variability in the human odorant receptor repertoire. *Nat. Neurosci.* 17, 114–120.

Majid, A., and Kruspe, N. (2018). Hunter-gatherer olfaction is special. *Curr. Biol.* 28, 409–413.e2.

Mashukova, A., Spehr, M., Hatt, H., and Neuhaus, E.M. (2006). Beta-arrestin2-mediated internalization of mammalian odorant receptors. *J. Neurosci.* 26, 9902–9912.

Maßberg, D., Simon, A., Häussinger, D., Keitel, V., Gisselmann, G., Conrad, H., and Hatt, H. (2015). Monoterpene (-)-citronellal affects hepatocarcinoma cell signaling via an olfactory receptor. *Arch. Biochem. Biophys.* 566, 100–109.

Maßberg, D., Jovancevic, N., Offermann, A., Simon, A., Baniahmad, A., Perner, S., Pungsrinont, T., Luko, K., Philippou, S., Ubrigg, B., et al. (2016). The activation of OR51E1 causes growth suppression of human prostate cancer cells. *Oncotarget* 7, 48231–48249.

Matarazzo, V., Clot-Faybessé, O., Marcet, B., Guiraudie-Capraz, G., Atanasova, B., Devauchelle, G., Cerutti, M., Etiévant, P., and Ronin, C. (2005). Functional characterization of two human olfactory receptors expressed in the baculovirus Sf9 insect cell system. *Chem. Senses.* 30, 195–207.

McRae, J.F., Mainland, J.D., Jaeger, S.R., Adipietro, K.A., Matsunami, H., and Newcomb, R.D. (2012). Genetic variation in the odorant receptor OR2J3 is associated with the ability to

detect the “grassy” smelling odor, cis-3-hexen-1-ol. *Chem. Senses.* 37, 585–593.

Menashe, I., Abaffy, T., Hasin, Y., Goshen, S., Yahalom, V., Luetje, C.W., and Lancet, D. (2007). Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biol.* 5, e284.

Neuhaus, E.M., Mashukova, A., Zhang, W., Barbour, J., and Hatt, H. (2006). A specific heat shock protein enhances the expression of mammalian olfactory receptor proteins. *Chem. Senses.* 31, 445–452.

Noe, F., Polster, J., Geithe, C., Kotthoff, M., Schieberle, P., and Krautwurst, D. (2017). OR2M3: a highly specific and narrowly tuned human odorant receptor for the sensitive detection of onion key food odorant 3-mercapto-2-methylpentan-1-ol. *Chem. Senses* 42, 195–210.

Nozaki, Y., and Nakamoto, T. (2016). Odor impression prediction from mass spectra. *PLoS One* 11, e0157030.

Roland, B., Deneux, T., Franks, K.M., Bathellier, B., and Fleischmann, A. (2017). Odor identity coding by distributed ensembles of neurons in the mouse olfactory cortex. *Elife* 6, e26337.

Saito, H., Chi, Q., Zhuang, H., Matsunami, H., and Mainland, J.D. (2009). Odor coding by a mammalian receptor repertoire. *Sci. Signal.* 2, ra9.

Sanchez-Lengeling, B., Wei, J.N., Lee, B.K., Gerkin, R.C., Aspuru-Guzik, A., and Wiltschko, A.B. (2019). Machine learning for scent: learning generalizable perceptual representations of small molecules. *arXiv*, arXiv:1910.10685.

Sanz, G., Schlegel, C., Pernollet, J.C., and Briand, L. (2005). Comparison of odorant specificity of two human olfactory receptors from different phylogenetic classes and evidence for antagonism. *Chem. Senses.* 30, 69–80.

Schaffer, E.S., Stettler, D.D., Kato, D., Choi, G.B., Axel, R., and Abbott, L.F. (2018). Odor perception on the two sides of the brain: consistency despite randomness. *Neuron* 98, 736–742.e3.

Schmiedeberg, K., Shirokova, E., Weber, H.P., Schilling, B., Meyerhof, W., and Krautwurst, D. (2007). Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2. *J. Struct. Biol.* 159, 400–412.

Shirasu, M., Yoshikawa, K., Takai, Y., Nakashima, A., Takeuchi, H., Sakano, H., and Touhara, K. (2014). Olfactory receptor and neural pathway responsible for highly selective sensing of musk odors. *Neuron* 81, 165–178.

Spehr, M., Gisselmann, G., Poplawski, A., Riffell, J.A., Wetzal, C.H., Zimmer, R.K., and Hatt, H. (2003). Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* 299, 2054–2058.

Stensmyr, M.C., Dweck, H.K.M., Farhan, A., Ibba, I., Strutz, A., Mukunda, L., Linz, J., Grabe, V., Steck, K., Lavista-Llanos, S., et al. (2012). A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. *Cell* 151, 1345–1357.

Stettler, D.D., and Axel, R. (2009). Representations of odor in the piriform cortex. *Neuron* 63, 854–864.

Suh, G.S.B., Wong, A.M., Hergarden, A.C., Wang, J.W., Simon, A.F., Benzer, S., Axel, R., and Anderson, D.J. (2004). A single population of olfactory sensory neurons mediates an innate avoidance behaviour in *Drosophila*. *Nature* 431, 854–859.

Tham, E.H., Dyjack, N., Kim, B.E., Rios, C., Seibold, M.A., Leung, D.Y.M., and Goleva, E. (2019). Expression and function of the ectopic

olfactory receptor OR10G7 in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* 143, 1838–1848.e4.

Topin, J., Demarch, C.A., Charlier, L., Ronin, C., Antonczak, S., and Golebiowski, J. (2014). Discrimination between olfactory receptor agonists and non-agonists. *Chem* 20, 10227–10230.

Trimmer, C., Keller, A., Murphy, N.R., Snyder, L.L., Willer, J.R., Nagai, M.H., Katsanis, N., Vosshall, L.B., Matsunami, H., and Mainland, J.D. (2019). Genetic variation across the human olfactory

receptor repertoire alters odor perception. *Proc. Natl. Acad. Sci. U S A* 116, 9475–9480.

Weiss, T., Snitz, K., Yablonka, A., Khan, R.M., Gafso, D., Schneidman, E., and Sobel, N. (2012). Perceptual convergence of multi-component mixtures in olfaction implies an olfactory white. *Proc. Natl. Acad. Sci. U S A* 109, 19959–19964.

Wilson, C.D., Serrano, G.O., Koulakov, A.A., and Rinberg, D. (2017). A primacy code for odor identity. *Nat. Commun.* 8, 1477.

iScience, Volume 23

Supplemental Information

**Predicting Human Olfactory Perception
from Activities of Odorant Receptors**

Joel Kowalewski and Anandasankar Ray

iScience, Volume ■ ■

Supplemental Information

Predicting Human Olfactory Perception from Activities of Odorant Receptors

Joel Kowalewski and Anandasankar Ray

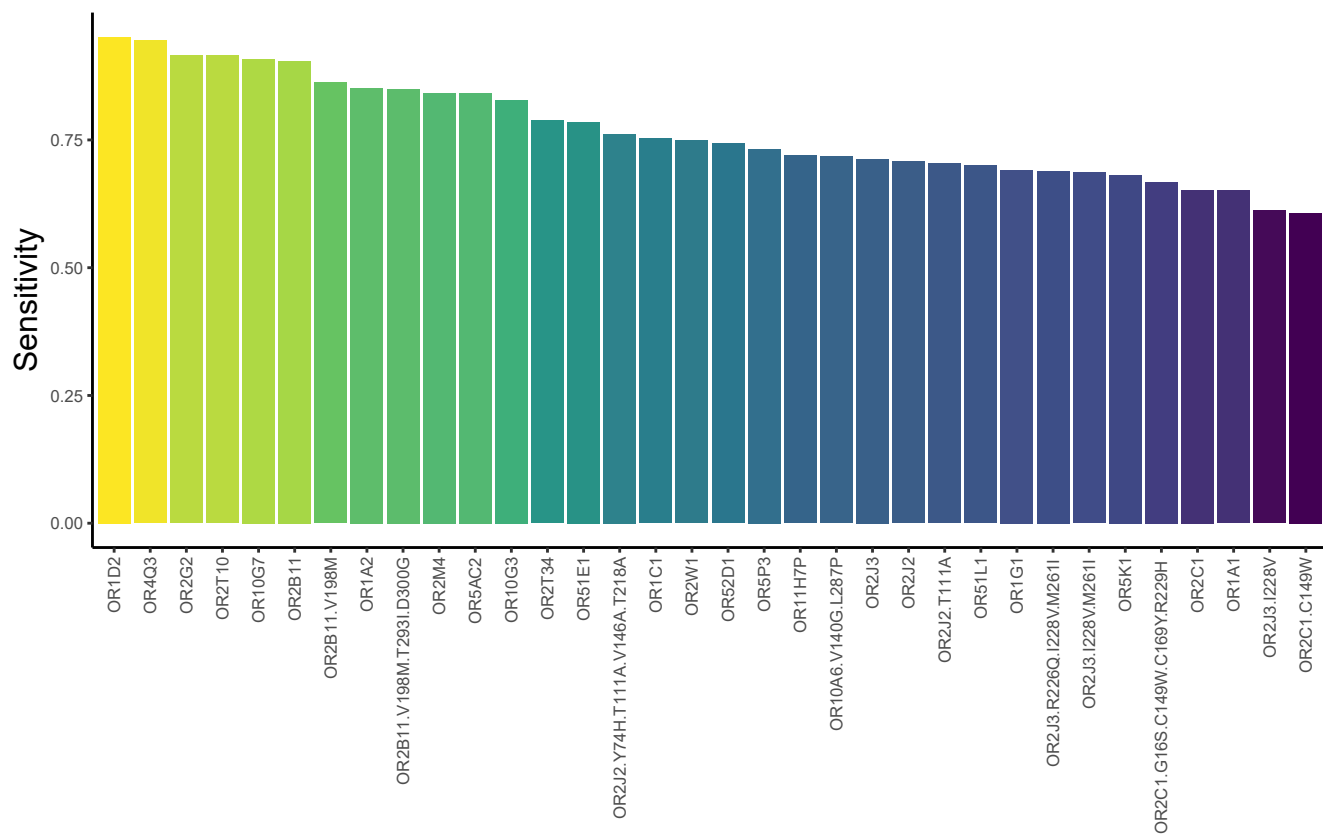
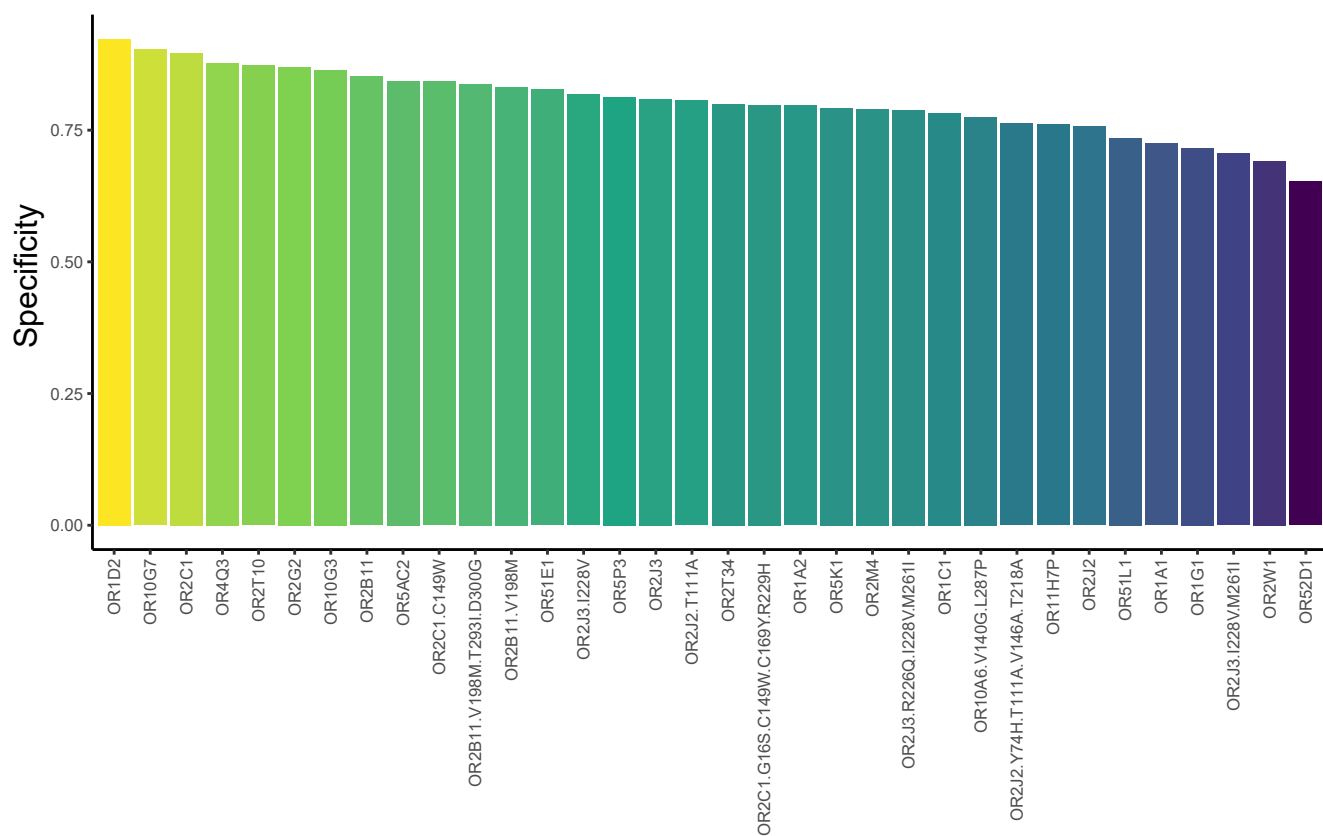
a**b**

Figure S1. Detailed performance of models predicting activity on 34 human ORs. (Related to Figure 1 and Table S1). A) The average sensitivity of the 34 OR models and B) average specificity over repeated cross validation folds (10-fold CV repeated 10 times).

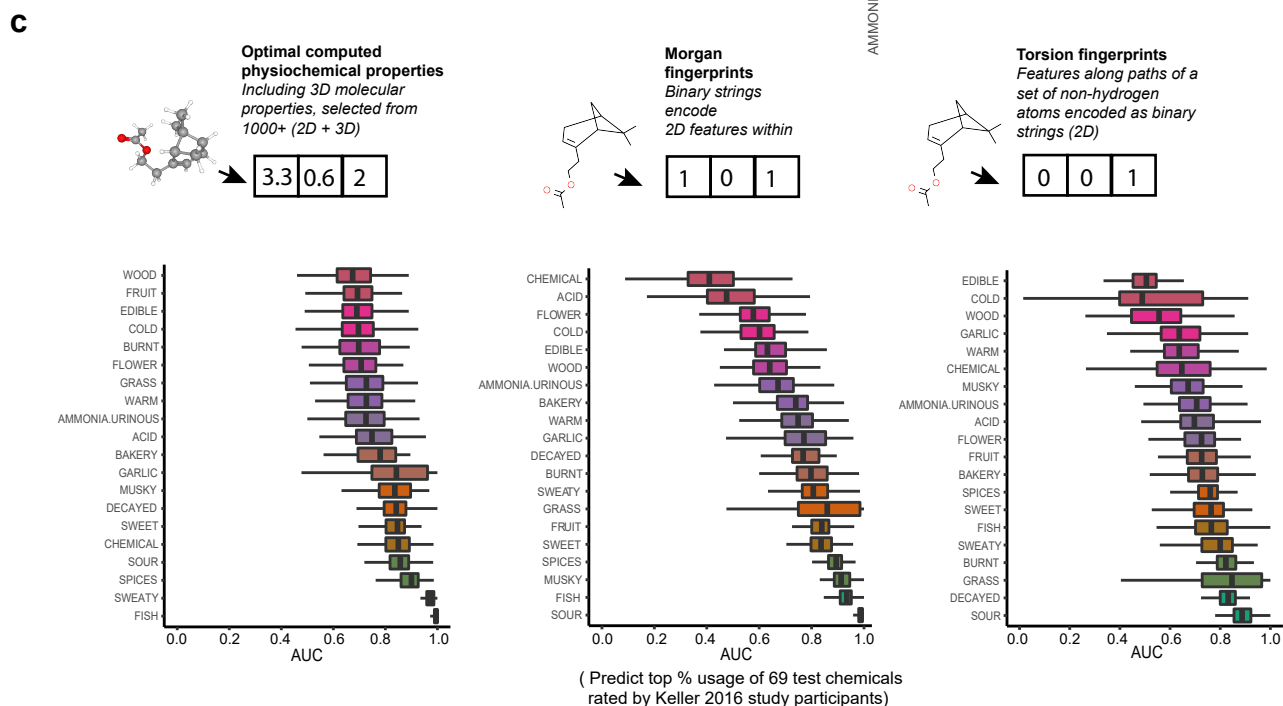
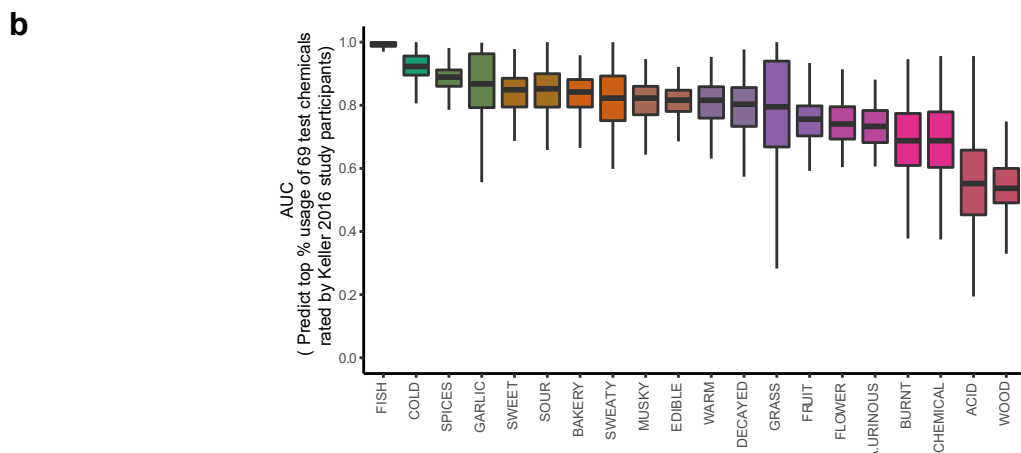
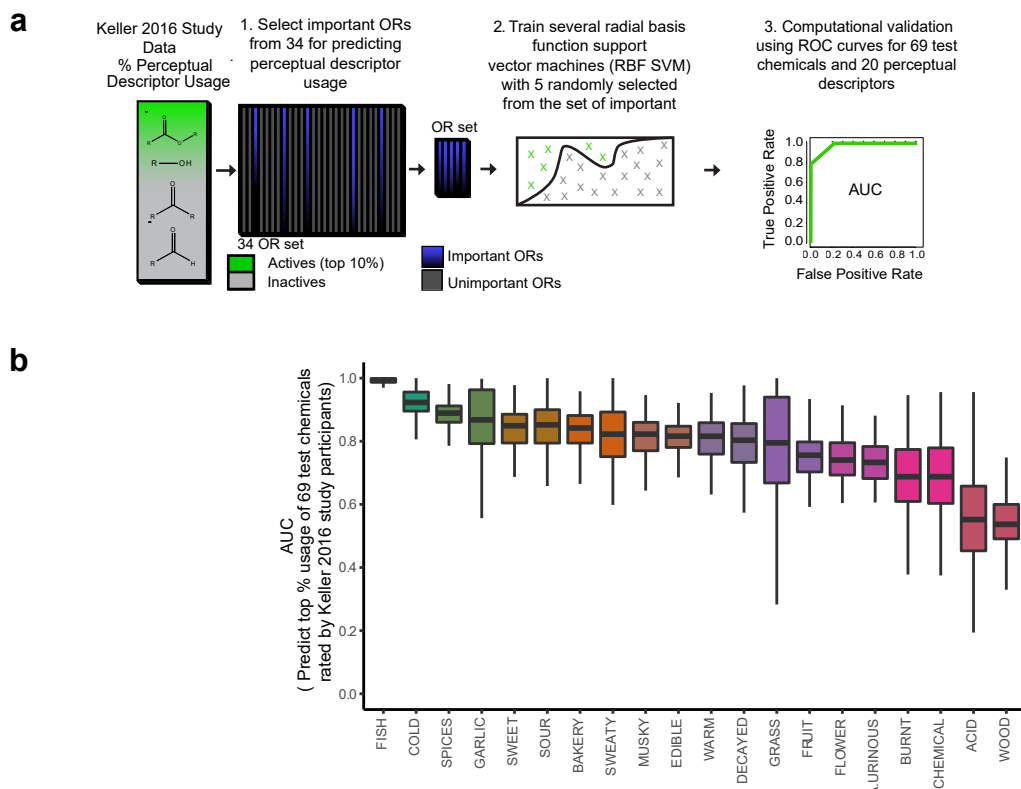


Figure S2. Human OR activity or chemical features predict perceptual data from Keller 2016 study participants. (Related to Figure 2). A) Pipeline for making predictive models for odor perception from ORs for the Keller 2016 perceptual data. Classification cutoffs for the 69 test chemicals are determined from 407 training chemicals. B) Classifying the top 10% of usage for 69 hidden test set chemicals; performance is reported as the area under the ROC curve (AUC). Prediction of the % usage is an aggregate of 5 SVM models, each sampling 5 ORs from the top 10. The OR ranking is determined by recursive feature elimination over cross validation (10-fold repeated 10 times) with 407 training chemicals. C) Prediction of the 69 test chemicals with models trained on various chemical feature representations. Left, physicochemical features are computed for optimized 3D structures and 5 SVM models sample 35 top ranked chemicals features. Plotted performance is the aggregated prediction. Middle, predictions from an SVM model trained on Morgan circular fingerprints. During training, low variance bit positions are dropped to improve the fit. Right, predictions from an SVM model trained on topological torsion fingerprints, dropping low variance bit positions during training. All plots display the standard deviation over 100 bootstrap samples of the 69 test chemicals.

a



Figure S3. Small subsets of ORs optimize predictions of most perceptual descriptors. (Related to Figure 2). A) Comparison between models fit with 10 or 138 ORs on ATLAS study data. Black colored dots show the performance using all ORs while blue dots show the performance using 10 ORs.

a

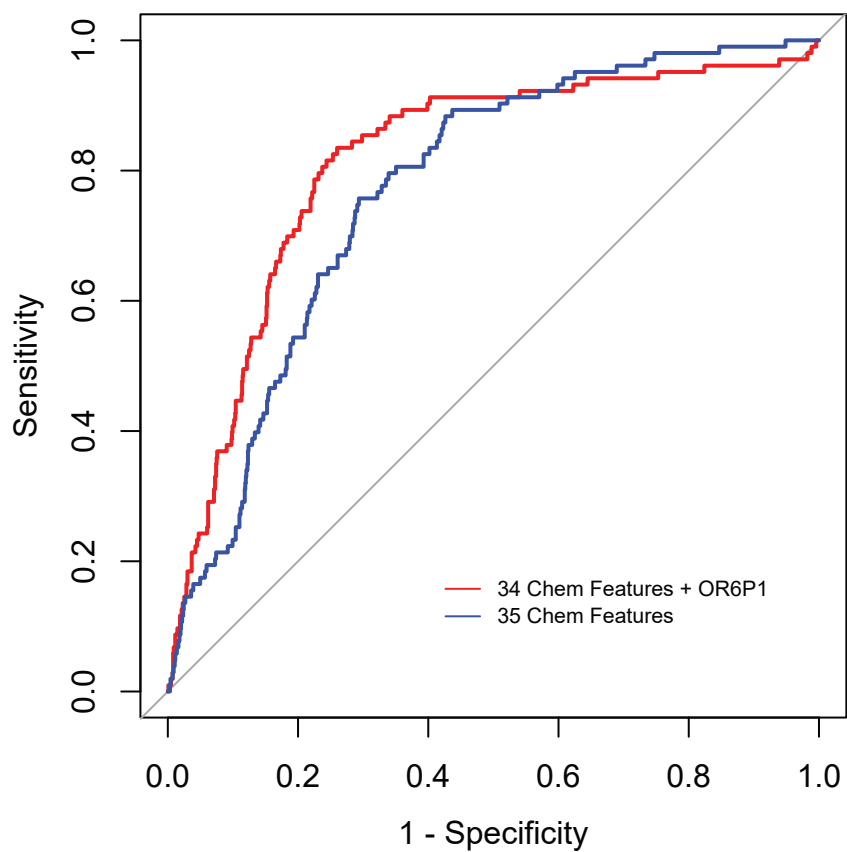


Figure S4. Adding an OR to a chemical feature model improves odor perception predictions. (Related to Figure 4). A) An OR (OR6P1) that was selected for predicting the % usage of the descriptor “Cinnamon” in the ATLAS study is validated with and without physicochemical features

Table S1

OR	Metric	Value
OR10G3	AUC	0.9764246
OR10G7	AUC	0.9571523
OR11H7P	AUC	0.8438965
OR1A2	AUC	0.8741326
OR1C1	AUC	0.7959467
OR1D2	AUC	0.9958090
OR1G1	AUC	0.7593999
OR2B11	AUC	0.9649722
OR2G2	AUC	0.9473866
OR2J2	AUC	0.7880475
OR2J3	AUC	0.8525910
OR2M4	AUC	0.8861412
OR2T10	AUC	0.9438511
OR2T34	AUC	0.8784901
OR2W1	AUC	0.7741072
OR4Q3	AUC	0.9717862
OR51E1	AUC	0.9142729
OR51L1	AUC	0.8378983
OR52D1	AUC	0.7448325
OR5AC2	AUC	0.9059954
OR5K1	AUC	0.7841736
OR5P3	AUC	0.8616590
OR10A6.V140G.L287P	AUC	0.8566833
OR2B11.V198M	AUC	0.9949333
OR2B11.V198M.T293I.D300G	AUC	0.9946667
OR2C1.C149W	AUC	0.9331333
OR2C1.G16S.C149W.C169Y.R229H	AUC	0.8975333
OR2J2.T111A	AUC	0.8400286
OR2J2.Y74H.T111A.V146A.T218A	AUC	0.8713667
OR2J3.I228V.M261I	AUC	0.8373600
OR2J3.I228V	AUC	0.8386250
OR2J3.R226Q.I228V.M261I	AUC	0.8731667
OR1A1	AUC	0.7773069
OR2C1	AUC	0.9636074
OR10G3	Sens	0.8283333
OR10G7	Sens	0.9080000
OR11H7P	Sens	0.7196667
OR1A2	Sens	0.8520000
OR1C1	Sens	0.7533333
OR1D2	Sens	0.9516667
OR1G1	Sens	0.6909206
OR2B11	Sens	0.9042286
OR2G2	Sens	0.9150000
OR2J2	Sens	0.7081444

OR2J3	Sens	0.7126667
OR2M4	Sens	0.8422222
OR2T10	Sens	0.9150000
OR2T34	Sens	0.7876190
OR2W1	Sens	0.7489190
OR4Q3	Sens	0.9450000
OR51E1	Sens	0.7851667
OR51L1	Sens	0.7000000
OR52D1	Sens	0.7425333
OR5AC2	Sens	0.8422222
OR5K1	Sens	0.6816667
OR5P3	Sens	0.7321667
OR10A6.V140G.L287P	Sens	0.7173333
OR2B11.V198M	Sens	0.8626667
OR2B11.V198M.T293L.D300G	Sens	0.8493333
OR2C1.C149W	Sens	0.6066667
OR2C1.G16S.C149W.C169Y.R229H	Sens	0.6680000
OR2J2.T111A	Sens	0.7034286
OR2J2.Y74H.T111A.V146A.T218A	Sens	0.7613333
OR2J3.I228V.M261I	Sens	0.6864000
OR2J3.I228V	Sens	0.6120000
OR2J3.R226Q.I228V.M261I	Sens	0.6893333
OR1A1	Sens	0.6514000
OR2C1	Sens	0.6520000
OR10G3	Spec	0.8646471
OR10G7	Spec	0.9032486
OR11H7P	Spec	0.7610861
OR1A2	Spec	0.7966139
OR1C1	Spec	0.7823186
OR1D2	Spec	0.9231667
OR1G1	Spec	0.7163778
OR2B11	Spec	0.8521181
OR2G2	Spec	0.8702770
OR2J2	Spec	0.7565145
OR2J3	Spec	0.8094917
OR2M4	Spec	0.7900806
OR2T10	Spec	0.8730319
OR2T34	Spec	0.7994472
OR2W1	Spec	0.6918247
OR4Q3	Spec	0.8771373
OR51E1	Spec	0.8272121
OR51L1	Spec	0.7352230
OR52D1	Spec	0.6525470
OR5AC2	Spec	0.8433472
OR5K1	Spec	0.7919028

OR5P3	Spec	0.8127833
OR10A6.V140G.L287P	Spec	0.7745800
OR2B11.V198M	Spec	0.8317600
OR2B11.V198M.T293L.D300G	Spec	0.8378600
OR2C1.C149W	Spec	0.8425800
OR2C1.G16S.C149W.C169Y.R229H	Spec	0.7978600
OR2J2.T111A	Spec	0.8063800
OR2J2.Y74H.T111A.V146A.T218A	Spec	0.7627000
OR2J3.I228V.M261I	Spec	0.7053000
OR2J3.I228V	Spec	0.8181000
OR2J3.R226Q.I228V.M261I	Spec	0.7882000
OR1A1	Spec	0.7249889
OR2C1	Spec	0.8959733

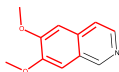
Table S1. Summary of ROC analysis for models predicting activity on 34 human ORs. (Related to Figure 1)

Averages for the prediction performance of Figure 1 models over validation, including the sensitivity (true positive rate), specificity (false positive rate = 1-specificity), and overall AUC.

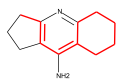
Image

IDs OR

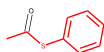
608374 OR10G7



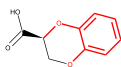
631771 OR1G1



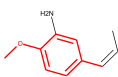
486255 OR2B11



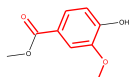
479671 OR2G2



540140 OR2M4



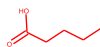
481045 OR2T10



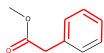
516498 OR2W1



478561 OR51E1



478221 OR51L1



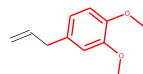
Image

IDs OR

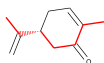
477085 OR52D1



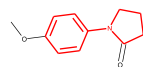
494660 OR5K1



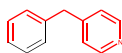
479790 OR5P3



631730 OR10A6.V140G.L287P



535081 OR2B11.V198M



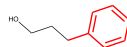
483236 OR2B11.V198M.T293I.D300G



509362 OR2C1.C149W



481575 OR2C1.G16S.C149W.C169Y.R229H



492074 OR2J2.T111A

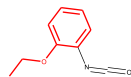


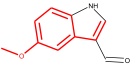



Image	IDs	OR
	591602	OR2J2.Y74H.T111A.V146A.T218A
	490819	OR2J3.I228V
	596094	OR2J3.I228V.M261I
	474648	OR2J3.R226Q.I228V.M261I

Table S2. Enriched substructures among predicted ligands for 34 human ORs (Related to Figure 1). Additional enriched cores/substructures for the labeled ORs, highlighting the core on an exemplar chemical from the eMolecules predictions (related to Figure 1D). The ID is the eMolecules identifier for the representative chemical, which is among the top 10 predictions for the labeled OR. Bonds and atoms are colored black. The enriched substructure is in red.

Perc.Descriptor	AUC.Actual.ORs	AUC.Perm.ORs	P (Actual v. Perm)	AUC.Random.ORs	P (Actual v. Random)
ALCOHOLIC	0.68	0.62	6.624121e-03	0.72	0.723178607
ALMOND	0.94	0.93	2.487726e-01	0.91	0.034877069
AMMONIA	0.82	0.76	7.437701e-02	0.81	0.411008447
ANIMAL	0.62	0.62	4.653889e-01	0.70	0.964195584
ANISE LICORICE	0.78	0.63	9.318801e-03	0.76	0.361514077
APPLE FRUIT	0.56	0.61	6.724894e-01	0.66	0.800616700
AROMATIC	0.71	0.61	7.262549e-02	0.73	0.689074811
BAKERY FRESH BREAD	0.57	0.62	9.059523e-01	0.58	0.719289158
BANANA	0.85	0.64	9.809746e-03	0.87	0.708005236
BARK BIRCH BARK	0.57	0.63	8.966433e-01	0.59	0.615803346
BEANY	0.67	0.67	5.570947e-01	0.61	0.156366133
BEERY	0.61	0.60	3.367454e-01	0.59	0.288927182
BITTER	0.76	0.61	4.906788e-03	0.72	0.274457885
BLACK PEPPER	0.69	0.63	3.041965e-02	0.61	0.060708625
BLOOD RAW MEAT	0.75	0.57	3.027819e-04	0.69	0.131544847
BURNT CANDLE	0.64	0.59	1.217738e-01	0.62	0.299863633
BURNT MILK	0.61	0.60	4.026164e-01	0.62	0.554476157
BURNT PAPER	0.86	0.67	3.743915e-03	0.84	0.304135032
BURNT RUBBER	0.76	0.64	1.367015e-03	0.68	0.006654245
BURNT SMOKY	0.84	0.69	1.459191e-02	0.81	0.130453646
BUTTERY FRESH BUTTER	0.68	0.76	9.155957e-01	0.66	0.366338507
CADAVEROUS DEAD ANIMAL	0.92	0.83	1.386069e-04	0.92	0.368750235
CAMPHOR	0.63	0.57	4.832609e-02	0.70	0.945089522
CANTALOUPE HONEY DEW MELON	0.79	0.73	3.018547e-02	0.82	0.767234025
CARAMEL	0.89	0.81	8.795053e-04	0.86	0.175338953
CARAWAY	0.82	0.57	9.910237e-07	0.82	0.467041687
CARDBOARD	0.63	0.60	1.545006e-01	0.65	0.643885830
CAT URINE	0.55	0.62	9.862892e-01	0.59	0.887941827
CEDARWOOD	0.78	0.66	7.725876e-02	0.76	0.392239690
CELERY	0.61	0.70	9.874586e-01	0.66	0.861059602
CHALKY	0.60	0.59	4.697072e-01	0.58	0.362309551
CHEESY	0.79	0.58	5.814025e-04	0.85	0.969998448
CHEMICAL	0.80	0.72	5.454888e-04	0.76	0.187394329
CHERRY BERRY	0.82	0.65	1.870925e-03	0.82	0.400904651
CHOCOLATE	0.65	0.79	9.961704e-01	0.74	0.953634802
CINNAMON	0.86	0.78	3.247772e-02	0.73	0.003079151
CLEANING FLUID	0.76	0.75	2.907639e-01	0.81	0.921645505
CLOVE	0.88	0.60	1.088595e-04	0.85	0.264212595
COCONUT	0.61	0.59	2.942493e-01	0.60	0.373755189
COFFEE	0.74	0.57	1.325299e-04	0.76	0.689137219
COLOGNE	0.82	0.74	4.557177e-02	0.75	0.034619751
COOKED VEGETABLES	0.79	0.60	5.312181e-04	0.79	0.539445780
COOL COOLING	0.77	0.64	5.920177e-03	0.80	0.795019055
CORK	0.70	0.61	5.878820e-02	0.61	0.022596784
CREOSOTE	0.80	0.81	6.339020e-01	0.83	0.893222941
CRUSHED GRASS	0.69	0.62	1.174218e-02	0.64	0.143682506
CRUSHED WEEDS	0.74	0.61	4.921776e-03	0.66	0.013566817
DILL	0.57	0.73	9.992697e-01	0.59	0.781535873
DIRTY LINEN	0.73	0.61	7.730652e-04	0.67	0.044159634
DISINFECTANT CARBOLIC	0.79	0.90	9.958961e-01	0.89	0.998115114
DRY POWDERY	0.64	0.63	4.372117e-01	0.62	0.342789968
EGGY FRESH EGGS	0.59	0.56	7.667215e-02	0.59	0.461463704
ETHERISH ANAESTHETIC	0.96	0.90	2.183660e-03	0.93	0.027905360
EUCALIPTUS	0.69	0.57	1.904247e-03	0.76	0.981148786
FECAL LIKE MANURE	0.70	0.77	9.765591e-01	0.71	0.546529155
FERMENTED ROTTEN FRUIT	0.83	0.64	1.552895e-04	0.73	0.068225780

Perc.Descriptor	AUC.Actual.ORs	AUC.Perm.ORs	P (Actual v. Perm)	AUC.Random.ORs	P (Actual v. Random)
FISHY	0.76	0.58	4.824185e-04	0.77	0.584032635
FLORAL	0.75	0.81	8.525523e-01	0.76	0.571314301
FRAGRANT	0.67	0.54	5.315367e-02	0.57	0.002180182
FRESH GREEN VEGETABLES	0.74	0.61	8.271564e-04	0.72	0.251533345
FRESH TOBACCO SMOKE	0.63	0.57	5.314839e-02	0.64	0.612188603
FRIED CHICKEN	0.64	0.62	2.208670e-01	0.64	0.503480983
FRUITY CITRUS	0.77	0.74	1.227108e-01	0.77	0.415384866
FRUITY OTHER THAN CITRUS	0.75	0.54	2.031716e-04	0.76	0.545991914
GARLIC ONION	0.97	0.63	2.531550e-05	0.96	0.231282678
GASOLINE SOLVENT	0.92	0.85	2.795957e-03	0.90	0.129023122
GERANIUM LEAVES	0.62	0.54	8.120492e-02	0.65	0.748064731
GRAINY AS GRAIN	0.65	0.68	7.357469e-01	0.61	0.241183239
GRAPE JUICE	0.60	0.63	8.148000e-01	0.58	0.249827462
GRAPEFRUIT	0.73	0.69	6.902408e-02	0.76	0.832562571
HAY	0.63	0.58	2.191222e-01	0.64	0.635096113
HEAVY	0.75	0.52	8.808084e-03	0.70	0.180959523
HERBAL GREEN CUT GRASS	0.66	0.57	2.178443e-02	0.61	0.135613486
HONEY	0.86	0.72	6.441975e-03	0.84	0.105893090
HOUSEHOLD GAS	0.91	0.81	1.735901e-05	0.93	0.730659466
INCENSE	0.74	0.75	5.484723e-01	0.77	0.716919634
KEROSENE	0.94	0.76	1.013985e-05	0.93	0.327626884
KIPPERY SMOKED FISH	0.57	0.59	7.332818e-01	0.58	0.577061232
LAUREL LEAVES	0.63	0.56	8.699926e-02	0.67	0.838973837
LAVENDER	0.73	0.71	2.760870e-01	0.74	0.619993047
LEATHER	0.68	0.61	1.345040e-02	0.79	0.992360918
LEMON	0.87	0.87	2.979404e-01	0.85	0.136759987
LIGHT	0.77	0.73	1.434928e-01	0.85	0.917788531
MALTY	0.66	0.70	7.085725e-01	0.65	0.398116520
MAPLE SYRUP	0.81	0.62	6.505141e-04	0.78	0.196660862
MEATY COOKED GOOD	0.68	0.55	1.619274e-03	0.69	0.532624907
MEDICINAL	0.81	0.70	2.199711e-02	0.72	0.094122938
METALLIC	0.62	0.59	3.428863e-01	0.63	0.625422059
MINTY PEPPERMINT	0.75	0.59	4.713820e-04	0.67	0.005963238
MOTHBALLS	0.76	0.75	4.015789e-01	0.80	0.939709284
MOUSE	0.85	0.69	6.052005e-05	0.74	0.001296670
MUSHROOM	0.67	0.61	5.507031e-02	0.68	0.612807182
MUSK	0.64	0.61	2.600264e-01	0.65	0.641149881
MUSTY EARTHY MOLDY	0.54	0.59	8.496130e-01	0.56	0.726368591
NAIL POLISH REMOVER	0.93	0.93	4.717329e-01	0.87	0.008637722
NEW RUBBER	0.69	0.60	1.422405e-02	0.69	0.488580423
NUTTY WALNUT ETC	0.79	0.80	6.001993e-01	0.79	0.541252697
OAK WOOD COGNAC	0.71	0.60	5.334112e-02	0.70	0.449304705
OILY FATTY	0.75	0.68	5.442055e-02	0.78	0.703181219
ORANGE	0.70	0.63	4.857428e-02	0.70	0.487048230
PAINT	0.91	0.79	8.734811e-05	0.91	0.461934408
PEACH FRUIT	0.79	0.58	5.484548e-05	0.76	0.266667783
PEANUT BUTTER	0.83	0.78	1.108906e-01	0.86	0.826427173
PEAR	0.86	0.80	1.820044e-02	0.84	0.381644634
PERFUMERY	0.82	0.78	1.393034e-02	0.81	0.429208350
PINEAPPLE	0.83	0.83	4.383805e-01	0.82	0.400655067
POPCORN	0.59	0.68	9.254312e-01	0.65	0.890432139
PUTRID FOUL DECAYED	0.87	0.80	2.422996e-03	0.92	0.971155085
RAISINS	0.57	0.59	7.519006e-01	0.60	0.789009023
RANCID	0.80	0.69	1.046710e-02	0.80	0.573930524
RAW CUCUMBER	0.84	0.61	5.236949e-04	0.76	0.027564181
RAW POTATO	0.63	0.56	8.758358e-03	0.66	0.714216712
ROPE	0.60	0.58	3.075816e-01	0.61	0.668058005

Perc.Descriptor	AUC.Actual.ORs	AUC.Perm.ORs	P (Actual v. Perm)	AUC.Random.ORs	P (Actual v. Random)
ROSE	0.67	0.59	1.072022e-01	0.65	0.330191386
SAUERKRAUT	0.86	0.88	8.345338e-01	0.83	0.187159140
SEASONING FOR MEAT	0.61	0.58	1.750854e-01	0.59	0.258163703
SEMINAL SPERM LIKE	0.68	0.59	2.868288e-02	0.75	0.911515570
SEWER	0.97	0.85	1.684756e-03	0.97	0.404678453
SHARP PUNGENT ACID	0.89	0.84	7.879184e-03	0.87	0.285085817
SICKENING	0.97	0.86	9.330160e-04	0.96	0.173703843
SOAPY	0.81	0.78	2.807849e-01	0.77	0.176364339
SOOTY	0.64	0.57	6.044304e-02	0.68	0.830745272
SOUPY	0.65	0.61	2.153393e-01	0.57	0.201392365
SOUR MILK	0.98	0.88	1.311887e-03	0.97	0.395627632
SOUR VINEGAR	0.86	0.78	2.440581e-02	0.88	0.781798703
SPICY	0.91	0.73	1.473271e-04	0.85	0.044485570
STALE	0.81	0.71	3.902490e-03	0.80	0.384288452
STALE TOBACCO SMOKE	0.68	0.65	1.798344e-01	0.71	0.741998826
STRAWBERRY	0.88	0.56	1.729606e-04	0.85	0.273525575
SULFIDIC	0.94	0.75	8.166529e-04	0.93	0.298005800
SWEATY	0.68	0.61	6.619308e-02	0.70	0.661394584
SWEET	0.73	0.65	3.036319e-02	0.72	0.397528666
TAR	0.67	0.66	3.736882e-01	0.66	0.451093414
TEA LEAVES	0.61	0.62	5.105959e-01	0.57	0.113352387
TURPENTINE PINE OIL	0.83	0.72	3.150109e-04	0.80	0.237355229
URINE	0.74	0.62	1.306574e-03	0.76	0.694702395
VANILLA	0.75	0.72	9.793389e-02	0.69	0.180342803
VARNISH	0.97	0.83	6.922921e-04	0.94	0.017280947
VIOLETS	0.71	0.74	7.678331e-01	0.71	0.449119936
WARM	0.72	0.57	1.437458e-03	0.70	0.319931645
WET PAPER	0.60	0.69	9.795118e-01	0.62	0.673009454
WET WOOL WET DOG	0.78	0.59	1.104795e-03	0.74	0.098478406
WOODY RESINOUS	0.65	0.60	1.157541e-01	0.68	0.734801220
YEASTY	0.86	0.84	2.672057e-01	0.89	0.690830166

Table S3. Detailed analysis of odor perception predictions using chemical features and ORs (Related to Figure 4). Combined OR and chemical feature model performance using the ATLAS study data. In one condition, ORs are replaced with those of lesser importance (“Random”). In the second condition, OR activities for the best combined set are permuted (shuffled). The chemical features are intact in both conditions. Training and testing chemicals are equivalent.

Best Predicted	Metric	Chem Features	Chem Features + ORs
Top 5	AUC	0.9502134	0.9708750
Top 10	AUC	0.9210808	0.9565362
Top 20	AUC	0.8969860	0.9253587
Top 25	AUC	0.8822129	0.9117761
Top 50	AUC	0.8314386	0.8558157
Top 5	R	0.6559822	0.7265567
Top 10	R	0.5977552	0.6949485
Top 20	R	0.5818210	0.6447922
Top 25	R	0.5734141	0.6267827
Top 50	R	0.5271159	0.5474018

Table S4. Comparing predictions of odor perception with chemical features or chemical features and ORs. (Related to Figures 2 and 4). Summary table containing the average test performance for the best predicted perceptual descriptors in the ATLAS study across two metrics (R and AUC) and using different predictor sets (e.g. OR vs chemical features). R is the correlation between the predicted and observed % usage of the perceptual descriptors. The AUC is the classification success for chemicals in the top 10% of usage.

SUPPLEMENTAL INFORMATION

Transparent Methods

Modeling OR ligands from chemical features (Figure 1)

We trained SVM models to learn physicochemical features of the confirmed ligands for a subset of ORs whose response profiles are currently better characterized (34 total). Different chemical features were encoded as binary fingerprints (1,0) (Klekota-Roth (Klekota and Roth, 2008), Morgan/Circular (Morgan, 1965), MACCs, Shortest Path, and Hybridization (Steinbeck et al., 2003). Chemical fingerprints can encode up to ~1000 bits and many are possibly uninformative. Kullback–Leibler (KL) divergence (Nisius and Bajorath, 2010) was used to select only those bits that maximized the distance between active and inactive compounds in the heterologous assay data. Predictions from these models provided probability scores for each OR-chemical pair for the ATLAS chemicals. This work relied on the chemistry development kit (CDK) (Steinbeck et al., 2003) as well as its R interface (Guha, 2007). The structural similarity filter to count the number of predicted ligands used functions in the ChemmineR package in R (Cao et al., 2008).

Enriched Substructures/Cores (Figure 1)

Enriched cores were analyzed using RDKit through Python (Landrum, 2006). The algorithm is an exhaustive search for the maximum common substructure among chemicals. In practice, larger chemical sets often yield less substantive cores. To remedy this, the algorithm includes a threshold parameter that relaxes the proportion of chemicals containing the core. We used a threshold of .5, requiring that half of the top predicted chemicals contained the core.

ORs as predictors of perception (Figures 2-4)

Despite several available data sources, most in vitro assays typically report a handful of ORs with multiple ligands and many others with few ligands (1 or 2 compounds that pass statistical thresholds). To incorporate the more narrowly tuned receptors, we computed an approximation of the 3D pharmacophore kernel (Mahé et al., 2006). Pharmacophore kernels are a versatile method for computing pairwise similarities among chemicals according to a set of standard features that are related to biological activity. Namely, similarity between ATLAS chemicals and known OR ligands was defined by the three-point Tanimoto coefficient, which is scaled to 0-1, with 1 being maximally similar. In cases where there were > 1 ligands for an OR the maximally similar ligand was used.

To incorporate the ORs with more ligands, we trained SVM models on physicochemical features of odorants with known activity. There were 34 ORs with sufficient training data for this approach. These models assigned probability scores for the 34 ORs to the perceptual study chemicals (ATLAS and Keller 2016). The Keller 2016 perceptual ratings were converted to the % usage, or the % of participants using a perceptual descriptor; that is, supplying a rating (0-100) for a given descriptor. The ATLAS study provides this metric.

The receiver operating characteristic (ROC) analysis or, in particular, the area under the curve (AUC) is based on transforming the rating that had been assigned to a perceptual descriptor by study participants into a classification label (active/inactive). The active chemicals are those within the top 10% of the ratings (% usage). However, as this cutoff is arbitrary, other metrics are supplied in supplementary materials for comparison. These, in addition to the classification-based metric (ROC analysis), are explained in detail in the metrics section alongside their strengths and weaknesses for this specific problem. Unless noted in the figure legends the importance of an OR is not based on classification. Specific methods for evaluating importance are discussed below.

Computing chemical features to predict perceptual descriptors (Figures 4, 6, S2, S4)

We computed chemical features using the Python wrapper for the open source RDKit software (Landrum, 2006). This included chemicals features that were raw values, pertaining to features such as functional group counts and 3D geometries, which closely resemble the proprietary DRAGON software; the whole library is accessible through the mordred module (Moriwaki et al., 2018). We also computed Morgan/circular (radius =2) and topological torsion fingerprints. These use a hash function to encode different chemical features as fixed length binary strings (1024 bits).

Selecting important ORs in prediction of human perception (Figures 2-4, S2-S3)

Important ORs were selected using a cross validated recursive feature elimination (10-fold, repeated 10 times), with the random forest (RF) algorithm or the support vector machine (SVM) algorithm. Random forest defines importance by permuting predictors and reporting the % increase in error. Random forest fits multiple decision trees on different bootstrap samples and supplies a consensus vote over the trees as the prediction. Bootstrap sampling leads to a portion of data being left out; the “out of bag” sample which is used to estimate the prediction performance. When a model is fit, the predictor importance (% increase in error) is computed. The support vector machine, however, does not include an ‘out of bag’ sample and therefore the OR/chemical feature importance is computed externally by fitting non-linear regression models for each predictor.

By including the model-fitting inside a cross validation loop the importance is computed over multiple folds or portions of the training data rather than on the complete training set, which reduces bias in the predictors that are selected. The importance is in this context redefined as a selection rate (e.g. the rate the predictor was highly ranked).

Clustering (Figure 5)

Clustering was performed with the hcust function in R using the Ward D2 method and the Euclidean distance for numerical matrices such as the perceptual ratings (Figure 5A) or 1-Jaccard distances for binary matrices (Figure 5B-D). 1000 bootstrap samples were used to select the optimal number of clusters, according to the gap statistic (1-standard error (SE) rule).

Quantification and statistical analysis

Support Vector Machine

Training the support vector machine (SVM) involves identifying a set of parameters that optimize a cost function, where cost 1 and cost 0 correspond to training chemicals labeled as “Active” and “Inactive,” respectively.

$$SVM\ Cost = \min_{\theta} C \sum_{i=1}^m y^{(i)} cost_1(\theta^T f^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T f^{(i)}) + \frac{1}{2} \sum_{j=1}^n \theta_j^2$$

Additionally, a kernel determines the shape of the decision boundary between the active and inactive chemicals from the training set. The radial basis function (RBF) or Gaussian kernel enables the learning of more complex, non-linear boundaries. It is therefore well suited for problems in which the physicochemical properties vary among the biologically active chemicals. This kernel computes the similarity for each chemical (x) and a set of landmarks (l), where σ^2 is a tunable parameter determined by the problem and data. The similarity with respect to these landmarks is used to predict new chemicals (“Active” vs. “Inactive”).

$$\text{Gaussian Kernel} = \exp\left(-\frac{\|x - l^{(1)}\|^2}{2\sigma^2}\right)$$

Metrics

The area under the roc curve (**AUC**) assesses the true positive rate (TPR or sensitivity) as a function of the false positive rate (FPR or 1-specificity) while varying the probability threshold (T) for a label (Active/Inactive). If the computed probability score (x) is greater than the threshold (T), the observation is assigned to the active class. Integrating the curve provides an estimate of classifier performance, with the top left corner giving an AUC of 1.0 denoting maximum sensitivity to detect all targets or actives in the data without any false positives. The theoretical random classifier is reported at AUC = 0.5.

$$\text{TPR}(T) = \int_T^\infty f_1(x) dx$$

$$\text{FPR}(T) = \int_T^\infty f_0(x) dx$$

Where T is a variable threshold and x is a probability score

However, we generated classifiers that are more authentic than theoretical random classification, shuffling the chemical feature (or OR) values in the models and statistically comparing the mean AUCs across multiple partitions of the data. This controls against optimally tuned algorithms predicting well simply because of specific predictor attributes (e.g. range, mean, median, and variance) or models that are of a specific size (number of predictors) performing well even with shuffled values. Additionally, biological data sets are often small, with stimuli or chemicals that—rather than random selection—reflect research biases, possibly leading to optimistic validation estimates without the proper controls. We used the AUC with classification-based training, such as to predict binary labels (Active/Inactive). For classification-based training we initially converted the % usage into a binary label (Active/Inactive) using the top 10% of the distribution as the cutoff. The basis for a classification-based performance metric was the often top-heavy distribution of the % usage. It is for instance possibly not as relevant for models to accurately predict chemicals with minimal % usage. Rather, it is preferable for models to accurately predict whether a chemical will smell “Sweet” or not.

To provide further clarity we also reported multiple performance metrics including the correlation between the predicted and observed % usage, the root mean squared error (RMSE), and mean absolute error (MAE): **RMSE**: Root mean squared error is the square root of the mean difference between predicted values and those observed. It is the average prediction error on the same scale as the target or outcome being predicted. We supplied this metric because the correlation coefficient (R) is not always an accurate representation of model performance and classification of exemplar chemicals required an arbitrary cutoff (e.g. 90th percentile). We reported the correlation coefficient, R , between the predicted and observed % usage due to its previous use with human perceptual data. **MAE**: Mean absolute error is the mean of the absolute difference between predicted and observed. It thus assigns equal weight to all prediction errors, whether large or small.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{N}}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}|; \text{ where, } \hat{y} = \text{predicted and } y = \text{observed}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}; \text{ where, } TP = \text{True Positive and } FN = \text{False Negative}$$

$$\text{Specificity} = \frac{TN}{TN+FP}; \text{ where, } TN = \text{True Negative and } FP = \text{False Positive}$$

Supplemental References

- Cao, Y., Charisi, A., Cheng, L.C., Jiang, T., Girke, T., 2008. ChemmineR: A compound mining framework for R. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btn307>
- Guha, R., 2007. Chemical informatics functionality in R. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v018.i05>
- Klekota, J., Roth, F.P., 2008. Chemical substructures that enrich for biological activity. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btn479>
- Landrum, G., 2006. RDKit: Open-source cheminformatics. <http://www.rdkit.org>.
- Mahé, P., Ralaivola, L., Stoven, V., Vert, J.P., 2006. The pharmacophore kernel for virtual screening with support vector machines. *J. Chem. Inf. Model.* 46, 2003–2014. <https://doi.org/10.1021/ci060138m>
- Morgan, H.L., 1965. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* 5, 107–113. <https://doi.org/10.1021/c160017a018>
- Moriwaki, H., Tian, Y.S., Kawashita, N., Takagi, T., 2018. Mordred: A molecular descriptor calculator. *J. Cheminform.* <https://doi.org/10.1186/s13321-018-0258-y>
- Nisius, B., Bajorath, J., 2010. Reduction and recombination of fingerprints of different design increase compound recall and the structural diversity of hits. *Chem. Biol. Drug Des.* <https://doi.org/10.1111/j.1747-0285.2009.00930.x>
- Steinbeck, C., Han, Y., Kuhn, S., Horlacher, O., Luttmann, E., Willighagen, E., 2003. The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics, in: *Journal of Chemical Information and Computer Sciences*. <https://doi.org/10.1021/ci025584y>