Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/4/8/eaaq1458/DC1

Supplementary Materials for

Hyperbolic geometry of the olfactory space

Yuansheng Zhou, Brian H. Smith, Tatyana O. Sharpee*

*Corresponding author. Email: sharpee@salk.edu

Published 29 August 2018, *Sci. Adv.* **4**, eaaq1458 (2018)
DOI: 10.1126/sciadv.aaq1458

This PDF file includes:

Supplementary Information

Supplementary Figures

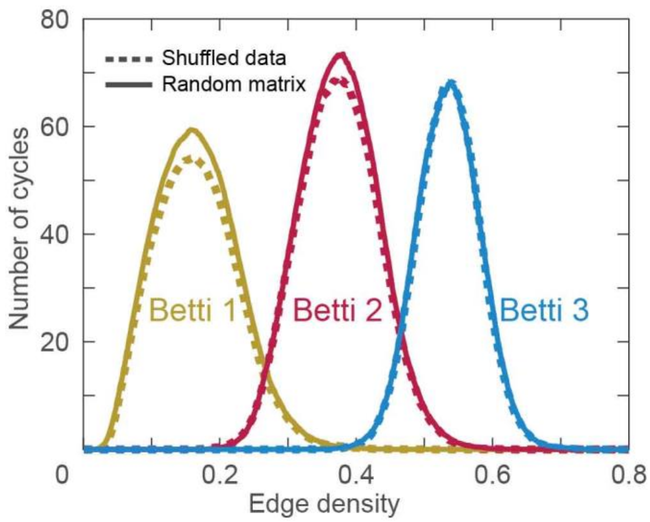


Fig. S1. No indications of hyperbolic geometry in shuffled odor data sets. Betti curves from shuffled blueberry data can be accounted for by random sampling and is not consistent with the Hyperbolic or Euclidean models. The correlation matrix between odors was computed by taking measurements of odor pairs from separate, randomly selected fruit samples. This removes correlations in the fluctuations of component concentrations between odors. Computing Betti curves from such shuffled correlation matrices produces Betti curves that are fully consistent with random matrices (p=0.4, 0.7, 0.9 for Betti curves one, two and three, respectively) and are not consistent ($p<0.02$) with either Euclidean or hyperbolic spaces with small noise amounts as necessary to account for real (unshuffled) data.

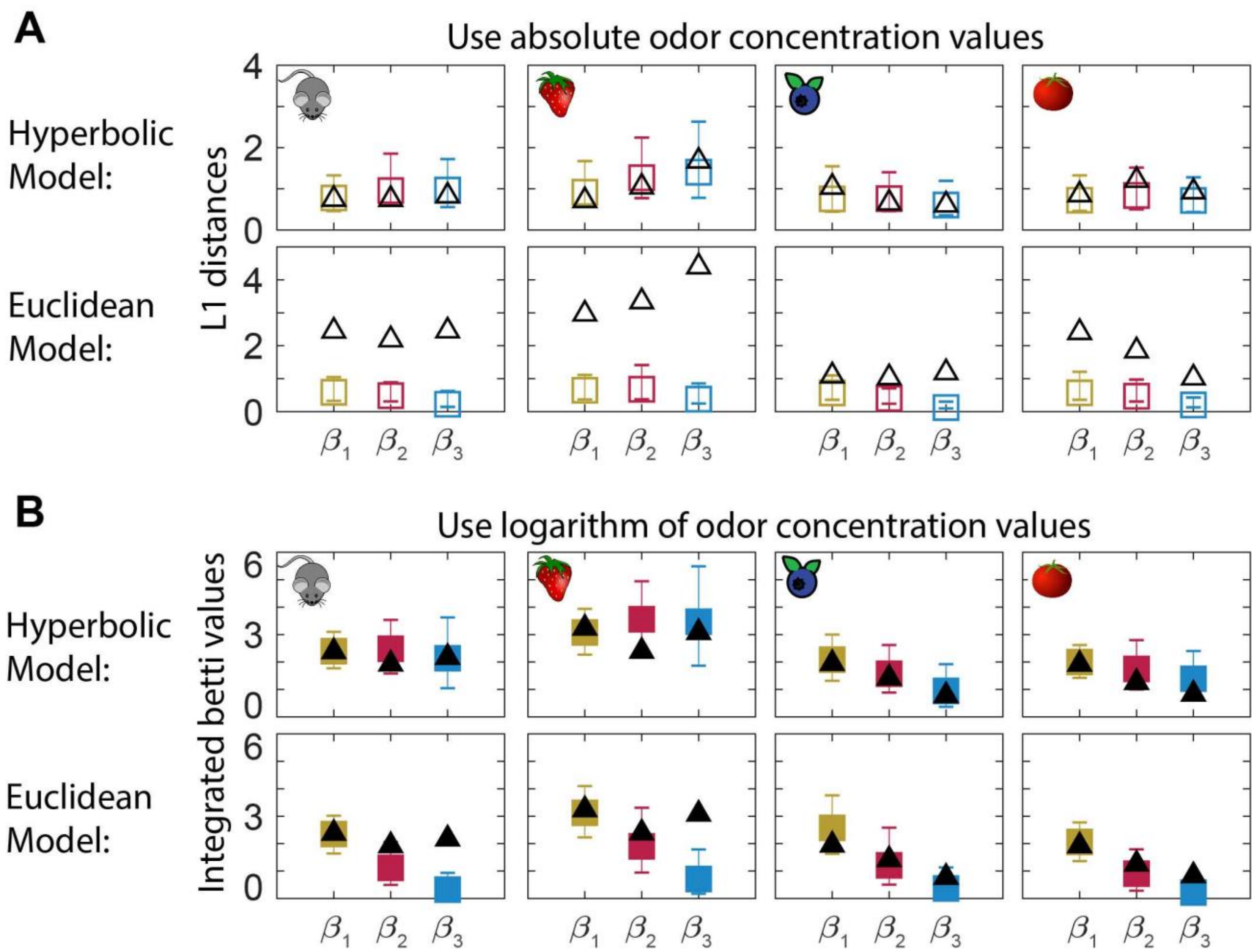
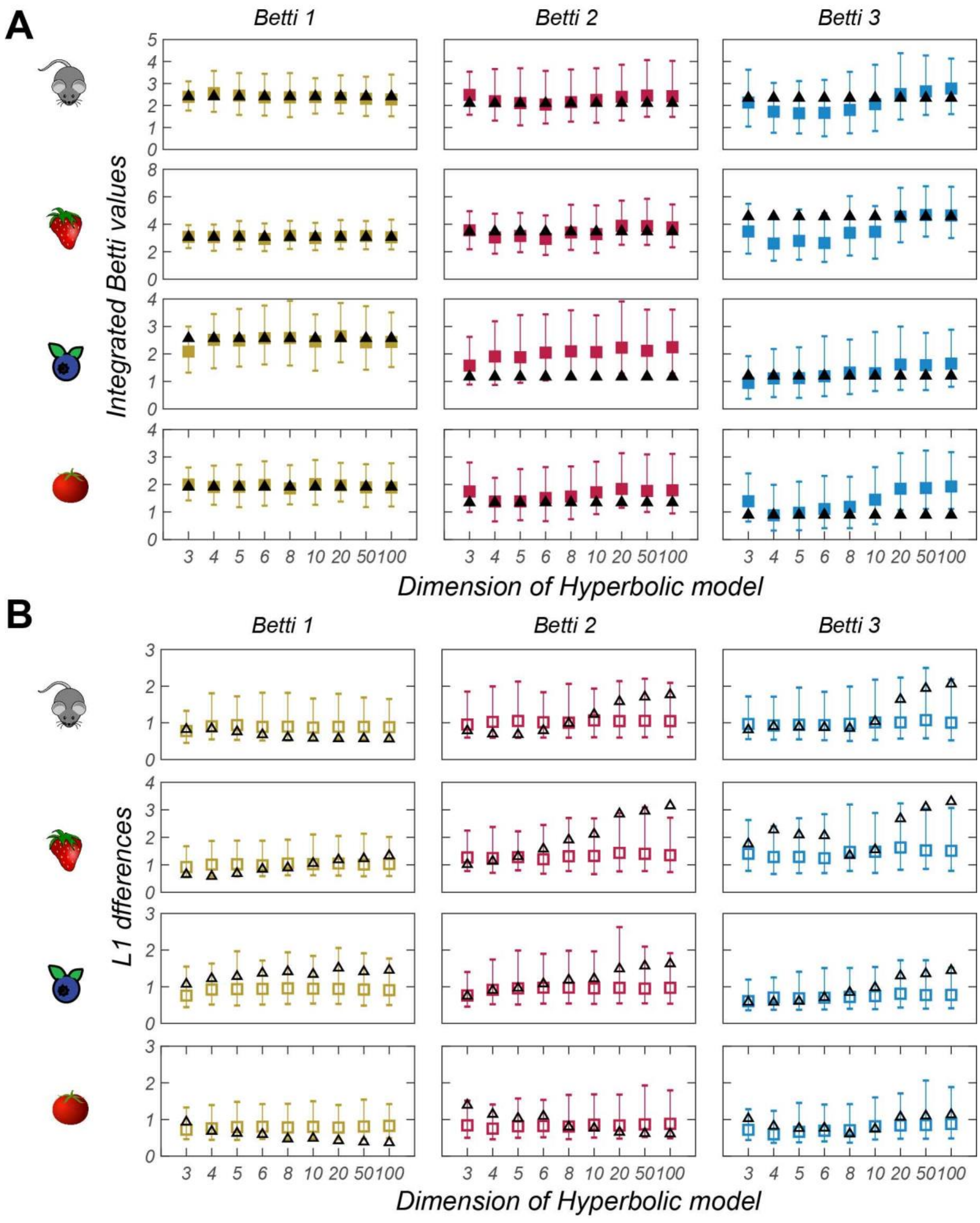


Fig. S2. Alternative ways of evaluating differences between Betti curves also support hyperbolic geometry of natural odor spaces. Error bar plots of Betti curves statistics of both geometric models using L1 distances **(A)** or logarithm of concentrations and integrated Betti values **(B)**. The same geometric parameters were used as in Fig. 2. (**A**) The gray triangles showed the L1 distance between Betti curves of data and the mean of 300 geometric models, the error bar plots showed the statistics of the L1 distances between Betti curves of all of the 300 models and model mean. The error bar showed the 95% confidence intervals (2.5% ~ 97.5%). (**B**) Use the logarithms of odors concentrations to generate Betti curves. The gray triangles and colored error bar plots show the statistics of integrated betti values in both geometric models.



A
Betti 1
Betti 2
Betti 3
Integrated Betti values
Dimension of Hyperbolic model
B
Betti 1
Betti 2
Betti 3
L1 dfferences
Dimension of Hyperbolic model

Fig. S3. Error bar plots of Betti curves statistics for the hyperbolic model of different dimensions. The parameters were determined by fitting the first integrated Betti value of the models to data (first columns), then the same parameters were used to evaluate the correspondence with the second and third Betti curves. (**A**) Error bar plots of integrated Betti values of data (gray triangles) and 300 model repeats (colored bars) with dimension ranging from 3D to 100D. (**B**) Error bar plots of L1 difference of Betti curves between data and model mean (gray triangles), and the distances between all 300 models and model mean (colored bars). The parameters were the same as in Fig. 2 and fig. S2.

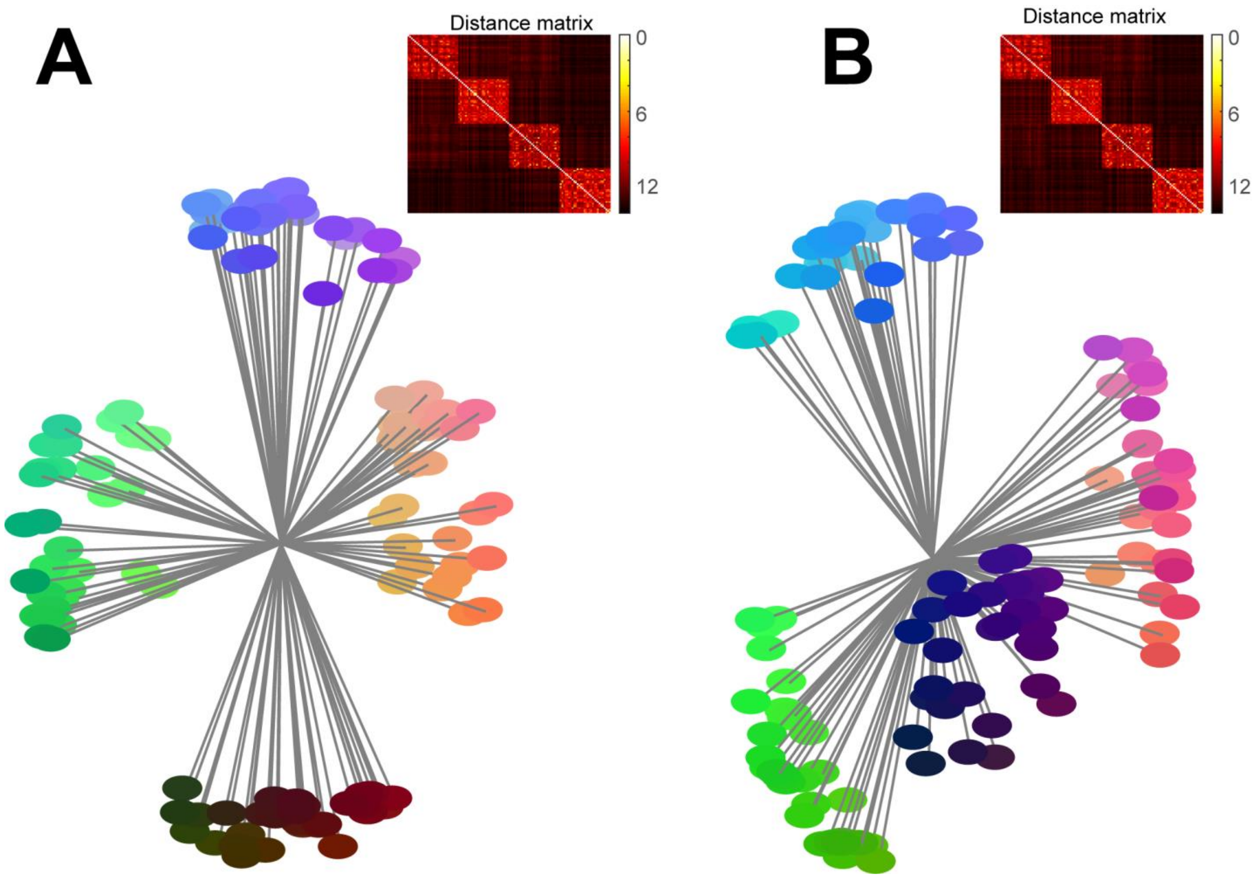


Fig. S4. Test of the nonmetric multidimensional scaling algorithm in the hyperbolic space on synthetic data. (**A**) 120 sampling points were generated near the surface of 3D hyperbolic sphere forming four clusters. The radii were distributed uniformly within 0.9 of the sphere radius, the same distribution we used to model natural odor mixtures. Inset in the top right shows the matrix of pairwise distances that indicates these four clusters. (**B**) Nonmetric multidimensional scaling can be used to embed points on the surface of a 3D hyperbolic sphere. The embedded points also form four clusters, albeit at different orientation in the space. The distance matrix (inset) is also reproduced.

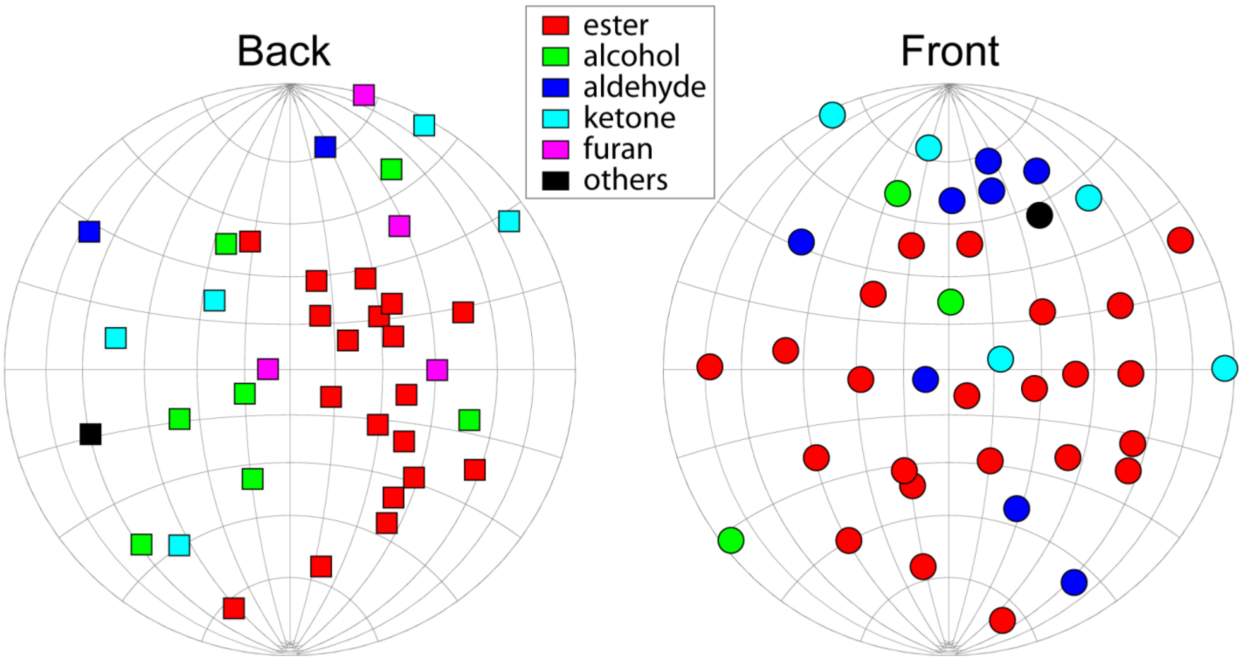


Fig. S5. Odors within the identified space do not cluster by functional group. The odors shown are from the strawberry data set. Circles and squares show the front and back side of the sphere, respectively.

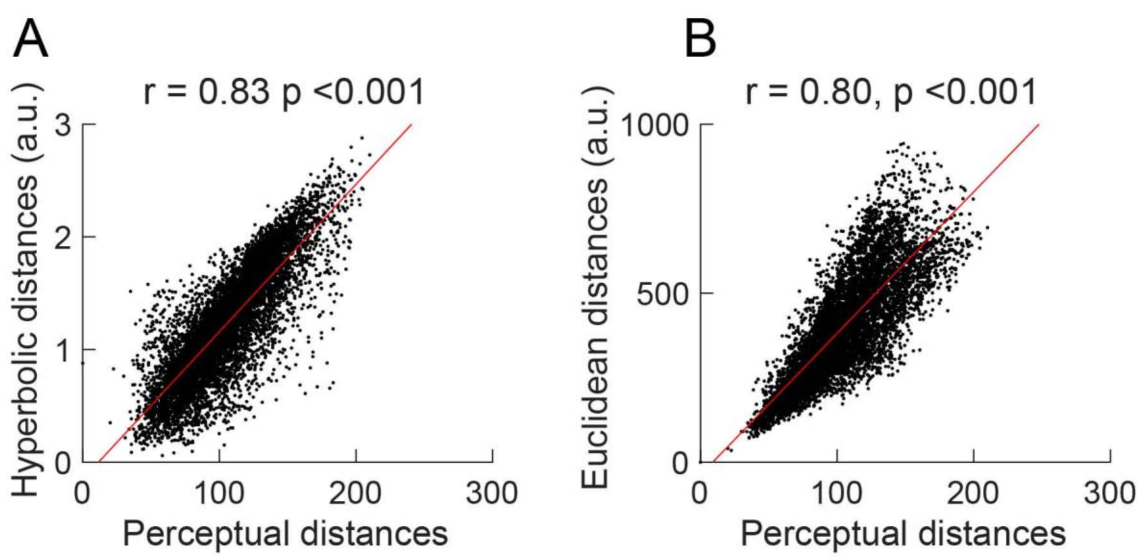


Fig. S6. Comparison between embedded geometric distances and reported perceptual distances. Nonmetric multidimensional scaling was used to embed the odors into 3D Hyperbolic space (**A**) or 3D Euclidean space (**B**).

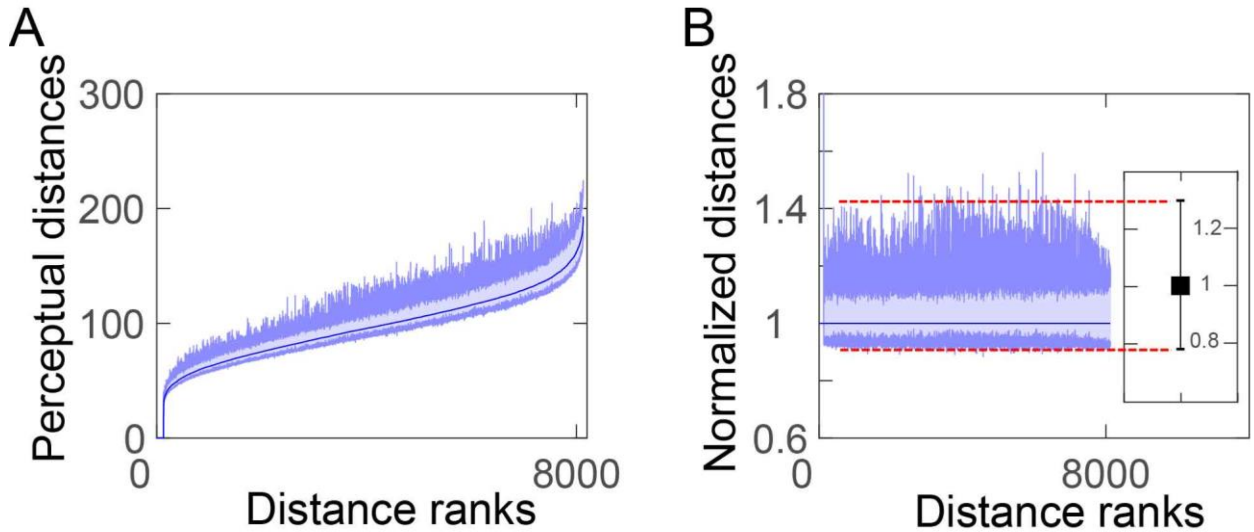


Fig. S7. Analysis of sensitivity of integrated Betti value to noise in the input distances. (**A**) The error bars show 95% confidence intervals for the distributions of Betti value computed based on 300 different partial samples of perceptual descriptors (120 out of 146 perceptual descriptors). Odor distances are rank ordered based on the medium distance across 300 samples. The blue line near the center shows the medium pairwise distances. (**B**) Pairwise distances normalized by their medium. The error bar plots show the 95% percent confidence intervals, blue line is the medium. Variability in normalized distances no longer depends on distance and matches the variability in the first integrated Betti value normalized by its medium (inset).

Supplementary Tables

Table S1. Statistical tests (*P* values) for consistency with hyperbolic models based on integrated Betti values.

p values		Dimension								
		3	4	5	6	8	10	20	50	100
Mouse urine	Betti 1	0.973	0.740	0.900	0.967	0.933	0.953	0.887	0.807	0.807
	Betti 2	0.453	0.853	0.980	1.000	0.920	0.773	0.540	0.573	0.593
	Betti 3	0.733	0.313	0.313	0.347	0.440	0.700	0.800	0.680	0.533
Strawberry	Betti 1	0.967	1.000	0.953	0.747	0.860	0.860	0.967	0.880	1.000
	Betti 2	0.893	0.613	0.713	0.420	1.000	0.813	0.707	0.647	0.680
	Betti 3	0.193	0.033	0.080	0.087	0.273	0.287	1.000	0.940	0.967
Blueberry	Betti 1	0.247	0.920	0.860	0.993	0.947	0.813	0.880	0.847	0.833
	Betti 2	0.260	0.173	0.167	0.127	0.080	0.093	0.040	0.027	0.013
	Betti 3	0.580	0.887	0.880	1.000	0.820	0.820	0.347	0.453	0.387
Tomato	Betti 1	0.747	0.973	1.000	0.880	0.913	0.840	0.867	0.947	1.000
	Betti 2	0.253	0.947	0.900	0.733	0.653	0.400	0.193	0.313	0.333
	Betti 3	0.207	1.000	0.873	0.633	0.533	0.187	0.007	0.033	0.000

Table S2. Statistical tests (*P* values) for consistency with hyperbolic models based on L1 distances between Betti curves.

p values		Dimension								
		3	4	5	6	8	10	20	50	100
Mouse urine	Betti 1	0.873	0.800	0.433	0.327	0.187	0.107	0.093	0.080	0.113
	Betti 2	0.347	0.147	0.173	0.413	0.967	0.613	0.287	0.180	0.147
	Betti 3	0.553	0.913	0.853	0.833	0.720	1.000	0.280	0.113	0.080
Strawberry	Betti 1	0.173	0.060	0.173	0.640	0.613	1.000	0.713	0.500	0.427
	Betti 2	0.287	0.767	1.000	0.413	0.327	0.233	0.060	0.067	0.013
	Betti 3	0.513	0.073	0.167	0.133	0.793	1.000	0.173	0.040	0.027
Blueberry	Betti 1	0.313	0.353	0.367	0.267	0.260	0.393	0.220	0.200	0.180
	Betti 2	0.967	0.947	1.000	0.700	0.587	0.493	0.327	0.213	0.153
	Betti 3	0.920	0.607	0.747	1.000	0.660	0.460	0.213	0.120	0.080
Tomato	Betti 1	0.373	0.713	0.373	0.300	0.067	0.060	0.007	0.007	0.007
	Betti 2	0.067	0.140	0.387	0.387	0.960	0.693	0.340	0.193	0.193
	Betti 3	0.220	0.340	0.607	0.733	0.673	0.807	0.493	0.453	0.460

Table S3. Statistical tests (*P* values) for evaluating consistency of experimental Betti curves with respect to 3D hyperbolic model or optimal optimal Euclidean model.

p values		Integrated Betti values		L1 differences	
		3D Hyperbolic	Optimal Euclidean	3D Hyperbolic	Optimal Euclidean
Mouse urine	Betti 1	0.973	0.907	0.873	0.000
	Betti 2	0.453	0.047	0.347	0.000
	Betti 3	0.733	0.000	0.553	0.000
Strawberry	Betti 1	0.967	0.893	0.173	0.000
	Betti 2	0.893	0.020	0.287	0.000
	Betti 3	0.193	0.000	0.513	0.000
Blueberry	Betti 1	0.247	0.993	0.313	0.060
	Betti 2	0.260	0.940	0.967	0.000
	Betti 3	0.580	0.027	0.920	0.000
Tomato	Betti 1	0.747	0.620	0.373	0.000
	Betti 2	0.253	0.333	0.067	0.000
	Betti 3	0.207	0.000	0.220	0.000

Table S4. Statistical tests (*P* values) for evaluating consistency of Betti curves computed based on logarithm of odor concentrations with respect to hyperbolic model. Consistency evaluated based on integrated Betti values

p values		Dimension								
		3	4	5	6	8	10	20	50	100
Mouse urine	Betti 1	0.947	0.693	0.840	1.000	1.000	0.993	1.000	0.867	0.840
	Betti 2	0.293	0.547	0.713	0.833	0.680	0.507	0.347	0.380	0.393
	Betti 3	0.960	0.473	0.453	0.427	0.627	0.853	0.547	0.513	0.360
Strawberry	Betti 1	0.707	0.740	0.773	0.553	0.867	0.633	0.807	0.780	0.740
	Betti 2	0.060	0.327	0.273	0.433	0.147	0.220	0.027	0.027	0.073
	Betti 3	0.680	0.640	0.713	0.633	0.767	0.653	0.107	0.047	0.080
Blueberry	Betti 1	0.713	0.307	0.287	0.293	0.233	0.287	0.200	0.367	0.300
	Betti 2	0.687	0.460	0.353	0.340	0.240	0.307	0.113	0.180	0.113
	Betti 3	0.607	0.453	0.420	0.333	0.233	0.167	0.087	0.087	0.040
Tomato	Betti 1	0.753	0.973	1.000	0.893	0.907	0.847	0.873	0.933	1.000
	Betti 2	0.207	0.793	0.753	0.547	0.540	0.320	0.120	0.193	0.213
	Betti 3	0.167	0.900	0.733	0.540	0.427	0.140	0.007	0.020	0.000

Table S5. *P* values of hyperbolic and Euclidean model using integrated Betti values for perceptual data set.

p values		Hyper 3D	Hyper 4D	Hyper 5D	Hyper 6D	Hyper 8D	Hyper 9D	Euc 2D	Euc 3D	Euc 4D
Odor atlas	Betti 1	0.920	0.860	0.893	0.987	0.767	0.680	0.000	0.120	0.000
	Betti 2	0.187	0.227	0.273	0.133	0.053	0.033	0.000	0.013	0.000