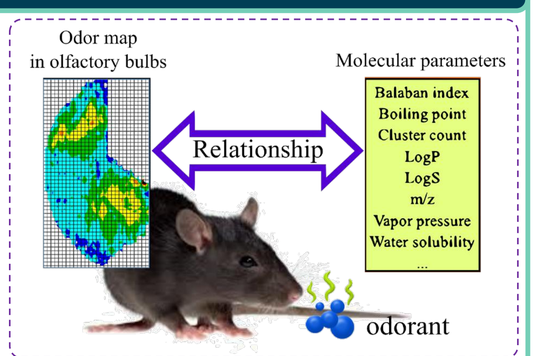
Statistical Clustering of Odorant Molecules Based on both Molecular Profile Feature Extraction and Olfactory Bulb odor Map Imaging Analysis

Liang Shang[1], Chuanjun Liu[1,2], Yoichi Tomiura[3], Kenshi Hayashi[1*]
[1]Department of Electronics, Graduate School of Information Science and Electrical Engineering, Kyushu University
[2]Research Laboratory, U.S.E. Co., Ltd., Tokyo
[3]Department of Informatics, Graduate School of Information Science and Electrical Engineering, Kyushu University

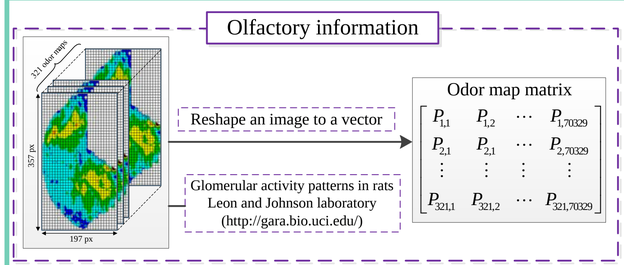
INTRODUCTION

- Investigations in molecular biology have demonstrated that the response pattern of odorants on olfactory bulb (odor map) are corresponding to their molecular structure.
- Detailed statistical analysis on both odor map and molecular parameter are carried out for an extensive understanding on the structure-odor relationship.
- Correlation coefficient analysis revealed that parameters were divided for 7 clusters, and each cluster showed relatively similar response pattern on olfactory bulb.
- T-distributed stochastic neighbor embedding (t-SNE) was employed for mapping odorants in 2D spaces by olfactory images and molecular parameters, respectively.
- Based on the features extracted by PCA or t-SNE, functional group identification models were calibrated by artificial neural network (ANN).
- The feasibility of odor maps and molecular parameters for odorant function group classification is discussed.

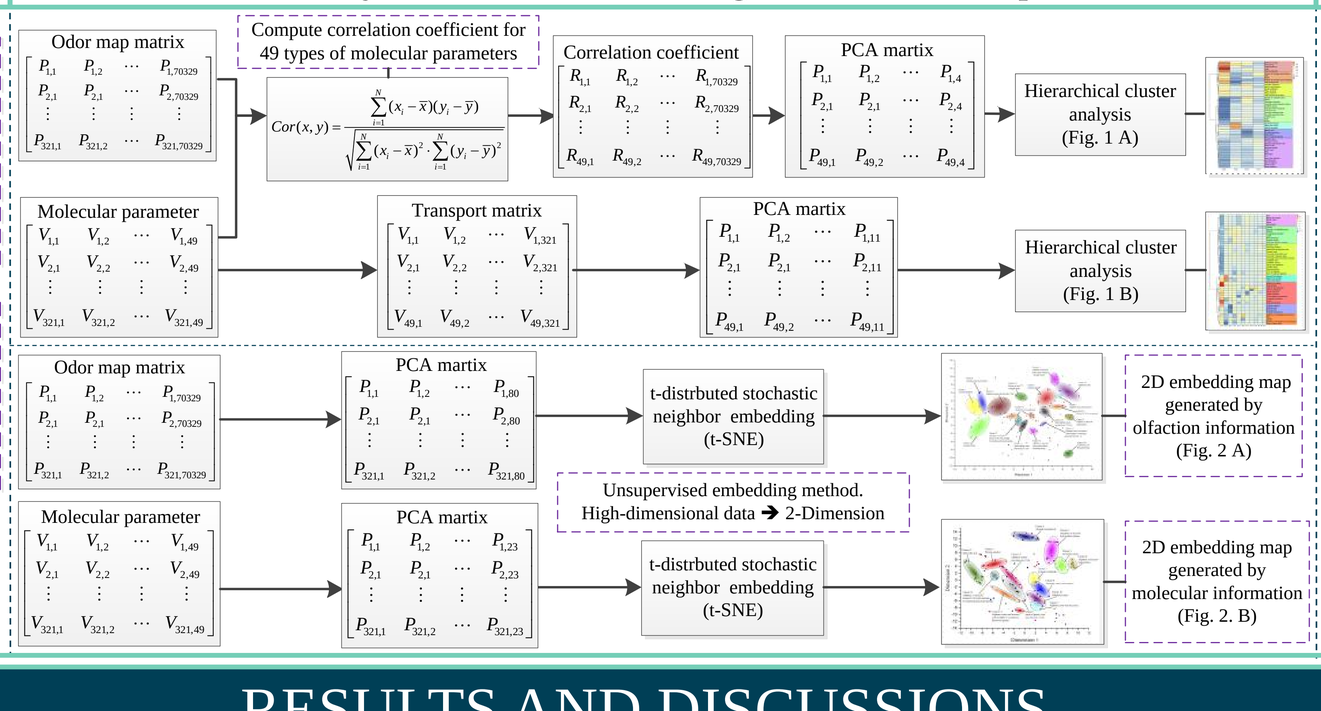


MATERIAL AND METHOD

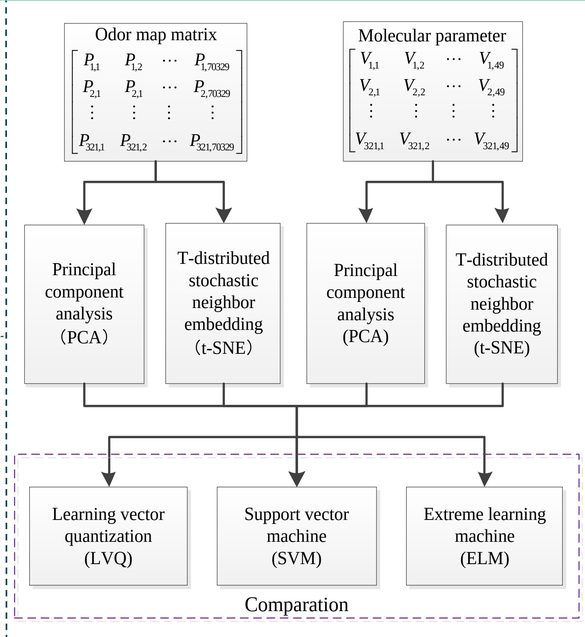
Data description



Statistical analyzation of olfaction images and molecular parameters



Discriminating model

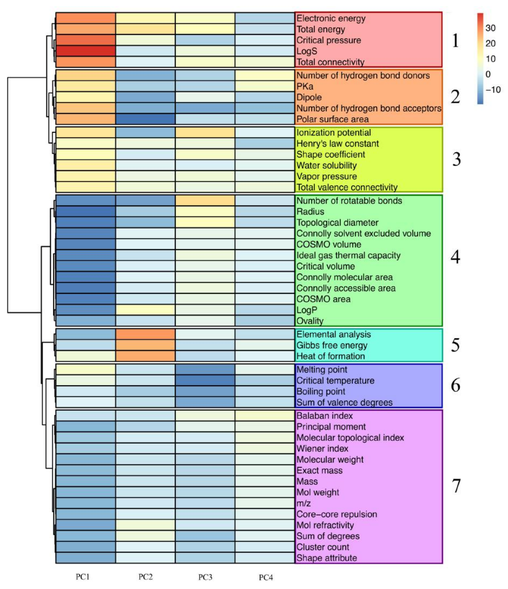


Reference

- Johnson, B. A. Spatial representations of odorants in olfactory bulbs of rats and mice: Similarities and differences in chemotopic organization. The Journal of comparative neurology 514, 658-673, (2009).

RESULTS AND DISCUSSIONS

A. Heat map of R-maps



B. Heat map of molecular parameters

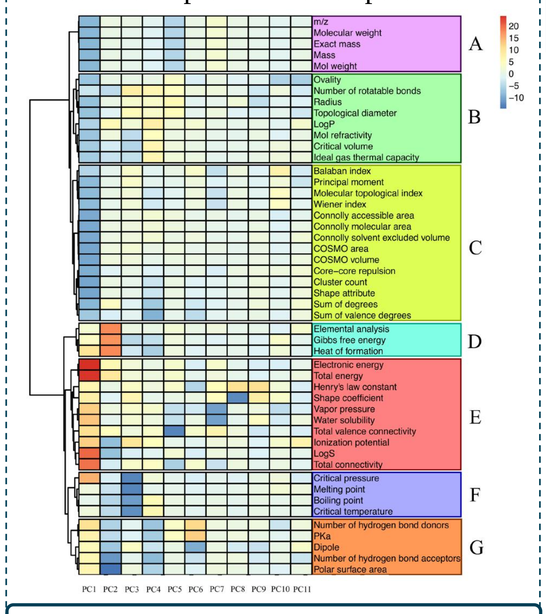
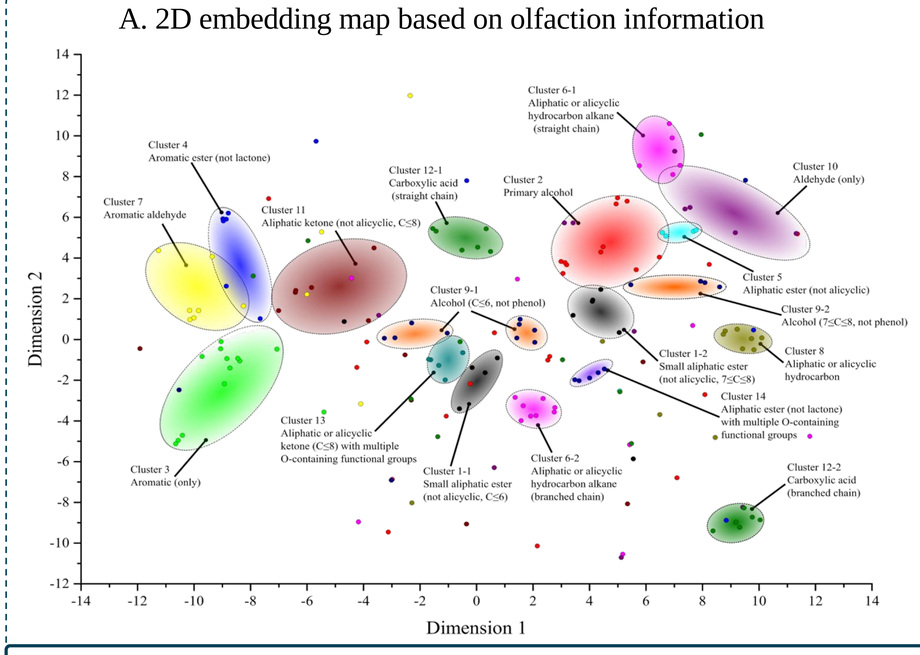


Fig 1. Heat map and hierarchical dendrogram

A. 2D embedding map based on olfaction information



B. 2D embedding map based on molecular information

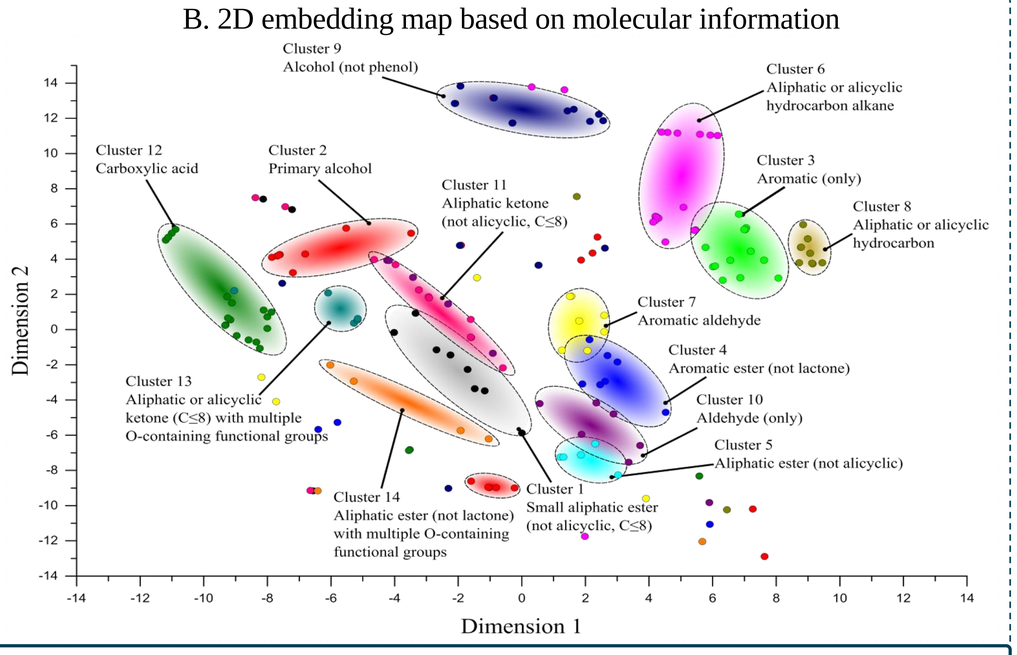


Fig. 2. 2D artificial odorant cluster map by t-SNE method

- In present study, samples from 14 varieties of functional groups were mapped in the t-SNE spaces generated by olfactory information and molecular information, respectively.
- In the olfactory space (A), odorants with the similar functional group are mostly (64.04%) clustered in one or two groups, respectively.
- The length of carbon chain would be vital for aliphatic esters (cluster 1) and alcohols (cluster 9), and the branched chains would play a role in hydrocarbon alkanes (cluster 6) and carboxylic acids (cluster 12) in rat`s olfaction .
- In the molecular space (B), odorants in the same cluster are mostly (68.00%) clustered together, and more overlappings were observed.
- It indicated that some information, such as the length of carbon chain, could not be included in molecular information map. However, some scatters are not clustered in a cluster.
- Hidden patterns would not been found because of the insufficiency of sample numbers.

Table 1. Identification accuracies of LVQ, SVM and ELM models.

Odorant descriptor	Pretreatment methods	Modeling approach	Accuracy (%)	
			Calibration set	Validation set
Olfactory image	PCA	LVQ	72.09	68.89
		SVM	29.63	18.60
		ELM	**94.81**	**93.02**
	t-SNE	LVQ	14.81	13.95
		SVM	14.81	13.95
		ELM	**72.59**	**90.70**
Molecular parameter	PCA	LVQ	65.93	69.77
		SVM	93.33	90.70
		ELM	**89.63**	**93.02**
	t-SNE	LVQ	21.48	23.26
		SVM	82.22	83.72
		ELM	**86.67**	**95.35**

- Cluster analysis was performed by Ward’s method on Euclidean distance.
- The results indicated that all the parameters were clustered in 7 clusters.
- The parameters in the same cluster described the similar information for odorants.
- Most of parameters contained energy information are in cluster 1, and parameters contained polar information are in cluster 2.
- Compared the heat map for R-maps, similar groups were observed. Just like cluster 2 and G, cluster 5 and D, cluster 6 and F.
- It indicated that the molecular parameters would be sensitive to olfactory information.

CONCLUSIONS

- 49 types of molecular parameters were clustered in 7 groups. Parameters in each cluster has a similar effect on olfactory images in rats.
- Some odorants are similar in molecular information space, but different in olfaction information space. However, most of odorants contained similar functional groups were clustered together.
- The results indicated that OI-PCA-ELM was the optimal model in distinguishing functional groups for odorants.
- An odorant would be described by molecular parameters to compare olfactory information.
- More molecular parameters would be considered to express some difficult features for odorant molecules.