

Biological Activity Predictions of Ligand based on Hybrid Molecular Fingerprint and Ensemble Learning

*Mengshan Li, Ming Zeng, Hang Zhang, Huijie Chen, Lixin Guan

College of Physics and Electronic Information, Gannan Normal University, Ganzhou, Jiangxi

341000, China,

List of contents

1. Source of data set	S2
2. The data processing	S2~S23

1. Source of data set

All the five datasets are derived from the supporting information of Reference 1¹. The website of the article is <https://pubs.acs.org/doi/full/10.1021/jm0497141>.

The activity values of each dataset are obtained from folder "jm0497141_si_001.pdf"; the molecular structure data file is obtained according to the file path "jm0497141_si_002\qsar_sets\3dqsar_mol2\3dqsar_mol2".

2. The data processing

On the basis of the original data, all the MOL2 files were converted into MACCS molecular fingerprints. The molecular fingerprint is shown in the following figure. Due to the length limitation of PDF output, the full length of the molecular fingerprint cannot be fully displayed.

2.1 ACE

Figure S1

Figure S2

Figure S3

2.2 AchE

Figure S5

Figure S6

Figure S7

2.3 BZR

Figure S8

Figure S9

Figure S10

Figure S11

Figure S12

2.4 COX2

Figure S13

Figure S14

Figure S15

Figure S16

Figure S17

Figure S18

Figure S19

Figure S20

Figure S21

Figure S22

Figure S23

2.5 DHFR

Figure S24

Figure S25

Figure S26

Figure S27

Figure S28

Figure S29

Figure S30

Figure S31

Figure S32

Figure S33

Figure S34

Figure S35

Figure S36

9-13m	6.49	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-13r	7.70	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-13s	7.77	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-13t	6.87	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-13u	6.11	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14a	6.92	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14b	7.54	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14c	6.80	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14h	8.40	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14k	6.64	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-14m	6.85	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-15a	6.64	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-15b	6.14	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-15c	6.09	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-16b	7.14	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-16c	7.64	0000000000000000000000000000000010000000000000000100010000000000011000000000001000
9-16d	6.34	0000000000000000000000000000000010000000000000000100010000000000011000000000001000

Figure S37

References

1. Sutherland, J. J.; O'Brien, L. A.; Weaver, D. F., A Comparison of Methods for Modeling Quantitative Structure-Activity Relationships. *Journal of Medicinal Chemistry* **2004**, *47* (22), 5541-5554.