

Supporting Information
Figures S1-S7
for
Scaffold and structural diversity of the secondary metabolite
space of medicinal fungi

R.P. Vivek-Ananth^{a,b}, Ajaya Kumar Sahoo^{a,b}, Shanmuga Priya Baskaran^{a,b}, Areejit Samal^{a,b,*}

^a *The Institute of Mathematical Sciences (IMSc), Chennai 600113, India*

^b *Homi Bhabha National Institute (HBNI), Mumbai 400094, India*

* To whom correspondence should be addressed: asamal@imsc.res.in

(a)

MeFSAT
Medicinal Fungi Secondary metabolites And Therapeutics

Home Basic search Advanced search Statistics Acknowledgement Download Help

ADVANCED SEARCH

Physicochemical filter Drug-like filter Chemical similarity filter Scaffold filter

Select Molecular scaffold type to filter secondary metabolites:

- Graph/Node/Bond level for a chemical has connectivity, element, and bond information
- Graph/Node level for a chemical has connectivity and element information but ignores bond information
- Graph level has connectivity information but ignores element and bond information

Graph/Node/Bond Graph/Node Graph

Choose scaffold

Choose from dropdown

Search

(b)

MeFSAT
Medicinal Fungi Secondary metabolites And Therapeutics

Home Basic search Advanced search Statistics Acknowledgement Download Help

Secondary metabolite: Antcin C

Summary Physicochemical properties Drug-likeness properties ADMET properties Descriptors

Summary

Molecular formula: C₂₉H₄₂O₅

SMILES: CC(C)=C(CC)C@H]1[C@@H]1CC[C@@H]2[C@@]1(C)CC(=O)C1=C2[C@@H](O)C[C@@H]2[C@@]1(C)CCC(=O)[C@@H]2C(C)=O

InChI: InChI=1S/C29H42O5/e1-15(17(3)27(33)34)7-8-16(2)19-9-10-20-25-23(31)13-21-18(4)22(30)11-12-28(2)5)26(25)24(3)14-29(19,20)6/h16-21,23,31H,1,7-14H2,2-6H3, (H,33,34)/t16-17,18+19-20+21+23+28+29-m/1s1

InChIKey: TWISSXUVVGIUBP-FNSKOMQISA-N

Molecular scaffolds

Scaffold Graph/Node/Bond level:
O=C1CCC2C3=C(C)CC2C1)C1CCCC1CC3=O

Scaffold Graph/Node level:
OC1CCC2C(C)CC3C4CCCC4CC(O)C23)C1

Scaffold Graph level:
CC1CCC2C(C)CC3C4CCCC4CC(C)C23)C1

Figure S1: Screenshots of the (a) Scaffold filter tab under the Advanced Search option in the updated MeFSAT database to filter secondary metabolites by selecting scaffolds of interest, and (b) the detailed information page for a secondary metabolite in the updated MeFSAT database displaying the identified scaffolds for the secondary metabolite.

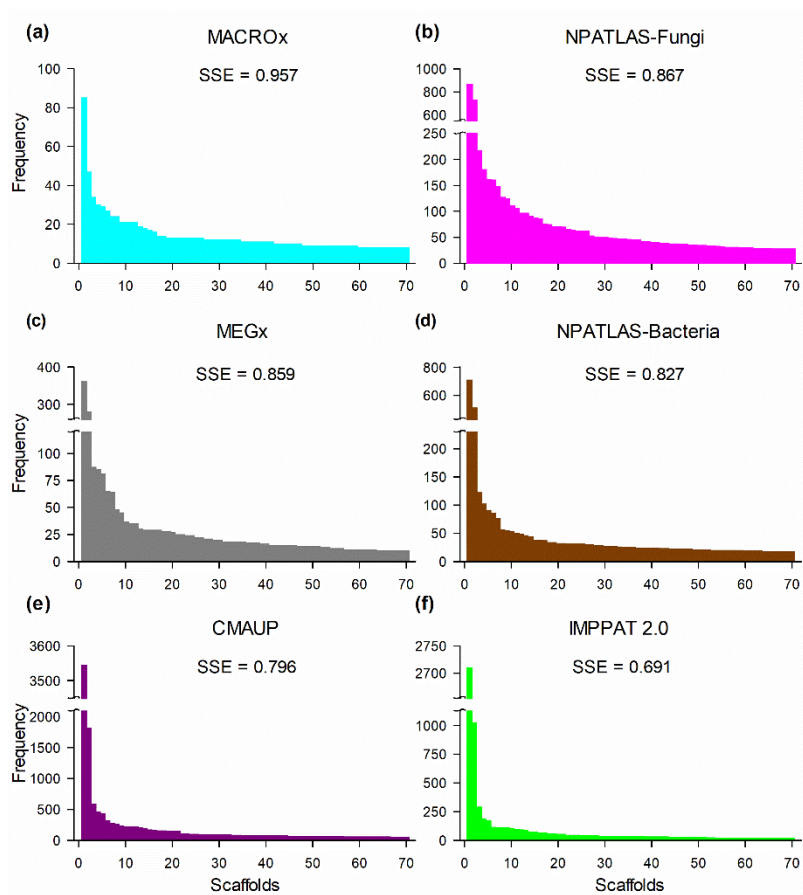


Figure S2: Distribution of chemicals across the top 70 most populated scaffolds in libraries: (a) MACROx, (b) NPATLAS-Fungi, (c) MEGx, (d) NPATLAS-Bacteria, (e) CMAUP, and (f) IMPPAT 2.0.

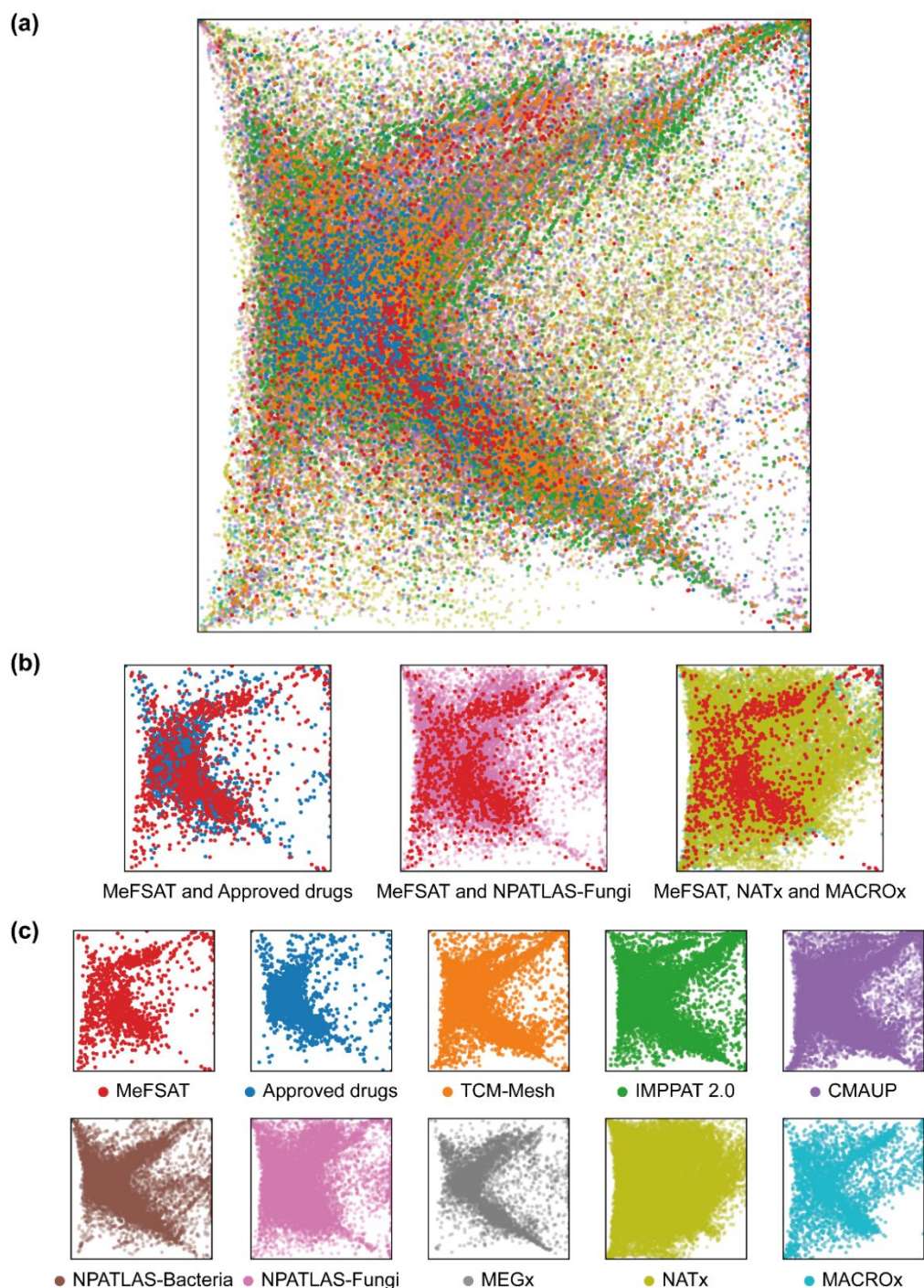


Figure S3: Visualization of the chemical spaces generated via GTM using molecular properties for the libraries analysed here. (a) Visualization of all chemical libraries analysed here. (b) Visualization of MeFSAT and Approved drugs, MeFSAT and NPATLAS-Fungi, and MeFSAT, NATx and MACROx. (c) Visualization of each individual chemical library. The colour used to represent each chemical library in the visualization is provided in part (c) along with the corresponding library name.

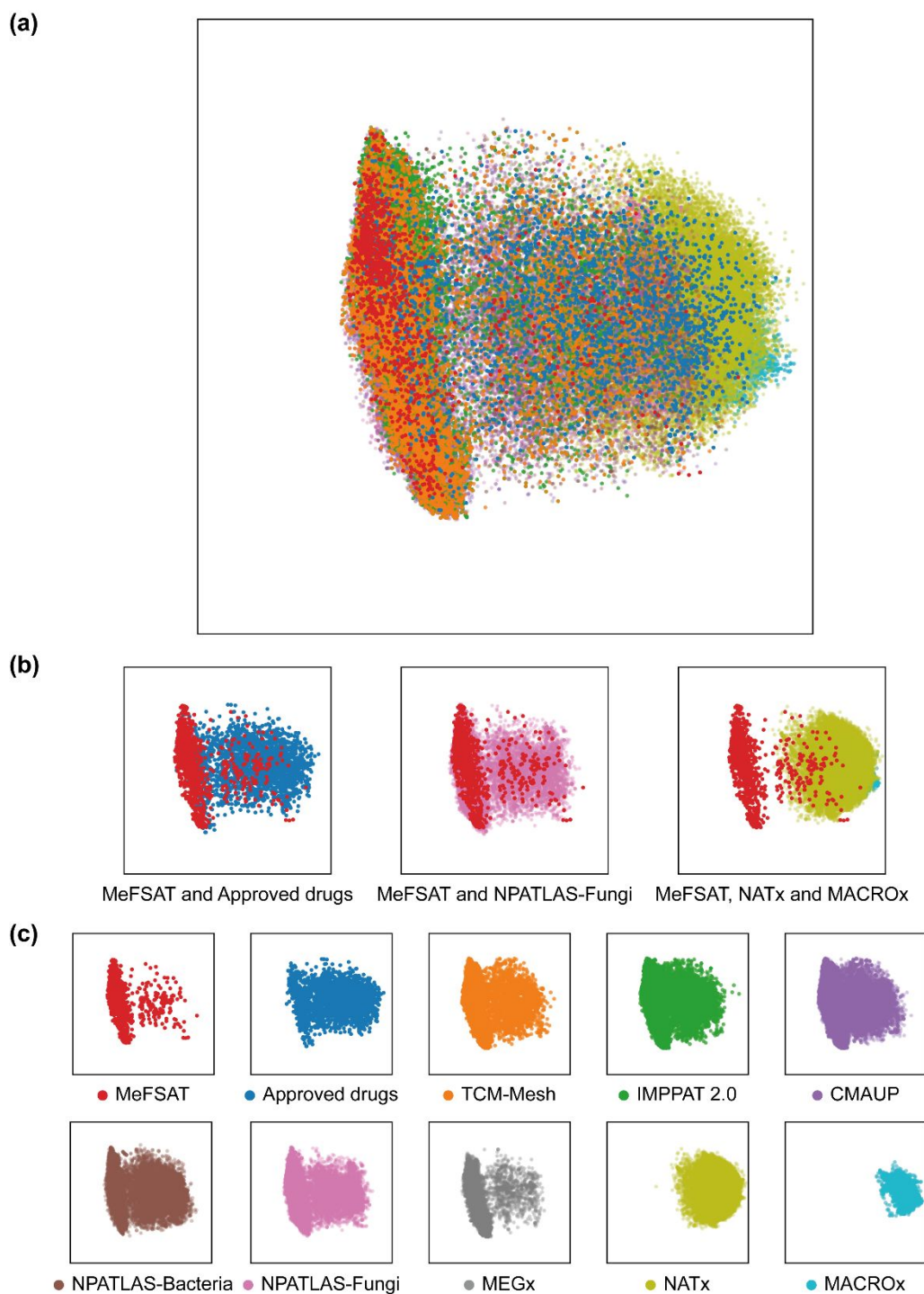


Figure S4: Visualization of the chemical spaces generated via PCA using MACCS keys structural fingerprints for the libraries analysed here. (a) Visualization of all chemical libraries analysed here. (b) Visualization of MeFSAT and Approved drugs, MeFSAT and NPATLAS-Fungi, and MeFSAT, NATx and MACROx. (c) Visualization of each individual chemical library. The colour used to represent each chemical library in the visualization is provided in part (c) along with the corresponding library name.

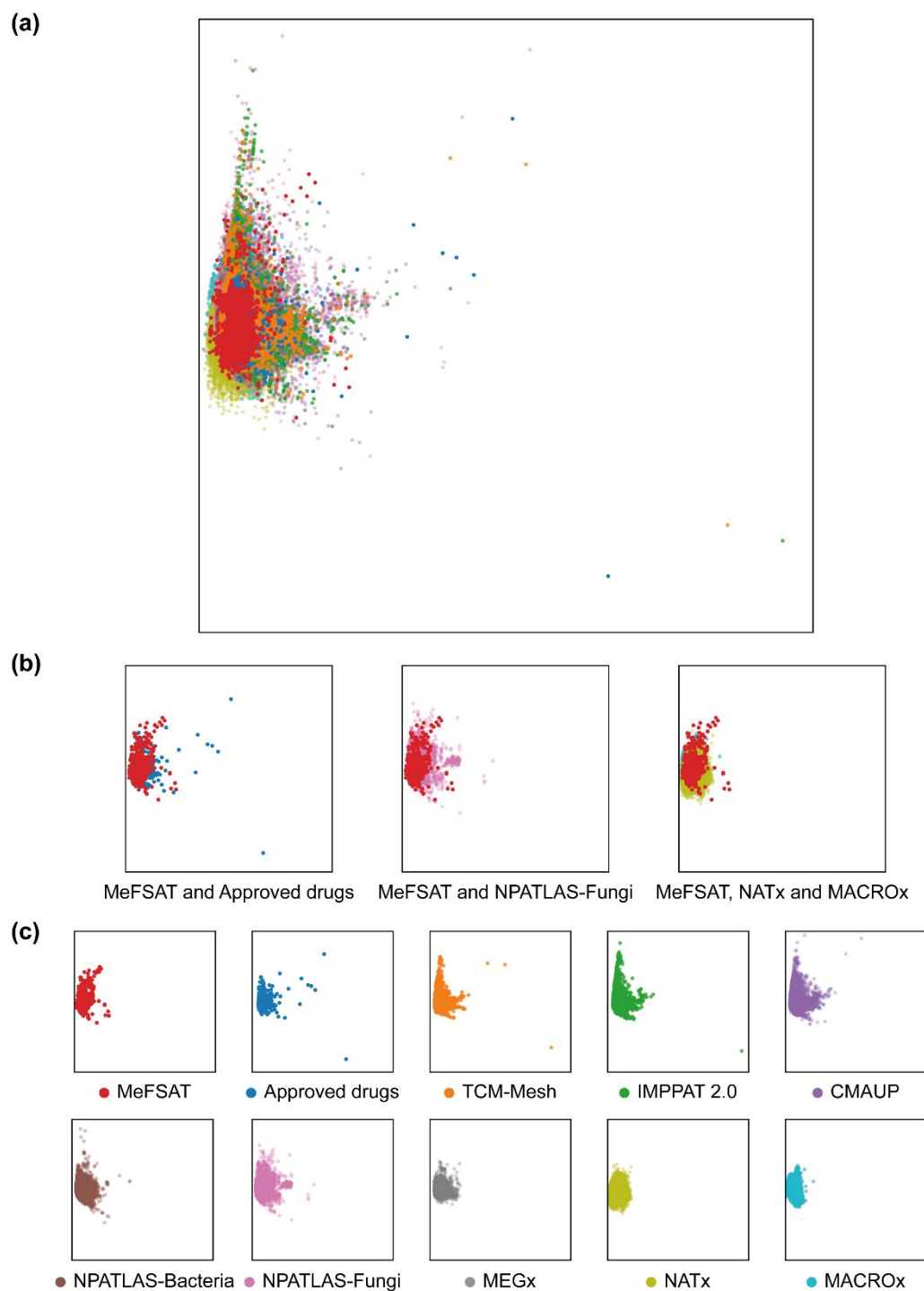


Figure S5: Visualization of the chemical spaces generated via PCA using molecular properties for the libraries analysed here. (a) Visualization of all chemical libraries analysed here. (b) Visualization of MeFSAT and Approved drugs, MeFSAT and NPATLAS-Fungi, and MeFSAT, NATx and MACROx. (c) Visualization of each individual chemical library. The colour used to represent each chemical library in the visualization is provided in part (c) along with the corresponding library name.

MeFSAT	NA									
Approved drugs	8	NA								
TCM-Mesh	82	181	NA							
IMPPAT 2.0	80	175	3423	NA						
CMAUP	130	248	5945	6703	NA					
NPATLAS-Bacteria	5	19	50	112	184	NA				
NPATLAS-Fungi	1202	10	102	121	451	2	NA			
MEGx	26	43	499	703	1058	59	199	NA		
NATx	0	0	0	0	0	0	0	0	NA	
MACROx	0	0	0	0	0	0	0	0	0	NA

MeFSAT **Approved drugs** **TCM-Mesh** **IMPPAT 2.0** **CMAUP** **NPATLAS-Bacteria** **NPATLAS-Fungi** **MEGx** **NATx** **MACROx**

Figure S6: Compound overlap between the chemical libraries analysed in this study. The triangular matrix provides the number of compounds common between different pairs of chemical libraries.

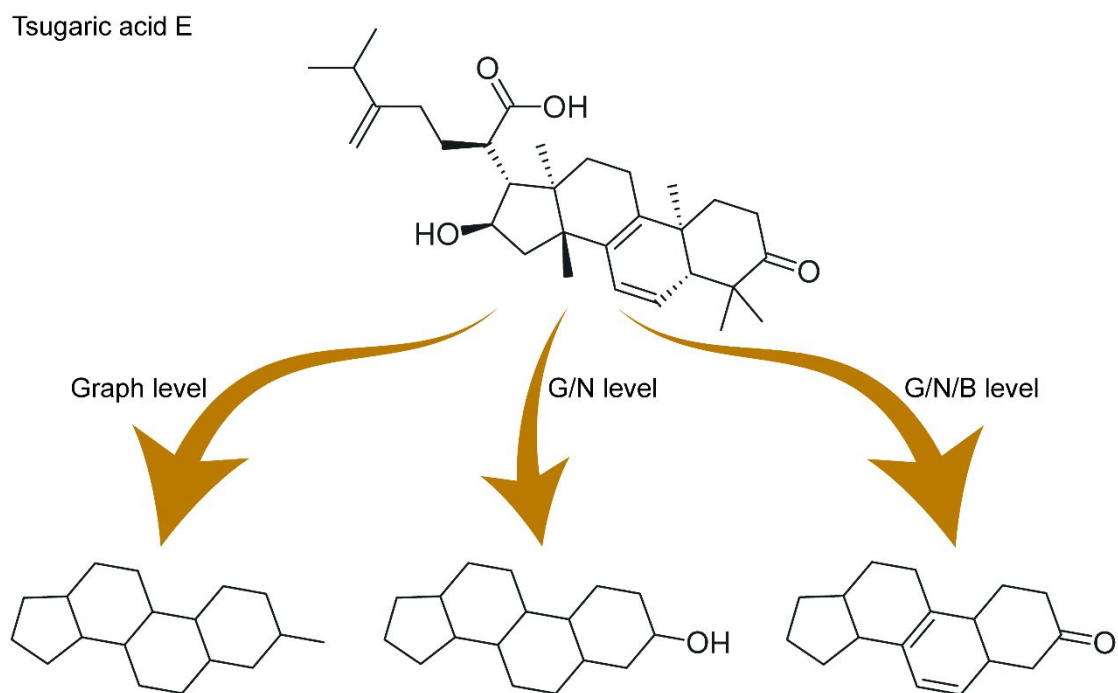


Figure S7: Molecular scaffold of the secondary metabolite Tsugaric acid E at three different levels: graph/node/bond (G/N/B) level (with connectivity, element and bond information), graph/node (G/N) level (with both connectivity and element information) and graph level (with only connectivity information).