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: <u>asamal@imsc.res.in</u>

MC	alcinati angi seci		apeutics			
Home	Basic search	Advanced search	Statistics	Acknowledgement	Download	
Dhusisashan	vieal filter Drug	like filter Chemical similar	ADVANCED S	SEARCH		
Thysicochen	inter brog	g une inter	Scallo			
Select Molecul • Graph/Node/Bo • Graph/Node lev • Graph level has	ar scaffold type to f nd level for a chemial ha el for a chemical has con connectivity information	ilter secondary metabolites: s connectivity, element, and bond informa nectivity and element information but ign but ignores element and bond informatio	ation lores bond information in			
Graph/Node/Bor	d Graph/Node	Graph				
Choose scaffol	l d	~				
Search						
50 St.						
-						
R N	NeFSAT					
R N		ondary metabolitas And The	rapeutice			
Re N	NeFSAT adicinal Fungi Sec	ondary metabolites And The	rapeutics			
Home	NeFSAT edicinal Fungi Sec Basic search	ondary metabolites And The Advanced search	rapeutics Statistics	Acknowledgement	Download	н
Home	NeFSAT adicinal Fungi Sec Basic search	ondary metabolites And The Advanced search	rapeutics Statistics	Acknowledgement	Download	н
Home	NeFSAT edicinal Fungi Sec Basic search	ondary metabolites And The Advanced search Sec	rapeutics Statistics ondary metabo	Acknowledgement	Download	н
Home Summary	NeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p	rapeutics Statistics ondary metabo roperties ADME	Acknowledgement blite: Antcin C T properties Descriptors	Download	н
Home Summary	NeFSAT Rdicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p	rapeutics Statistics ondary metabo roperties ADME	Acknowledgement Dite: Antcin C T properties Descriptors	Download	н
Home	NeFSAT edicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p	rapeutics Statistics ondary metabo roperties ADME	Acknowledgement blite: Antcin C T properties Descriptors Summary	Download	н
Home Summary	NeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C299 Data De conferences	rapeutics Statistics ondary metabo roperties ADME	Acknowledgement Ditte: Antcin C T properties Descriptors Summary	Download	н
Ken Me Home	NeFSAT edicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C29H SMILES: CC(C=OCCC) SMILES: CC(C=OCCC)	rapeutics Statistics ondary metabo roperties ADME 4205 HIJI(C@HJJCCC@@H	Acknowledgement blite: Antcin C T properties Descriptors Summary 4/21(©)11(C)C(==0)(1=C2(C@))	Download	(C)(
Ken Me Home	AeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C29H SMILES: CC[C=C]CC[C] [C@+1]2C]C<[C=0]	rapeutics Statistics ondary metabo roperties ADME 4205 sHijlC@HjlCCCC@#H	Acknowledgement Diffe: Antcin C T properties Descriptors Summary 1/21C(8)11(C)CCI=0)C1=C2[C8/	Download	(C)(
Home Summary	AeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p SMLES: C7(C1=07C100) InCh: InCh: IS/C29H42 2021 InCh: InCh: IS/C29H42 2021 InCh: InCh: IS/C29H42 2021 InCh: InCh: IS/C29H42	rapeutics Statistics ondary metabor roperties ADME 14205 pH](CGH]1CC(CGH 05k11-15(72)27(33) 05k1-15(72)27(33)	Acknowledgement Ditte: Antcin C T properties Descriptors Summary 4/2(C@)14(C)=C)C1=C2(C@) 194(7)=0.6(2)19-0.00-20-3 194(7)=0.100-20-50-3 194(7)=0.100-20-50-3	Download	H (C)0
Kong Kong Kong Kong Kong Kong Kong Kong	AeFSAT Edicinal Fungi Sec Basic search Physicochemic	andary metabolites And The Advanced search Sec at properties Drug-likeness p SMILES: CC[(=C)CC[6] SMILES: CC[(=C)C	rapeutics Statistics ondary metabol 4205 HI(IC@H]1CC[C@H 05k1-15[[7]3]27[33 -29(19.20)6h16-27.81-3 -20:41-2016h16-27.81-3	Acknowledgement blite: Antcin C T properties Descriptors Summary 4/2(C@)14(C)CC(=-0(C1=C2/C@)(134)7-8-16(2)19-9-10-20-25-7 333141.7-4442.2-643. 9-mb/b1	Download 9H1(O)C[C(\$)(9H2)(C(\$)) (3(31)13-21-18(4)22(30)	H ((C)(
Home Summary	AeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C29H SMILES: CC(C=C)CC(C (CM=1/2C)C(C=0)0 InCR: Inc.Ini=15/C29H4/ 28(21,5)26(25)24(22)14 (1+33.3+)1(6:-17/18+1 InCR/Kerz TV/SSQU/WOL	rapeutics Statistics ondary metabo voperties ADME 44205 H-II(C@H)ICC(C@H) 05k1-15(17(3)27(33) -29(19.20)6/h16-21; 9-20-21+.23+.28-20 SulPer-PNsKOWIGSA	Acknowledgement blite: Antcin C T properties Descriptors Summary 4/2(C6)1(C)CC(=0)C1=C2/C6/4 (34/7-8-16/2(19-9-10-20-25-2 331)H.17-14H2.2-6H3, 9-m1b1 N	Download	H (C)(
Konger Horne Summary	AeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Motecular formula: C29H SMILES: CC[C=C]CC[C] [CC=H]2C]C[C=00] [ICC=H]2C]C[C=00] [ICC=H][C]S[25]24(25]24(25]24(25)24(2	rapeutics Statistics ondary metabo (4205 HI[(C@H]1CC(C@#) 05k1-15(17(3)27(33 -29(19.20(6)16-21.3) 9-20+21+23+28+2 SIUBP-FNSKOMQISA	Acknowledgement Diffe: Antcin C T properties Descriptors Summary 4/21(G):14(5):Cc1=0;C1=C2[C)(124(7):8-16(2):19-9:05-20-25-2 123.11+1.7-1442.2-643, 19-m1/s1 -N	Download	H ((C)(
Konger Home Summary	AeFSAT adicinal Fungi Sec Basic search Physicochemic	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C29H SMILES: C2(C)=C2CC(C) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) [C@H]2C]C(C)=O) Molecular formula: C29H SMILES: C2(C)=C2CC(C) [C@H]2C]C(C)=O) Molecular formula: C29H SMILES: C2(C)=C2CC(C) [C@H]2C]C(C)=O) Molecular formula: C29H SMILES: C2(C)=C2CC(C) [C@H]2C]C(C)=O) Inchikey: TWISSNUW/O	rapeutics Statistics ondary metabo roperties ADME 44205 HI[(C@H]1CC(C@H 05k1-15[17(3)27(33 9-20+21+23+28+2 9-20+21+23+28+2 SIUBP-FNSKOMQISA	Acknowledgement Dilite: Antcin C T properties Descriptors Summary -1/21(G):14(5)CCI=0/C1-C2[C](4) 1/4(7-8-16/2)19-9-10-20-25-2 1/3.11+.17-14/12.2-6H3, 19-m1/s.1 Molecular scaffolds	Download	(C)(
Konsentration	AeFSAT adicinal Fungi Sec Basic search Physicochemic	Advanced search Advanced search Sec al properties Drug-likeness p Molecular formula: C239 Molecular fo	rapeutics Statistics ondary metabor reperties ADME 44205 9H][[C@H]1CC[C@#] 005k1-15[73]27(33) -29(19.20)6h16-21; 9-20-212-23-28-3; GUUP-FNSKOMQISA nd levet;	Acknowledgement Uitte: Antcin C T properties Descriptors Summary 42(C(8)11(C)CC(=0)C1-C2(C(8) 13(4)7-8-16(2)19-9-10-20-25-2 13(3)1(-1)7-14(-12-6H3) 9-m0.51 N Molecular scaffolds	Download	(C)
Kenter Ke	AeFSAT edicinal Fungi Sec Basic search Physicochemic Charles Content Physicochemic	andary metabolites And The Advanced search Sec at properties Drug-likeness p SMILES: CC(C+C)CC(6 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)524(21)4 (C)@412C(C)C(-C) ICCNE: InCNI=15/C29142 24(21,5)2(42)524(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)54(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)54(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)54(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)54(21)4 (C)@412C(C)C29142 24(21,5)2(42)54(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 24(21,5)2(42)54(21)4 (C)@412C(C)CC) ICCNE: InCNI=15/C29142 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)@412C(C)CC) ICCNE: InCNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (C)CNI=15/C2914 (rapeutics Statistics ondary metabol reperties ADME 44205 HH[(C@H]1CC[C@H 056:1-15[(7/3)27(33 -29(19.20(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:30(6n16-21:1) -30:2-12:12:12:10(6n16-21:1) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7) -30:2-12:12:10(7)	Acknowledgement blite: Antcin C T properties Descriptors Summary 4/2(C@)1(C)CC(=-0)C1=C2/C@)((34/7-8-16(2)19-9-10-20-25-2 333H1.7-7.444.2-643, (9-m1/a) Molecular scaffolds	Download	(C)
Konne Horne Summary	AeFSAT adicinal Fungi Sec Basic search Physicochemic $\downarrow \downarrow \downarrow \downarrow$ $\downarrow \downarrow \downarrow$ $\downarrow \downarrow$	ondary metabolites And The Advanced search Sec al properties Drug-likeness p Molecular formula: C29H SWILES: CC(C=C)CC(E) (CM+12C)C(C=0) InCk: Inch-15/C29H4 (H33.34)H:6-172.18+1 InChiKey: TWISSNUW/C0 Scaffold Graph/Node/Boo OcclCCC23+cC(CCC2) Scaffold Graph/Node/Sec	rapeutics Statistics ondary metabo roperties ADME 44205 HH[(C@H]1CC[C@@H 05k1-15[17(3)27(33 -29(19.20)6h-16-21:3 9-20+21+23+28+2 GIUBP-FNSKOMQISA nd levet: Licclcccc1cc3=0 et	Acknowledgement blite: Antcin C T properties Descriptors Summary 4/2(CG)14(C)CC(=O)C1=C2/CG)4 (34)7-8-16(2)19-9-10-20-25-2 2331H1,7-1442,2-643, 9-m1/21 	Download	H (C)

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.



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).