Supporting Information

Chemical Diversity of Metabolites from Fungi, Cyanobacteria, and Plants Relative to FDA-Approved Anticancer Agents

Tamam El-Elimat, Xiaoli Zhang, David Jarjoura, Franklin J. Moy, Jimmy Orjala, A. Douglas Kinghorn, Cedric J. Pearce, and Nicholas H. Oberlies

Contents:

Experimental Procedures

Table S1. Molecular Descriptors Utilized in the Current Study Compared to Related PCA Studies in the Literature.

Table S2. Pearson Correlation Coefficients for Raw Data.

Table S3. Pearson Correlation Coefficients for Molecular Weight Standardized Data.

Table S4. Summary Statistics of Different Properties among Secondary Metabolites from Filamentous Fungi (n = 105), Cyanobacteria (n = 75), Tropical Plants (n = 163) and Anticancer Drugs (n = 96).

Figure S1. Plot of the first and fourth principal components of the isolated secondary metabolites from filamentous fungi (n = 105), cyanobacteria (n = 75), tropical plants (n = 163) and anticancer drugs (n = 96).

Figure S1. Chemical structures of the anticancer drugs that were not overlapping in the chemical space with the investigated compounds in the PCA plots in Figures 1, 2, and S1.

EXPERIMENTAL PROCEDURES

The natural products utilized in this study were isolated in pursuit of anticancer leads. The recent literature may be examined for examples of descriptions of the isolation and structure elucidation of compounds from filamentous fungi, cyanobacteria, and tropical plants by our collaborative team. The FDA-approved anticancer agents were selected from two complementary sources: those posted on the website of the Developmental Therapeutic Program NCI/NIH and the recently approved anticancer drugs from the orange book on the FDA website. All molecular descriptors, except number of chiral centers, were calculated using Dragon software for molecular descriptor calculation (version 6.0 – 2010; Talete srl, Milano, Italy). The number of chiral centers was tabulated by inspection. SAS was used for principal component analysis varimax rotation of the 9 variable correlation matrix (Proc Princomp and Factor). Components score were obtained after variance standardization.

Table S1. Molecular Descriptors Utilized in the Current Study Compared to Related PCA Studies in the Literature [14-17]

Molecular descriptors	Feher and Schmidt[14]	Singh et al.[16]	Tan[15]	Singh and Culberson[17]	Current study
molecular weight (MW)	---	×	×	×	×
number of chiral centers (nCC)	×	---	×	×	×
number of rotatable bonds (NRBN)	×	×	×	×	×
number of solvated acceptor atoms for H-bonds (N,O,F) [N_{Acc},solv]	×	×	×	×	×
number of solvated donor atoms for H-bonds (N and O) [N_{Don},solv]	×	×	×	---	×
topological polar surface area [TPSA(NO)]	---	×	×	×	×
octanol/water partition coefficient	---	×	×	×	×
number of nitrogen atoms (nN)	---	---	×	×	×
number of oxygen atoms (nO)	---	---	×	×	×
number of C-N bonds (nCN)	×	---	---	---	---
number of C-O bonds (nCO)	×	---	---	---	---
number of C-halogen bonds (nCX)	×	---	---	---	---
number of C-S bonds (nCS)	×	---	---	---	---
ratio of aromatic atoms to ring atoms	×	---	---	---	---
ring fusion degree	×	---	---	---	---
normalized bond flexibility	---	---	---	×	---

Table S2. Pearson Correlation Coefficients for Raw Data

Correlation coefficient	MW	nRBN	nN	nO	nH_{Don}	nH_{Acc}	TPSA	MLOGP	nCC
MW	1.0000	0.8713	0.8185	0.8878	0.9001	0.9466	0.9459	-0.6667	0.6500
nRBN	0.8713	1.0000	0.8174	0.7533	0.8493	0.8681	0.8702	-0.6292	0.4024
nN	0.8185	0.8174	1.0000	0.6271	0.8768	0.8777	0.8795	-0.7712	0.3450
nO	0.8878	0.7533	0.6271	1.0000	0.8061	0.9202	0.9097	-0.7564	0.6632
nH_{Don}	0.9001	0.8493	0.8768	0.8061	1.0000	0.9258	0.9607	-0.7850	0.4993
nH_{Acc}	0.9466	0.8681	0.8777	0.9202	0.9258	1.0000	0.9906	-0.8427	0.5682
TPSA	0.9459	0.8702	0.8795	0.9097	0.9607	0.9906	1.0000	-0.8373	0.5591
MLOGP	-0.6667	-0.6292	-0.7712	-0.7564	-0.7850	-0.8427	-0.8373	1.0000	-0.3916
nCC	0.6500	0.4024	0.3450	0.6632	0.4993	0.5682	0.5591	-0.3916	1.0000

Table S3. Pearson Correlation Coefficients for Molecular Weight Standardized Data

Correlation coefficient	MW	nRBN	nN	nO	nH_{Don}	nH_{Acc}	TPSA	MLOGP	nCC
MW	1.0000	0.2423	0.1755	0.0879	0.2571	0.2449	0.2201	-0.4706	0.1445
nRBN	0.2423	1.0000	-0.0071	0.0249	0.0669	0.0604	0.0296	0.0694	-0.2509
nN	0.1755	-0.0071	1.0000	-0.3057	0.4915	0.4680	0.4842	-0.3778	-0.2184
nO	0.0879	0.0249	-0.3057	1.0000	0.2136	0.6711	0.6170	-0.5665	0.0529
nH_{Don}	0.2571	0.0669	0.4915	0.2136	1.0000	0.5692	0.7763	-0.5713	-0.1417
nH_{Acc}	0.2449	0.0604	0.4680	0.6711	0.5692	1.0000	0.9172	-0.8048	-0.1255
TPSA	0.2201	0.0296	0.4842	0.6170	0.7763	0.9172	1.0000	-0.7766	-0.1471
MLOGP	-0.4706	0.0694	-0.3778	-0.5665	-0.5713	-0.8048	-0.7766	1.0000	-0.0515
nCC	0.1445	-0.2509	-0.2184	0.0529	-0.1417	-0.1255	-0.1471	-0.0515	1.0000

Table S4. Summary statistics of different properties among secondary metabolites from filamentous fungi (n = 105), cyanobacteria (n = 75), tropical plants (n = 163) and anticancer drugs (n = 96)

source	Variables	Raw data		After MW standardization	
		Mean	Std. Dev.	Mean	Std. Dev.
Anticancer Drugs (n = 96)	MW	412.76	248.04	1.000	0.000
	RBN	5.31	5.12	0.012	0.008
	nN	3.02	2.79	0.009	0.008
	nO	4.92	4.73	0.011	0.007
	nHDon	3.09	3.46	0.008	0.008
	nHAcc	8.00	5.69	0.020	0.009
	TPSA	113.88	89.62	0.295	0.164
	MLOGP	1.12	2.24	0.003	0.007
	CC	3.30	4.87	0.006	0.007
Cyanobacteria (n = 75)	MW	628.78	338.39	1.000	0.000
	RBN	7.77	7.35	0.011	0.007
	nN	3.79	4.44	0.005	0.004
	nO	6.01	6.10	0.007	0.005
	nHDon	5.12	5.27	0.007	0.004
	nHAcc	9.12	10.49	0.011	0.008
	TPSA	150.41	152.65	0.193	0.106
	MLOGP	2.30	3.55	0.006	0.006
	CC	6.28	4.31	0.010	0.004
Filamentous Fungi (n = 105)	MW	660.36	555.30	1.000	0.000
	RBN	12.14	17.60	0.014	0.012
	nN	4.19	7.72	0.003	0.004
	nO	8.98	6.87	0.014	0.004
	nHDon	5.74	8.11	0.007	0.004
	nHAcc	13.17	14.29	0.018	0.004
	TPSA	196.95	218.11	0.263	0.073
	MLOGP	0.53	3.65	0.004	0.005
	CC	5.69	3.92	0.010	0.006
Tropical Plants (n = 163)	MW	450.42	152.31	1.000	0.000
	RBN	5.90	4.25	0.014	0.011
	nN	0.06	0.27	0.000	0.001
	nO	6.57	3.68	0.014	0.005
	nHDon	2.37	2.10	0.005	0.003
	nHAcc	6.60	3.65	0.014	0.005
	TPSA	100.87	52.82	0.220	0.064
	MLOGP	2.70	1.84	0.007	0.004
	CC	4.58	5.46	0.009	0.008

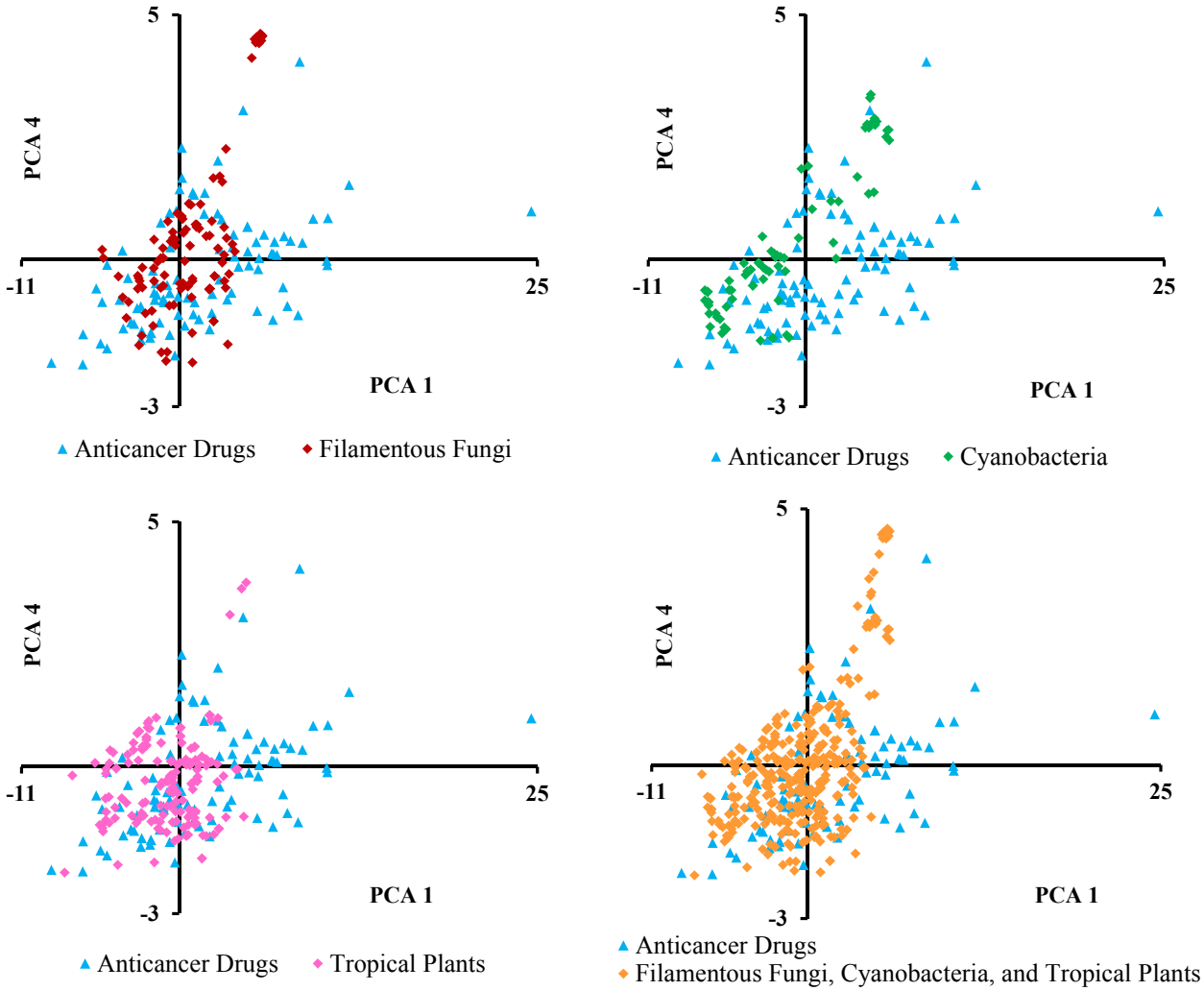

Figure S1. Plot of the first and fourth principal components of the isolated secondary metabolites from filamentous fungi (n = 105), cyanobacteria (n = 75), tropical plants (n = 163) and anticancer drugs (n = 96).

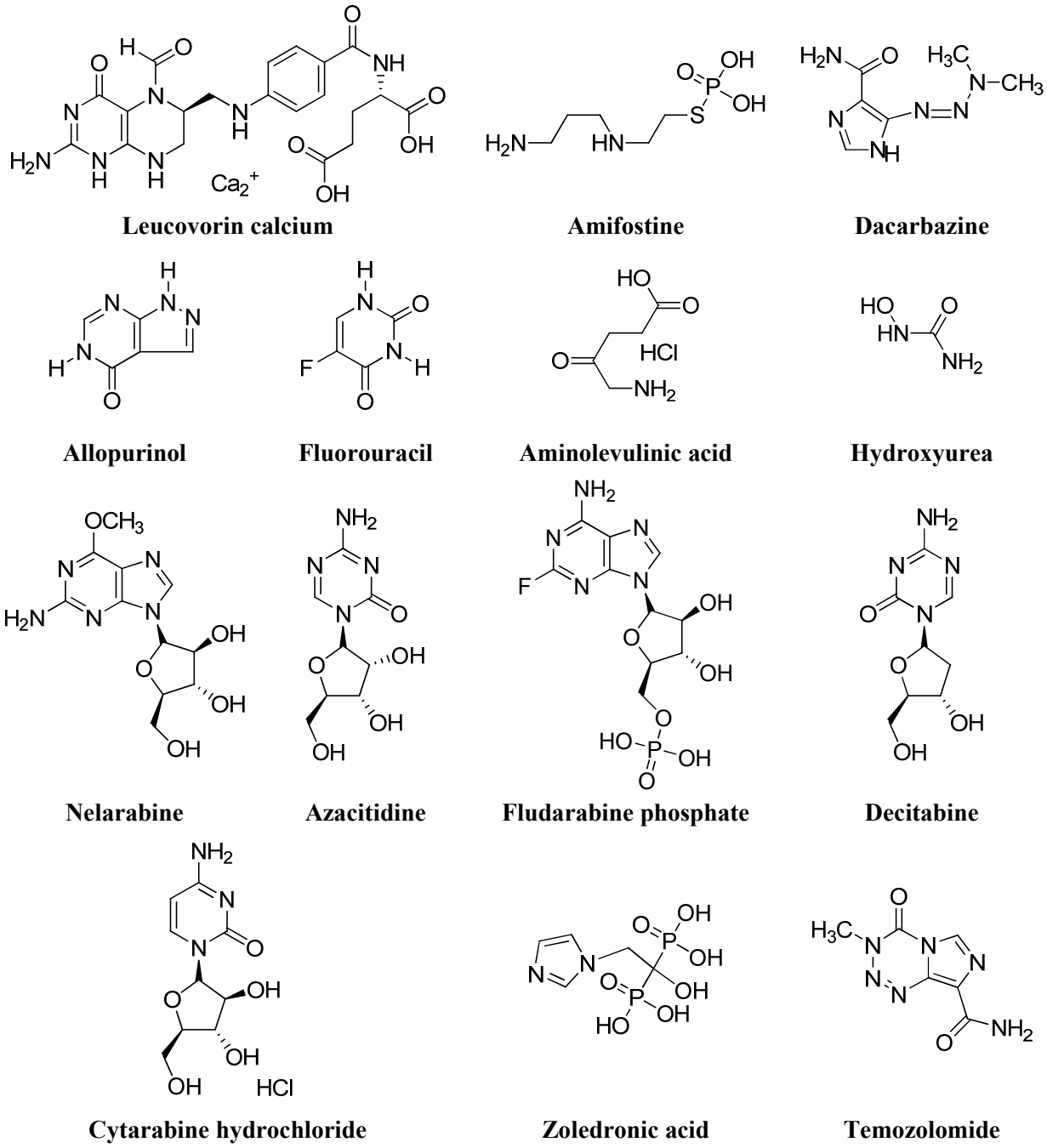

Figure S2. Chemical structures of the anticancer drugs that were not overlapping in the chemical space with the investigated compounds in the PCA plots in Figures 1, 2, and S1.