SUPPORTING INFORMATION

Epigenetic relevant chemical space: A chemoinformatic characterization of inhibitors of DNA methyltransferases

Eli Fernández-de Gortari and José L. Medina-Franco[*]

Contents

HBA

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	4.0×10^{-14}	-	-	-
Epigenetic	4.6×10^{-14}	1.1×10^{-6}	-	-
Inhibitors	6.6×10^{-14}	0.65	3.1×10^{-8}	-
DNMT	$< 2 \times 10^{-16}$	$< 2 \times 10^{-16}$	0.35	$< 2 \times 10^{-16}$

HBD

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	6.2×10^{-5}	-	-	-
Epigenetic	4.7×10^{-14}	1.7×10^{-12}	-	-
Inhibitors	4.7×10^{-7}	0.91	2.6×10^{-11}	-
DNMT	$< 2 \times 10^{-16}$	$< 2 \times 10^{-16}$	0.79	$< 2 \times 10^{-16}$

RB

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	3.9×10^{-14}	-	-	-
Epigenetic	0.06611	0.63338	-	-
Inhibitors	0.00061	1.8×10^{-6}	0.88426	-
DNMT	6.9×10^{-8}	0.35236	0.99556	0.09890

SlogP

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	3.7×10^{-14}	-	-	-
Epigenetic	3.9×10^{-14}	6.7×10^{-6}	-	-
Inhibitors	$<2 \times 10^{-16}$	8.5×10^{-5}	0.0233	-
DNMT	4.0×10^{-14}	0.1470	0.0033	0.6462

TPSA

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	6.5×10^{-14}	-	-	-
Epigenetic	0.00459	1.00000	-	-
Inhibitors	0.00050	0.00126	0.39842	-
DNMT	$<2 \times 10^{-16}$	7.7×10^{-14}	0.00041	1.2×10^{-14}

MW

	Approved	Clinical	Epigenetic	Inhibitors
Clinical	$<2 \times 10^{-16}$	-	-	-
Epigenetic	1.5×10^{-5}	0.57842	-	-
Inhibitors	4.6×10^{-14}	0.00047	0.99736	-
DNMT	4.3×10^{-14}	0.39607	0.98740	0.50709

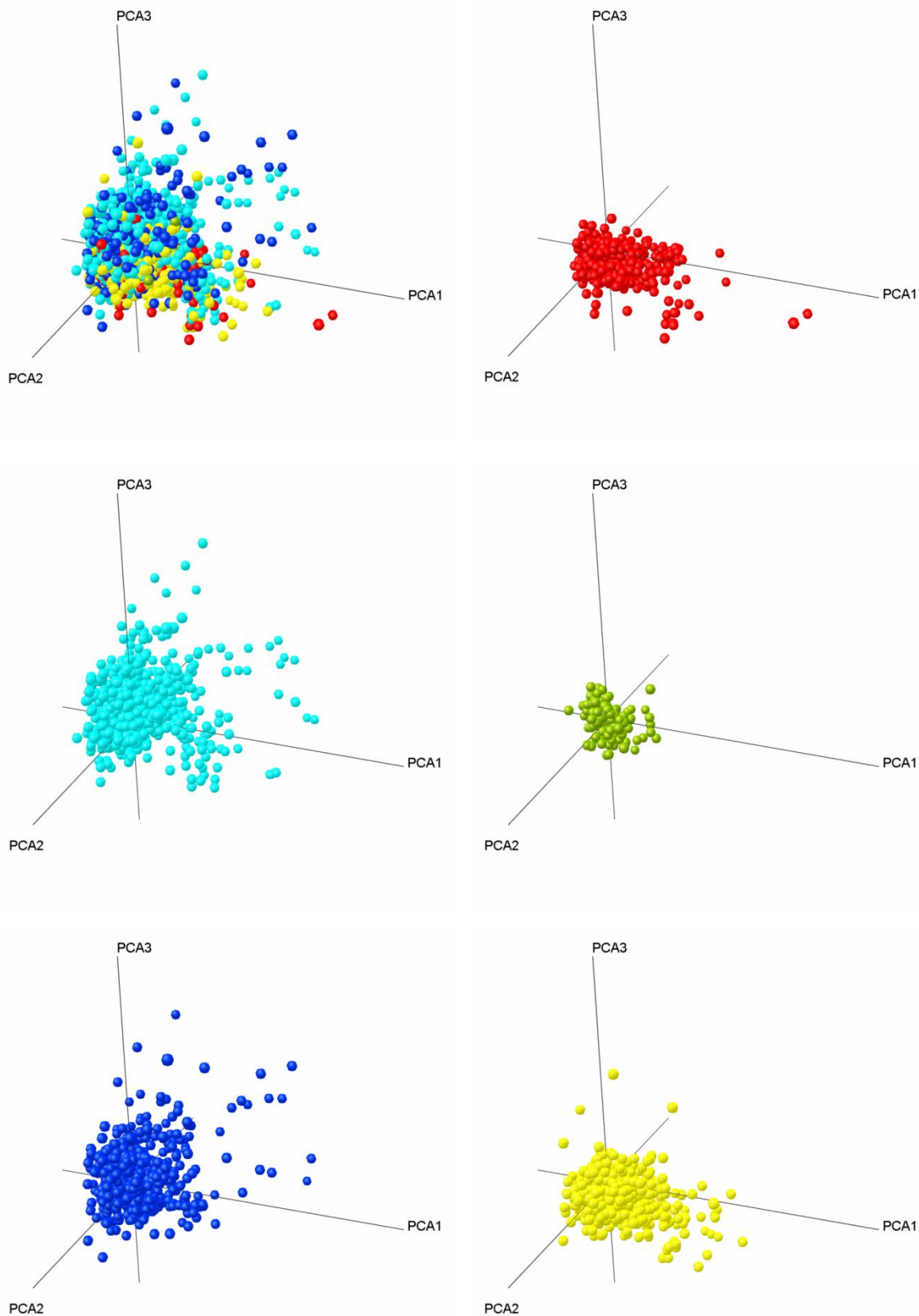**Figure S1.** Kruskal-Wallis analysis and a post hoc Nemenyi test.

Figure S2. 3D representation of the property space of the compounds data sets. DNMT1 inhibitors (red); approved drugs (cyan); focused on epigenetic targets (green); clinical (blue); general screening (yellow). All plots are in the same coordinate system.
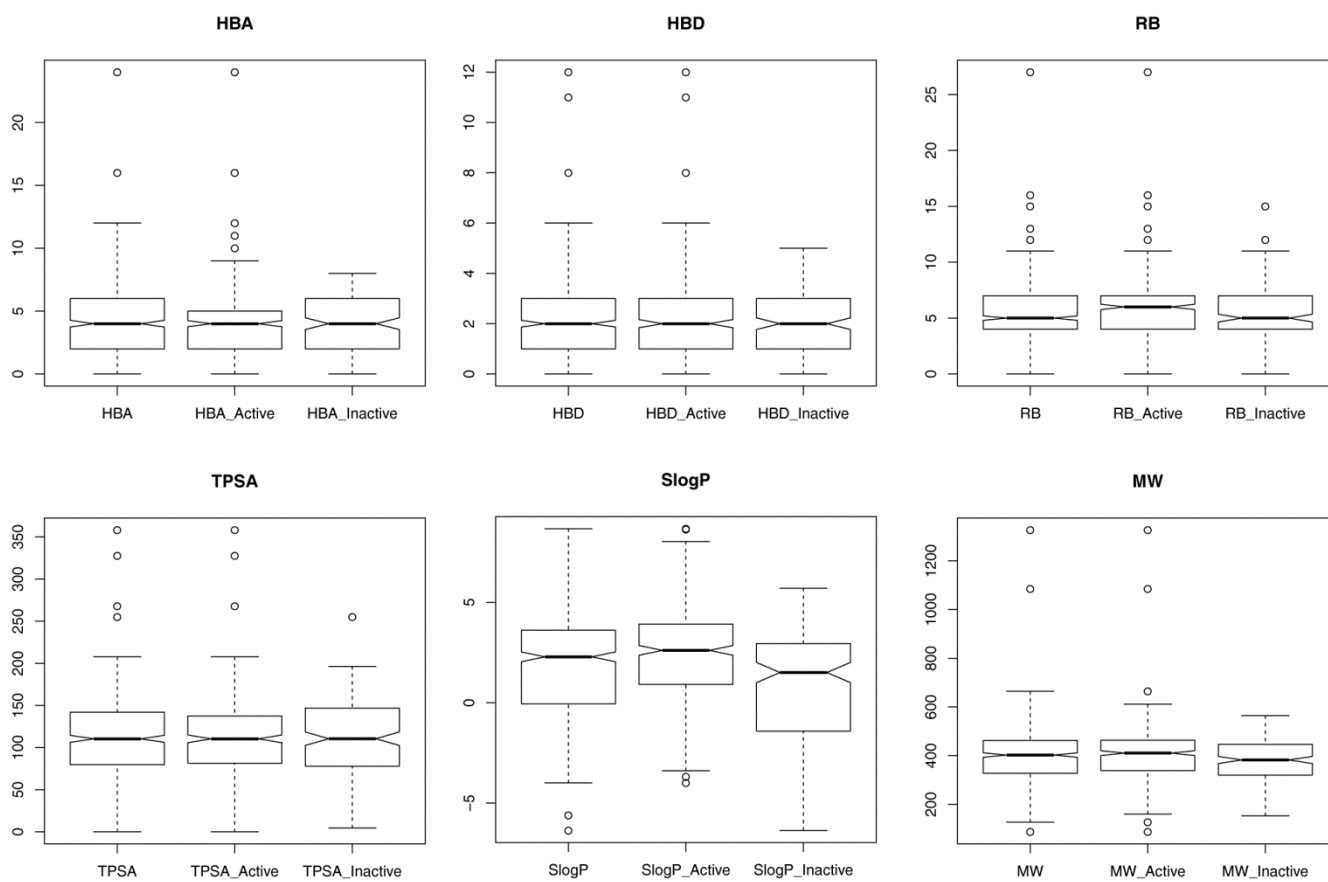
Figure S3. Comparison of the physicochemical properties of all compounds in the DNMT1 data set (left) *vs*. active (IC$_{50}$ < 10 µM) (middle) and inactive (right) compounds in the same set.

TGD	TGD_Epi	TGD_App	TGD_Cli	TGD_DNMT	TGD_Inhi
Min	0.2070	0.0000	0.0000	0.0910	0.2100
1Q	0.5910	0.4540	0.5230	0.5610	0.5430
Median	0.6680	0.5570	0.6190	0.6360	0.6290
Mean	0.6441	0.5371	0.5949	0.6290	0.6155
3Q	0.7190	0.6430	0.6960	0.7000	0.6980
Max	0.9960	0.9580	0.9220	1.0000	0.9540
SD	0.1131	0.1530	0.1443	0.1208	0.1159

MACCS	MACCS_Epi	MACCS_App	MACCS_Cli	MACCS_DNMT	MACCS_Inhi
Min	0.0770	0.0000	0.0000	0.0000	0.0380
1Q	0.3110	0.2200	0.2790	0.3100	0.2820
Median	0.3940	0.3040	0.3730	0.3980	0.3650
Mean	0.3969	0.3107	0.3686	0.4144	0.3688
3Q	0.4790	0.3970	0.4610	0.5060	0.4490
Max	0.9140	0.8750	0.9780	1.0000	0.9480
SD	0.1250	0.1301	0.1322	0.1602	0.1213

GpiDAPH3	Gpi_Epi	Gpi_App	Gpi_Cli	Gpi_DNMT	Gpi_Inhi
Min	0.0000	0.0000	0.0000	0.0000	0.0000
1Q	0.1940	0.0000	0.0000	0.1660	0.0000
Median	0.2760	0.1350	0.2230	0.2440	0.2390
Mean	0.2557	0.1374	0.1987	0.2463	0.2128
3Q	0.3460	0.2530	0.3010	0.3220	0.3210
Max	0.9470	1.0000	1.0000	1.0000	0.7110
SD	0.1366	0.1470	0.1392	0.1762	0.1443

ECFP4	EC4_App	EC4_Cli	EC4_DNMT	EC4_Epi	EC4_Inhi
Min	0.0000	0.0000	0.0000	0.0300	0.0000
1Q	0.0400	0.0400	0.0500	0.0600	0.0500
Median	0.0600	0.0500	0.0600	0.0700	0.0600
Mean	0.0585	0.0555	0.0729	0.0796	0.0618
3Q	0.0700	0.0700	0.0700	0.0800	0.0700
Max	0.6100	0.2100	0.7600	0.4300	0.4500
SD	0.0282	0.0212	0.0628	0.0713	0.0250

Figure S4. Summary statistics of the similarity values computed with TGD, MACCS, GpiDAPH3 and ECFP4 fingerprints for the five compound databases in this study: approved drugs ('App'), compounds in clinical trials ('Clin'), screening compounds focused on epigenetic targets ('Epi'), general screening collection ('Inhi'), and inhibitors of DNMT1 ('DNMT').
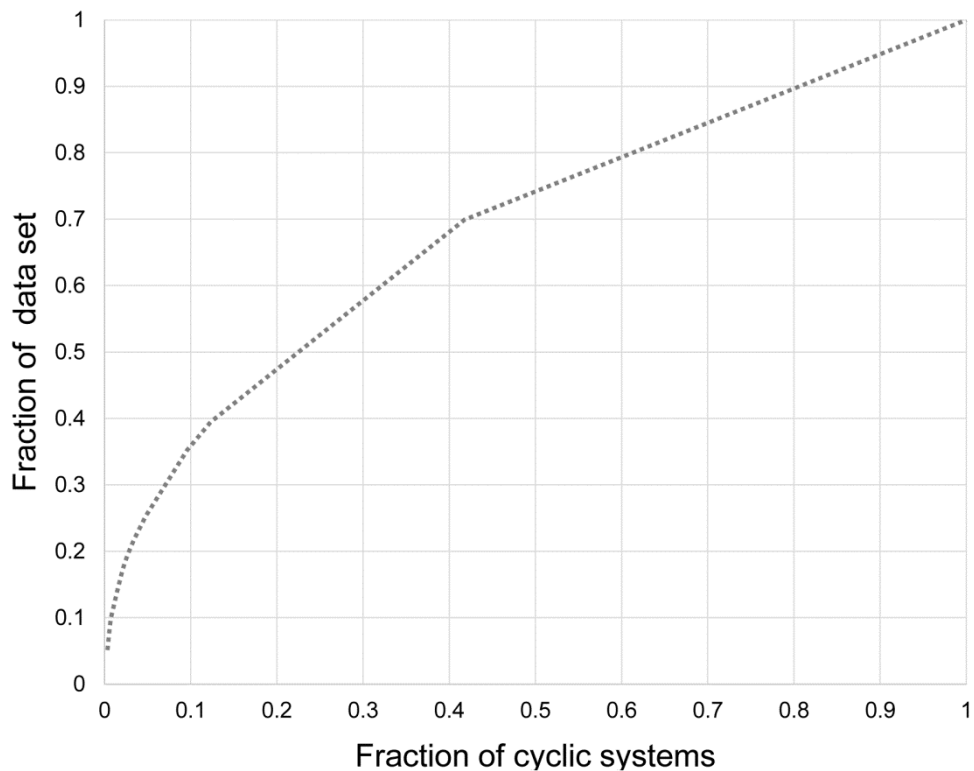
Figure S5. Cyclic System Recovery curve for the compounds tested as inhibitors of DNMT1. The Area Under the Curve (AUC) is 0.689 and the F_{50} value is 0.23.