Supplementary Materials: Pharmaceutical Machine Learning: Virtual High-Throughput Screens Identifying Promising and Economical Small Molecule Inhibitors of Complement Factor C1s.

Jonathan Jun Feng Chen 1, tyndsey N. Schmucker 2, and Donald P. Visco Jr 3, +

Table S1. Training Data. All data used to train the first-round classification and regression vHTS model. The same data is augmented with the *in vitro* experimental results and used to train the second-round classification and regression vHTS models. SID= PubChem Substance ID

Structure	SID	Classification	Activity $IC_{50}[\mu M]$
S N H CI	3711963	Inactive	N/A
H ₃ C N	3712600	Inactive	N/A

N S N N N N N N N N N N N N N N N N N N	3712681	Inactive	N/A
HN CH ₃	3712850	Inactive	N/A
H ₂ C NH NH NH	3713318	Inactive	N/A

H ₃ C N S HN S CH ₃	3713502	Inactive	N/A
H ₁ C OH ₁	3714494	Inactive	N/A
N N CH ₃	3715220	Inactive	N/A

H ₃ C ON CH ₃	3715428	Inactive	N/A
SH ₃ C N CH ₃	3715753	Inactive	N/A
CH ₃	3715870	Inactive	N/A

CH ₃	3716178	Inactive	N/A
N S II O HN S	3716847	Inactive	N/A
CH ₃	3717180	Inactive	N/A

CH ₃	3717586	Active	2.14
H ₃ C N H ₃ C N H ₃ C N H ₃ C	4240573	Inactive	N/A
H ₃ C CH ₃	4241692	Inactive	N/A

H ₃ C CH ₃	4242011	Inactive	N/A
H ₃ C H _N S S	4242319	Inactive	N/A
CH ₃ CH ₃ CH ₃ CH ₃	4243864	Inactive	N/A

H ₃ C O O O O O O O O O O O O O O O O O O O	4244211	Inactive	N/A
CH ₃	4245215	Inactive	N/A
CH ₃ O CH ₃ O N N N N N N N N N N N N N N N N N N N	4246549	Inactive	N/A

H ₃ C N S CH ₃	4246823	Inactive	N/A
H N N N CH ₃	4246912	Inactive	N/A
H ₃ C - 0	4247384	Inactive	N/A

N CH ₃	4248033	Inactive	N/A
HN CH3	4248495	Inactive	N/A
CH ₃	4249007	Active	34.72

S CH ₃	4249302	Inactive	N/A
F N H N S S S O CH₃ CH₃	4250394	Inactive	N/A
CH ₃ CH ₃ CH ₃ O CH ₃ O CH ₃	4251421	Inactive	N/A

O H N CH ₃	4251548	Inactive	N/A
H ₃ C N N N N CI	4251738	Inactive	N/A
H N CH ₃	4254582	Inactive	N/A

H ₃ C CH ₃	4254762	Inactive	N/A
O CH ₃ CH ₃	4254843	Inactive	N/A
O CI	4255208	Inactive	N/A

N O O CH ₃	4255209	Inactive	N/A
H ₃ C	4255516	Active	30.99
H_3C H_3C N	4255537	Inactive	N/A

H ₃ C O O O O O O O O O O O O O O O O O O O	4255573	Inactive	N/A
CH ₃	4255800	Inactive	N/A
H ₃ C O O O O O O O O O O O O O O O O O O O	4256088	Inactive	N/A

O-CH ₃	4256347	Inactive	N/A
H ₃ C N N N CH ₃ CH ₃	4256651	Inactive	N/A
S	4256969	Inactive	N/A

NH NH NH	4257420	Inactive	N/A
S S HN	4257713	Inactive	N/A
H N N CH ₃	4257844	Inactive	N/A

HN O CH ₃	4258135	Inactive	N/A
S N CH ₃	4258729	Inactive	N/A
N N N CH ₃	4258821	Inactive	N/A

ON NH CH ₃	4258988	Active	0.38
H ₃ C N NH O CH ₃	4259517	Inactive	N/A
CH ₃	4259584	Inactive	N/A

S N N	4259744	Active	14.91
S N N	4260000	Active	3.83
H ₃ C O CH ₃	4260130	Active	10.1

N S HN O HN O O	4261230	Inactive	N/A
	4261270	Inactive	N/A
	4261321	Active	37.87

H ₃ C O NH CH ₃	4261362	Inactive	N/A
H ₃ C O O O O O O O O O O O O O O O O O O O	4261432	Inactive	N/A
H ₃ C N N N CH ₃	4261564	Inactive	N/A

F F F	4261661	Inactive	N/A
H ₃ C CH ₃	4261818	Inactive	N/A
H ₃ C SH	4261899	Inactive	N/A

	4262775	Inactive	N/A
H ₃ C CH ₃	4263151	Active	5.35
N N N CH ₃	4263449	Active	5.59

H ₃ C N H ₃ C	4263981	Inactive	N/A
CH ₃	4264329	Inactive	N/A
O N H CH ₃	4264469	Inactive	N/A

CH ₃ N CH ₃ CH ₃	4264491	Inactive	N/A
O OH H N CH3 CH3	4264513	Inactive	N/A
H ₃ C CH ₃	4265032	Inactive	N/A

HN CH ₃ CH ₃ CH ₃	4265342	Inactive	N/A
NH OSS CH _B	4265617	Inactive	N/A
$0 \stackrel{\circ}{=} \stackrel{\circ}{>} - N$ $0 \stackrel{\circ}{=} \stackrel{\circ}{>} - N$ CH_3	4265630	Inactive	N/A

S S S S S S S S S S S S S S S S S S S	4265764	Inactive	N/A
H ₃ C CH ₃	4265870	Inactive	N/A
S O O O O O O O O O O O O O O O O O O O	7965115	Inactive	N/A

H N CH ₃	7965493	Inactive	N/A
N HO	7966122	Inactive	N/A
CH ₃	7967020	Inactive	N/A

CH ₃ O NH NH NH CH ₃ CH ₃ O CH ₃ O NH NH NH CH ₃	7969138	Inactive	N/A
CH ₃ OH OH OH CH ₃ CH ₃	7971047	Inactive	N/A
O H N O CH ₃	7971114	Inactive	N/A

S S N CH ₃	7971311	Inactive	N/A
O NH	7972455	Inactive	N/A
H ₃ C H ₃ C	7972585	Inactive	N/A

O N S I O H N S S N N N N N N N N N N N N N N N N	7972860	Inactive	N/A
NH ONLING ON THE PARTY OF THE P	7972957	Inactive	N/A
H S CH ₃	7973018	Inactive	N/A

	7973318	Active	4.62
H ₃ C CH ₃ N — OH	7973389	Inactive	N/A
N NH ₂	7973713	Inactive	N/A

CH ₃	7973891	Active	25.95
H N N N N N N N N N N N N N N N N N N N	7973916	Inactive	N/A
$H_2N - N$ N N N N N N N	7974410	Inactive	N/A

H ₃ C S N O S	7975174	Inactive	N/A
O CH ₃ NH H ₃ C	7976940	Inactive	N/A
	7977382	Active	0.85

S O CH ₃	7977978	Inactive	N/A
	844155	Inactive	N/A
H ₃ C CH ₃ CH ₃	844595	Inactive	N/A

N CI	845431	Inactive	N/A
H ₃ C CH ₃	846984	Inactive	N/A
O NH	847433	Inactive	N/A

N HN O	847954	Inactive	N/A
O CH ₃ O CH ₃	848337	Active	8.54
CI CH ₃	848667	Inactive	N/A

H N NH CH ₃	849229	Inactive	N/A
CH ₃ I H ₃ C N CH ₃ O CH ₃	849832	Inactive	N/A
HN III N	850997	Inactive	N/A

	851186	Inactive	N/A
O N H S O CH ₃	851359	Inactive	N/A
NH O HO	851511	Inactive	N/A

CH ₃	851650	Active	17.07
H ₃ C N N N N N N N N N N N N N N N N N N N	852054	Inactive	N/A
HO H ₃ C CH ₃	852663	Inactive	N/A

NH C CH ₃	852804	Inactive	N/A
O HO N CH ₃ CH ₃	852974	Inactive	N/A
H ₁ C CH ₃	855949	Inactive	N/A

O OH NH	856413	Inactive	N/A
HO N S OH	856534	Inactive	N/A
N N CH	857178	Active	4.26

HO P CH ₃	859659	Inactive	N/A
CH ₃ HO N CH ₃ HO N CH ₃	860659	Inactive	N/A
H ₃ C N N N N N N N N N N N N N N N N N N N	862522	Inactive	N/A

H ₃ C O	863457	Inactive	N/A
H ₃ C N	863929	Inactive	N/A
CH ₃ CH ₃ CH ₃ CH ₃	864485	Inactive	N/A

H ₃ C N N CH ₃	865412	Inactive	N/A
H ₃ C N NH NH	866130	Inactive	N/A

1