

Building your own AI models

M3. Real-world cases

26.09.2023

Three real-world datasets

- Open Source Mycetoma
- HDAC1 inhibitors
- ACE2-Spike PPI inhibitors



Exercise

- Split in 3 groups.
- Understand the source of your data.
- Binarize the output variable.
- Divide your dataset into a training set and a test set.
- Train an AI model with LazyQSAR.
- Evaluate performance.
- Retrain model with full dataset and download it.



Tomorrow: present your results!



Deployment in the Ersilia Model Hub!

Target Report Card

Components

→ ↺ 🏠 🔒 ebi.ac.uk/chembl/target_report_card/CHEMBL325/

ChEMBL

Search in ChEMBL

🔍

⌵

☰

EBI > Databases > Chemical Biology > ChEMBL Database > ChEMBL325

Target Report Card

Name And Classification

Components

Relations

Drugs And Clinical Candidates

Activity Charts

Ligand Efficiencies

Associated Compounds

Gene Cross References

Protein Cross References

Domain Cross References

Structure Cross References

Name And Classification

-

🔗

ID:

CHEMBL325

Type:

SINGLE PROTEIN

Preferred Name:

Histone deacetylase 1

Synonyms:

HD1 HDAC1 Histone deacetylase 1 Protein deacetylase HDAC1 Protein decrotonylase HDAC1 RPD3L1 Synonyms=RPD3L1

Organism:

Homo sapiens

Species Group:

No

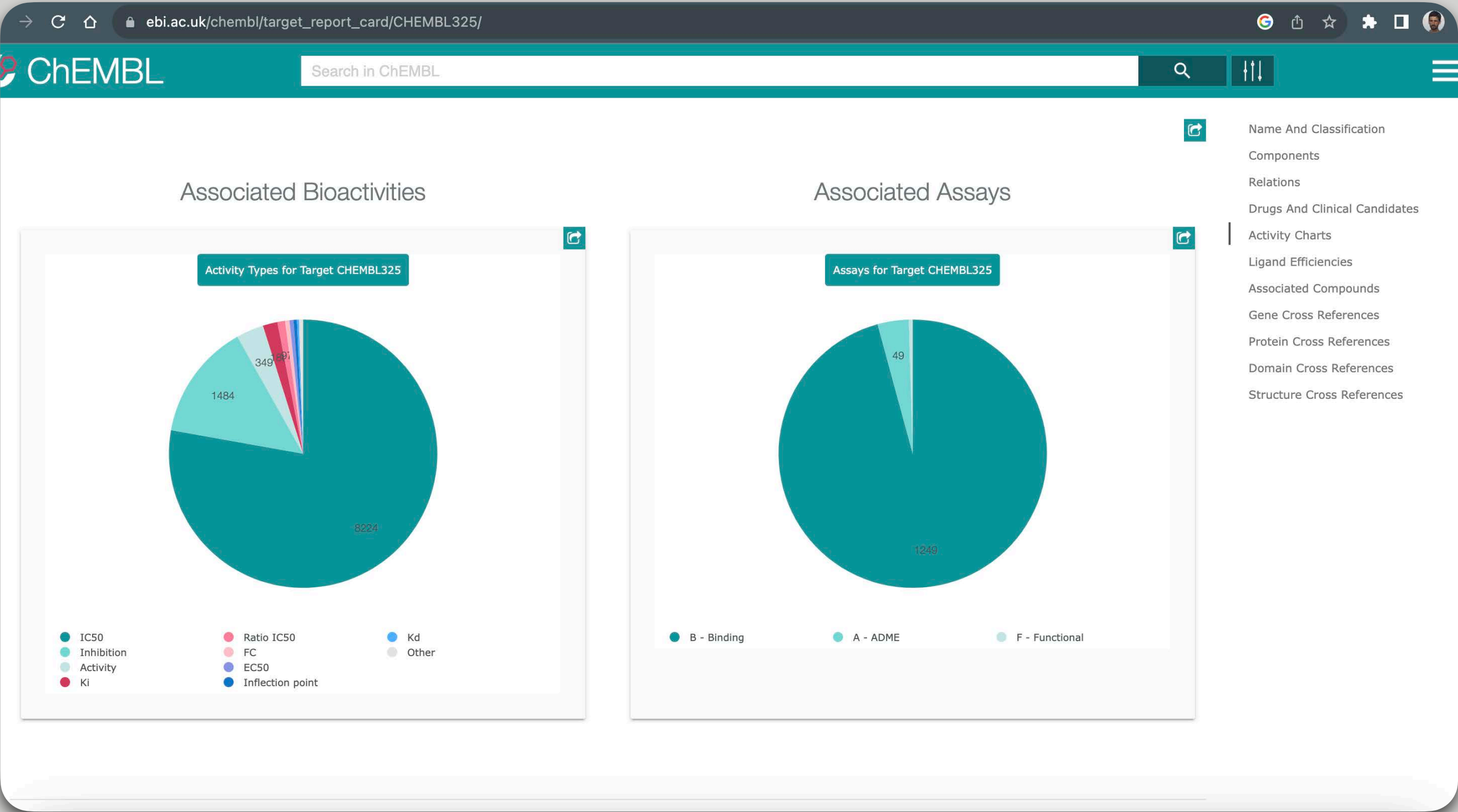
Protein Target Classification:

- Epigenetic regulator > Eraser > Histone deacetylase > HDAC class I

Components

-

HDAC1 inhibitors from ChEMBL (target 325)



Search in ChEMBL

Activities

8,224 items

Tools

RELATED COMPOUNDS

ASSAYS

SHARE

CSV

TSV

Filters

Custom

Custom Filtering

Per Property

Standard Type

HISTOGRAM

FIND_TERM

IC50

8,224

Target Type

Items per page

20

Showing 1-20 out of 8,224 records

<

1

2

...

411

412

>

<input type="checkbox"/>	Molecule ChEMBL ID ↑	Compound Key	Standard Type	Standard Relation	Standard Value	Standard Units	pChEMBL Value	Comment	Assay ID
<input type="checkbox"/>	<div><div></div><div>CHEMBL260957</div></div>	61c	IC50	=	3000.0	nM	5.52	---	CHE
<input type="checkbox"/>	<div><div></div><div>CHEMBL260957</div></div>	61a	IC50	=	1000.0	nM	6.00	---	CHE

MycetOS, series 1

An open source project

github.com/OpenSourceMycetoma

OpenSourceMycetoma

Type to search

OverviewRepositories12ProjectsPackagesPeople1

Open Source Mycetoma

New Medicines for Mycetoma without Secrecy. First timers: visit the "General Start Here" link below.

12 followers <http://opensourcemycetoma.org/> opensourcemycetoma@gmail.com

Follow

Pinned

Series-1-FenarimolsPublic

Open Source Mycetoma's First Series of Molecules

9 stars 2 forks

Series-2-AminothiazolesPublic

Series 2

Repositories

Find a repository...

TypeLanguageSort

Series-3-PhenothiazinesPublic

Project for the development of phenothiazine (and related tricyclic) anti-fungal drugs

0 stars 0 forks 4 (1 issue needs help) 0 watchers Updated on Aug 14

Monthly-zoom-meetingsPublic

Since MycetOS is diversifying we decided as from 2023 to place the monthly meetings no longer under a compound series but separately in this repository.

People

Top languages

C

Rich Text Format

Report abuse

MycetOS, series 1

An open source project

mycetoma Molecules

File Edit View Insert Format Data Tools Extensions Help

100% Calibri 11

		S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	
MYOS Codes		100 µM A	100 µM B	100 µM C	Percentage growth of M.mycetomatis at .. uM of compound												
!!READ ME before entering new MYOS codes!!	Other ID				5 µM C	25 µM mean	25 µM stdev	16µM	8 µM	4µM	2µM	1µM	0.5µM				
MYOS_00001_00_01	HPD14_1	-0.3			18.7	19.5	11.0	-1.6	-0.9	-1.0	-0.8	-0.7	3				
MYOS_00001_00_02	EPL-BS0495	2.8	0.2		-1.5	1.0	6.2	1.0	12.8	9.9	5.4	59.4	6				
MYOS_00002_00_01	EPL-BS0038	-0.9	-5.7			-3.1	0.1	-4.7	-5.2	-4.9	-5.3	-4.8	-				
MYOS_00003_00_01	EPL-BS0800	3.4	-0.9		2.5	-2.4	6.0	4.5	1.0	1.1	62.9	84.4	8				
MYOS_00003_00_02	HPD10_1	1.1			5.1	8.2	2.9	3.5	0.6	0.2	1.3	42.4	7				
MYOS_00004_00_01	EPL-BS1025	3.7	-0.2		1.6	-1.5	6.4	4.8	4.4	6.3	73.4	97.4	1				
MYOS_00005_00_01	HPD9-1	2.0			2.0			54.6	38.1	60.7	51.1	11.7	2.6	3.7	3.2	2.0	46.6
MYOS_00006_00_01	EPL-BS0237	15.1	4.3	11.2	10.2	5.5		9.7	13.4	11.1	9.7	1.9	15.6	9.6	14.0	13.5	43.7
MYOS_00007_00_01	EPL-BS0240	-1.6	-5.8		-3.7	3.0		-4.8	-6.2		-4.8	1.0	-4.3	-5.4	-6.8	17.0	60.4
MYOS_00008_00_01	EPL-BS0178	-4.9	3.3	-0.9	-0.8	4.1		19.5	64.7	-0.9	19.5	39.2	0.8	2.7	70.6	87.2	86.7
MYOS_00009_00_01	HPD20_1	17.5	15.5	8.5	13.9	4.7		7.4	6.0	7.2	6.9	0.7	-0.9	-1.9	-2.4	31.3	83.9
MYOS_00010_00_01	EPL-BS1246, MMV689244	-9.9	-1.2	-1.0	-4.0	5.1		0.5	12.2	-1.0	0.5	11.0	2.8	37.4	70.4	108.2	86.7
MYOS_00011_00_01	HPD12_1	-0.4			-0.4			65.2	61.4	20.7	49.1	24.7	-1.0	-1.3	4.5	41.9	59.1
MYOS_00012_00_01	HPD17_2F3	3.3	6.5	12.8	7.5	4.8		17.8	9.4	6.9	11.4	5.7	-0.5	-1.9	32.4	109.1	92.3
MYOS_00013_00_01	EPL-BS0549	0.1	6.5	7.9	4.8	4.2		16.9	40.7	15.2	16.9	14.2	8.1	9.2	50.1	64.8	73.9
MYOS_00014_00_01	EPL-BS0115	-9.4	-7.9		-8.6	1.0		-4.0	-3.9		-4.0	0.1	-3.7	-4.5	60.5	61.0	51.9
MYOS_00015_00_01	HPD30_1	5.1	5.8	6.5	5.8	0.7		7.8	4.4	18.4	10.2	7.3	-7.5	17.9	88.4	88.1	114.7
MYOS_00016_00_01	P4_C_003	2.5			58.9			78.7			78.7		82.0	79.7	53.3	63.0	87.3
MYOS_00016_00_01	DM7-1	58.9			2.5			18.0			18.0		0.1	-1.9	45.5	113.5	101.9
MYOS_00017_00_01	HPD25_1	9.2	13.9	7.4	10.2	3.4		4.0	6.7	8.1	6.3	2.1	-0.7	53.3	101.1	169.0	107.9
MYOS_00018_00_01	EPL-BS0132	-5.5	-5.9		-5.7	0.3		-4.8	-4.5		-4.8	0.3	-3.9	56.9	54.3	66.4	66.1
MYOS_00019_00_01	HPD8-1	0.4			0.4			63.9	60.2	53.9	59.3	5.0	1.9	56.1	56.0	65.1	73.4
MYOS_00020_01_01	Cetirizine.HCl	47.6			47.6			69.5	84.4	70.0	74.6	8.5	23.6	57.8	67.3	63.4	77.8
MYOS_00021_00_01	HPD3-4	1.6			1.6			41.9	43.5	39.0	41.5	2.3	2.5	65.7	67.7	72.0	73.8
MYOS_00022_00_01	EPL-BS0447	4.0	12.7	-0.1	5.5	6.5		22.2	53.8	14.5	22.2	28.5	7.2	69.1	85.9	85.5	82.7
MYOS_00023_00_01	EPL-BS0271	-5.5	-7.2		-6.3	1.2		4.1	-4.5		4.1	12.2	37.6	73.2	93.6	64.4	64.9
MYOS_00024_00_01	P4_C_001	28.1			28.1			115.5			115.5		54.5	68.8	69.8	71.6	78.2
MYOS_00024_00_01	DM6-1	83.8			83.8			46.8			46.8						
MYOS_00025_00_01	P4_B_012	89.7			89.7			76.2			76.2		67.8	72.5	70.4	76.1	82.7

% of Growth at 25uM
20% growth

12 MyOS Compounds Bio assay SI (Paper) Sheet8 MMVBoxes screened Full Epichem Library MYOS In vivo Sheet3 In vivo potency & LogD values

nature microbiology

View all journalsSearchLog in

Explore contentAbout the journalPublish with us

Sign up for alertsRSS feed

nature > nature microbiology > articles > article

Article | Published: 05 January 2023

Fluorogenic reporter enables identification of compounds that inhibit SARS-CoV-2

[Junjiao Yang](#), [Yinghong Xiao](#), [Peter V. Lidsky](#), [Chien-Ting Wu](#), [Luke R. Bonser](#), [Shiming Peng](#), [Miguel A. Garcia-Knight](#), [Michel Tassetto](#), [Chan-I Chung](#), [Xiaoquan Li](#), [Tsuguhisa Nakayama](#), [Ivan T. Lee](#), [Jayakar V. Nayak](#), [Khadija Ghias](#), [Kirsten L. Hargett](#), [Brian K. Shoichet](#), [David J. Erle](#), [Peter K. Jackson](#), [Raul Andino](#) ✉ & [Xiaokun Shu](#) ✉

Nature Microbiology **8**, 121–134 (2023) | [Cite this article](#)

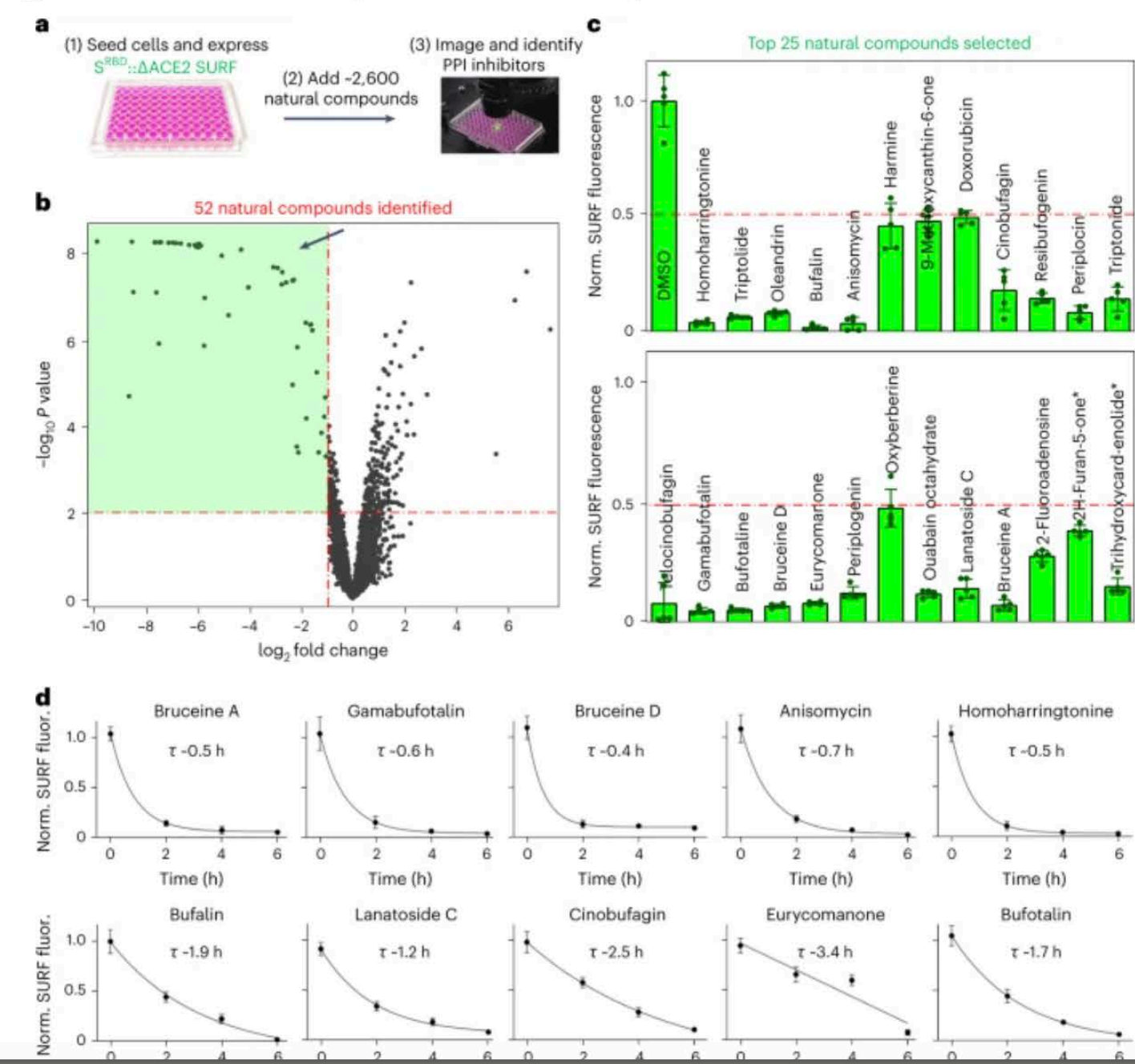
5742 Accesses | 3 Citations | 29 Altmetric | [Metrics](#)

AbstractMainResultsDiscussionMethodsData availabilityReferencesAcknowledgementsAuthor informationEthics declarations

The coronavirus SARS-CoV-2 causes the severe disease COVID-19. SARS-CoV-2 infection is initiated by interaction of the viral spike protein and host receptor angiotensin-converting enzyme 2 (ACE2). We report an improved bright and reversible fluorogenic reporter, named

ACE2-Spike protein-protein interaction

Fig. 2: HPS of natural compounds that inhibit spike and ACE2 interaction.



Sections

Figures

References

[Abstract](#)

[Main](#)

[Results](#)

[Discussion](#)

[Methods](#)

[Data availability](#)

[References](#)

[Acknowledgements](#)

[Author information](#)

Advertisement



ACE2-Spike protein-protein interaction

41564_2022_1288_MOESM6_ESM

HomeInsertDrawPage LayoutFormulasDataReviewViewAutomateAcrobatTell me

Paste

Calibri (Body)11

B

I

U

General

%

0.00

0.00

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

Σ

Sort & Filter

Find & Select

Analyse Data

Sensitivity

Create PDF and share link

CommentsShare

N16

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Data are presented as Ex488/561 ratio normalized to DMSO control																				
2	drug CAS	drug name	normalized Ex488/561 ratio					aver													
3			rep1	rep2	rep3	rep4	rep5														
4	-	DMSO-1	1,2453825	0,8468458	1,0341802	1,003861	0,8697306	-													
5	-	DMSO-2	0,8973923	1,0186663	0,9241451	1,0744147	1,0853816	-													
6	-	DMSO-3	0,8556126	0,836461	1,1588629	1,3637858	0,7852778	-													
7	146-78-1	2-Fluoroade	0,0013533	4,388E-06	0,0009403	0,0026946	2,977E-05	0,001													
8	465-11-2	Gamabufot	1,152E-05	0,0020265	0,0006183	3,709E-07	0,0092557	0,002													
9	218600-44-	Bardoxolon	0,0047179	0,0035831	0,0021486	0,0015305	0,0009023	0,0025788	5,100E-05	5,000000E-05	5,000000E-05	5,000000E-05	5,000000E-05								
10	464-74-4	Arenobufag	0,0019898	0,0075244		0,0002397	0,0009384	0,0026731	7,662E-08	-8,5472745	7,1156717										
11	0407-0024	3,12,14-Trih	0,0119533	0,0048782	0,0002098	0,0005202	0,0062113	0,0047545	5,314E-09	-7,716476	8,2746126										
12	38748-32-2	Triptolide	0,0103638	0,0010921	0,0014085	0,0003536	0,0111511	0,0048739	5,336E-09	-7,6807213	8,2728201										
13	11018-89-6	Ouabain octahydrate		0,0084743	0,0002389	0,0001962	0,01103	0,0049849	7,842E-08	-7,6482297	7,1055834										
14	465-21-4	Bufalin	0,0110539		0,0016779	0,0033035		0,0053451	1,207E-06	-7,5475661	5,918342										
15	84633-29-4	EURYCOMA	0,0022862	0,0046049	0,0031705	0,0087305	0,0085262	0,0054637	5,303E-09	-7,515919	8,275446										
16	13137-64-9	Periplocin	0,0066532	0,0068448	0,0070189	0,0030352	0,0046325	0,0056369	5,294E-09	-7,4708736	8,2762441										
17	470-37-1	Cinobufagin	0,0143832	0,0024584	0,0019284	0,0016794	0,0142099	0,0069319	5,471E-09	-7,1725414	8,2619401										
18	38647-11-9	Triptonide	0,0123639	0,0004323	0,0048107	0,0109471	0,0105457	0,0078199	5,451E-09	-6,9986267	8,2634889										
19	471-95-4	Bufotaline	0,0186271	0,0101922	0,0116832	0,0011288	0,0007743	0,0084811	5,575E-09	-6,8815283	8,253767										
20	13190-34-6	Goitrin	0,0188075	0,0177691	0,0019906	0,0013337	0,0083233	0,0096448	5,664E-09	-6,6960305	8,2468511										
21	465-16-7	Oleandrin	0,0076275	0,022397	0,0004146	0,0001494	0,0299763	0,0121129	6,127E-09	-6,367306	8,2127205										
22	514-39-6	Periplogenin	0,0198943	0,0079685	0,0016416	0,0333346	0,0073685	0,0140415	6,155E-09	-6,1541592	8,2107548										
23	17575-22-3	Lanatoside	0,0414286	0,0023382	0,0056684	0,0233967	0,0003235	0,0146311	6,678E-09	-6,0948201	8,1753839										
24	472-26-4	Telocinobuf	0,0090175	0,0087866	0,0130127	0,0269224	0,0191804	0,0153839	5,898E-09	-6,0224339	8,2293142										
25	21499-66-1	Bruceine D	0,0027256	0,0002299	0,0010757	0,0389398	0,0343851	0,0154712	6,968E-09	-6,0142686	8,1568884										
26	0407-0023	3-[(9S)-3,14	0,0171385	0,0126213	0,0128531	0,0151403	0,0208338	0,0157174	5,765E-09	-5,9914947	8,2392324										
27	26833-87-4	Homoharrir	0,0293807	0,0036175	0,0150109	0,002814	0,0322404	0,0166127	6,39E-09	-5,9115682	8,1945241										
28	22862-76-6	Anisomycin		0,0342691		0,0055195	0,0145847	0,0181244	1,353E-06	-5,7859195	5,868599										
29	62996-74-1	Staurosporine		0,010291	0,0041961	0,0043746	0,0547216	0,0183958	1,044E-07	-5,7644785	6,981378										
30	25316-40-9	Doxorubicin	0,0066792	0,0021658	0,009806	0,0514215	0,0756002	0,0291345	1,089E-08	-5,1011256	7,9628803										
31	25514-31-2	Bruceine A		0,0025467	0,1212423	0,0002626	0,0160866	0,0350345	2,64E-07	-4,8350789	6,5783535										

Fig2bFig2cFig2d+

50% PPI inhibition

log2(0.5) = -1

Let's get started!

- Work on your Colab notebook
- Prepare a presentation