Computational Chemical Biology Research Group

Connecting Compound Structures to Biological Effects by Gene-Expression Profiling

David Jaramillo

Mentor: Mathias Wawer

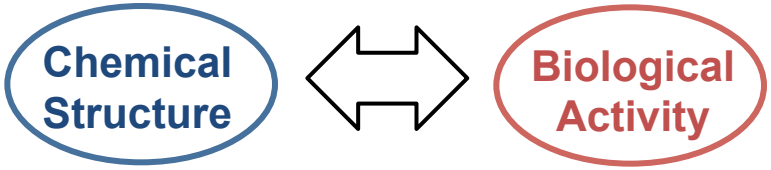
PI: Paul Clemons

Summer Research Program in Genomics, 2012

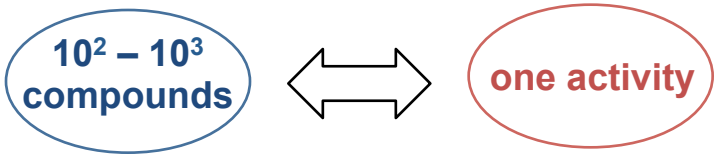
BROAD INSTITUTE

Introduction: structure-activity relationships

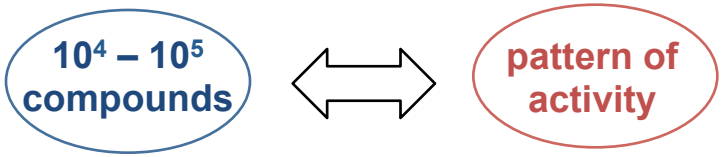
SARs



Traditional SARs



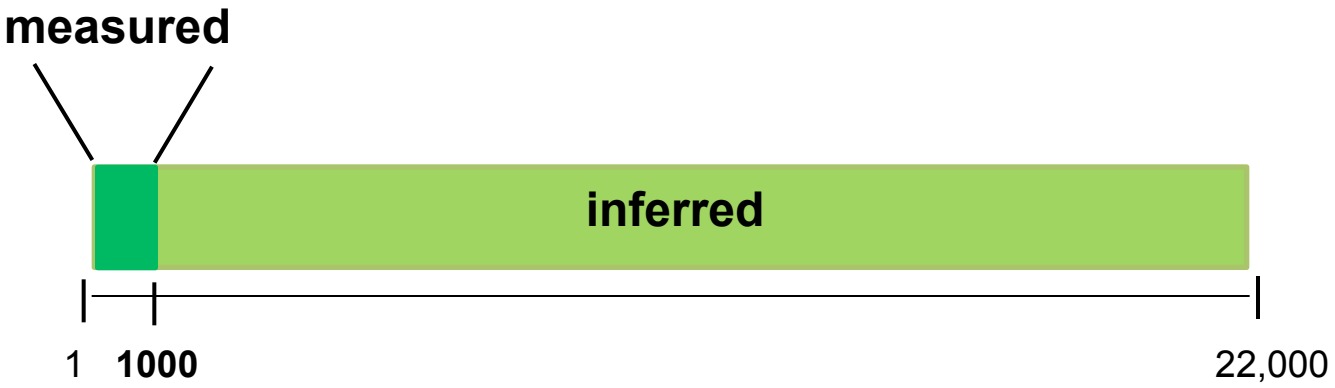
Our novel approach to SARs



But how can this be done in a high-throughput fashion?

Luminex 1000 (L1000)

1000 genes capture 80% of all gene-expression variation

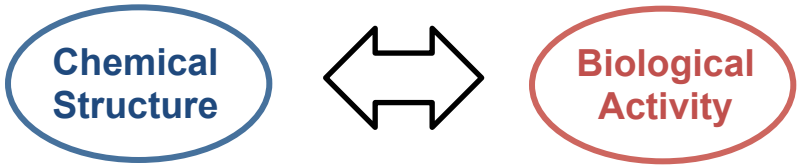


Only need 1000 gene-expression levels to have a strong idea of the overall state of the cell

Done in collaboration with GAP and the Cmap team

Importance

SARs



Importance: Suggest new research efforts

- Propose biological function of novel compounds
- Advise the synthetic chemist on which chemical features to prioritize in their synthetic efforts

Compounds

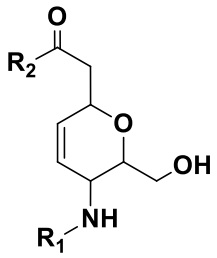
~19,000 compounds

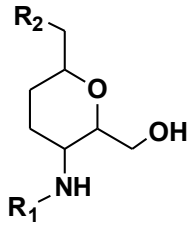
- 2,000+ bioactives
- 17,000+ compounds originated from diversity-oriented synthesis (DOS)
  - Many analogs for many different chemotypes
  - Defined structural relationships

Compounds

~19,000 compounds

- •2,000+ bioactives
- •17,000+ compounds originated from diversity-oriented synthesis (DOS)
  - •Many analogs for many different chemotypes
  - •Defined structural relationships : **cores**

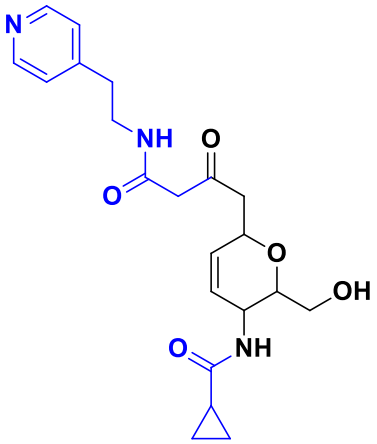


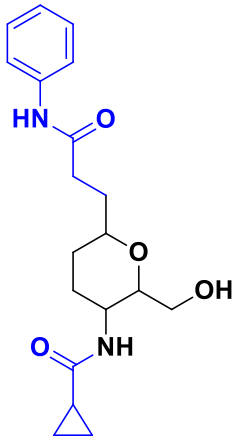


Compounds

~19,000 compounds

- 2,000+ bioactives
- 17,000+ compounds originated from diversity-oriented synthesis (DOS)
  - Many analogs for many different chemotypes
  - Defined structural relationships : cores, **appendages**

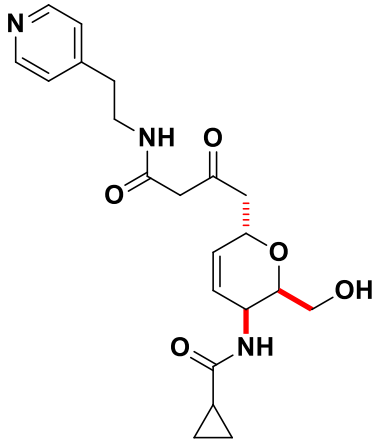


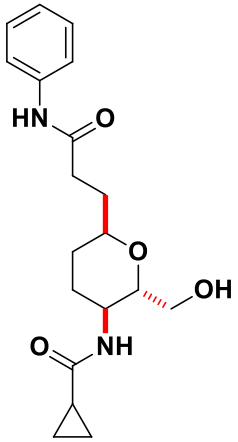


Compounds

~19,000 compounds

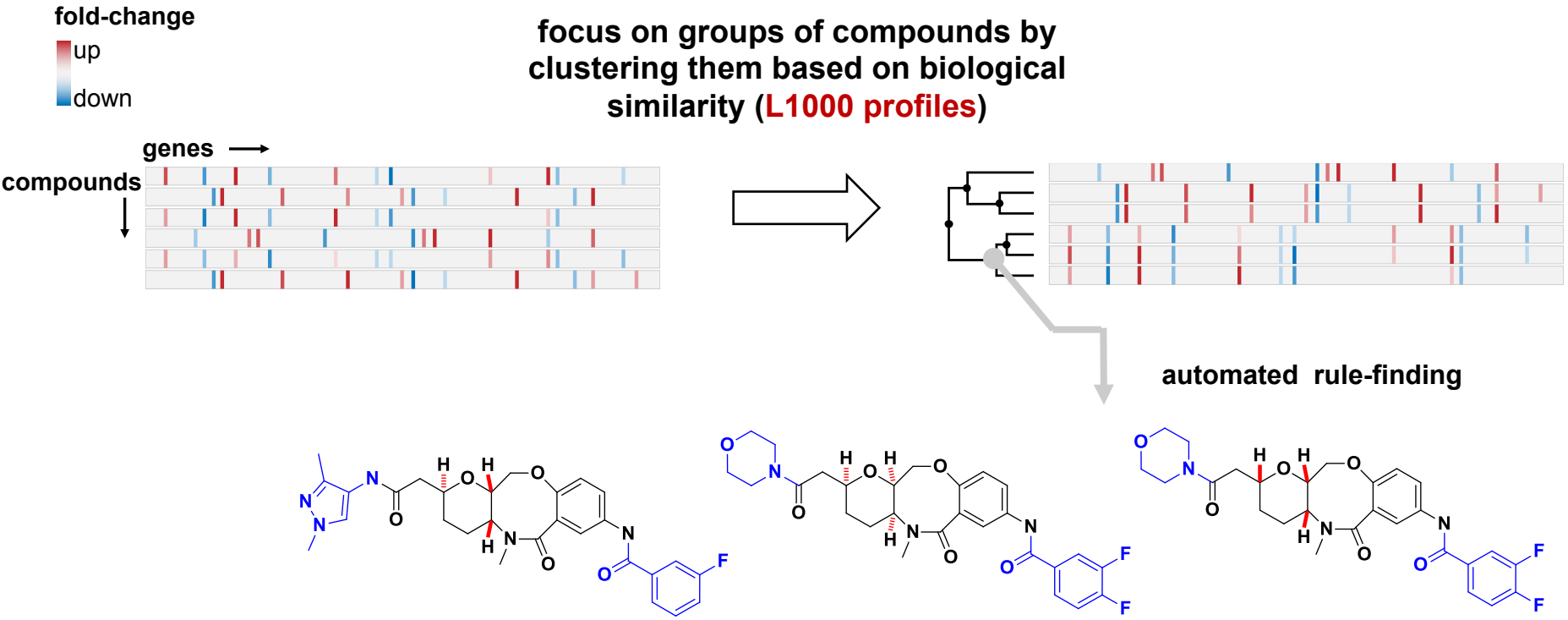
- 2,000+ bioactives
- 17,000+ compounds originated from diversity-oriented synthesis (DOS)
  - Many analogs for many different chemotypes
  - Defined structural relationships : cores, appendages, and **stereochemistry**





Connecting structure to activity

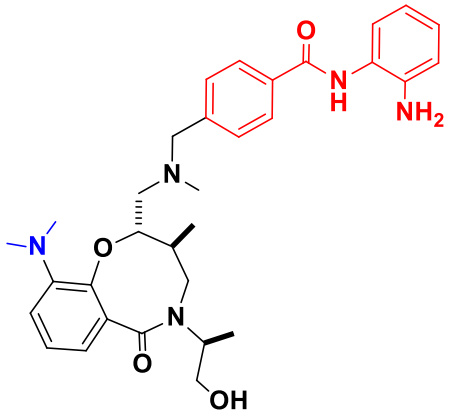
22,000+ compounds ⟺ L1000 profiles



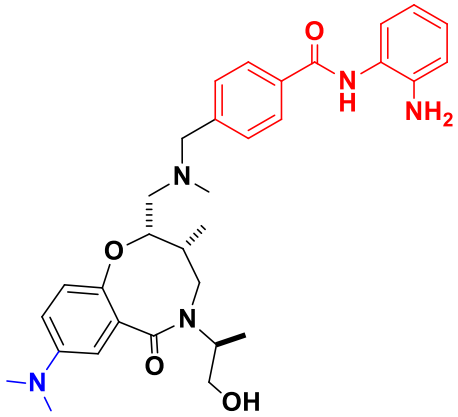
8 DOS compounds co-cluster with known 2 HDAC inhibitors

- The 8 DOS compounds consist of 3 similar cores
- All 10 compounds have appendage 1 ("benzanilide")
- 7 DOS compounds have appendage 1 ("benzanilide") and appendage 2

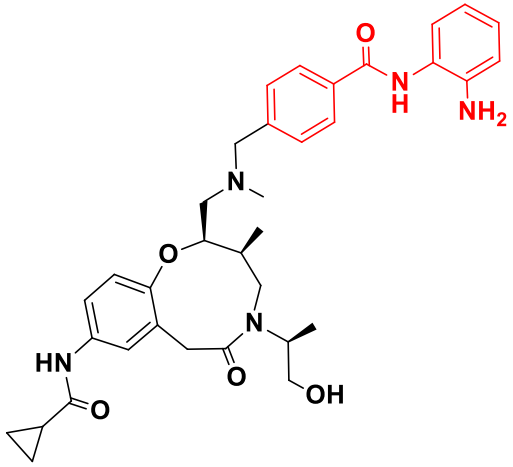
Core 1
5 stereoisomers



Core 2
2 stereoisomers



Core 3
1 stereoisomer

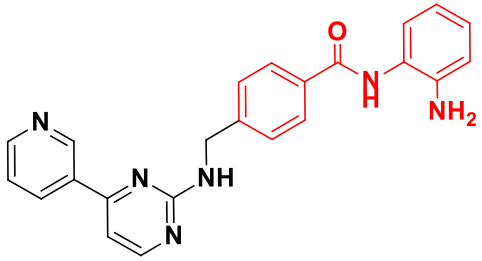


focus on these stereoisomers ⇨ SARs

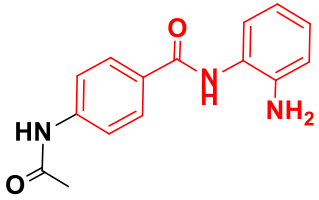
8 DOS compounds co-cluster with known 2 HDAC inhibitors

Cluster contains 2 known chemotherapeutic agents, histone deacetylase inhibitors (HDACi)

mocetinostat



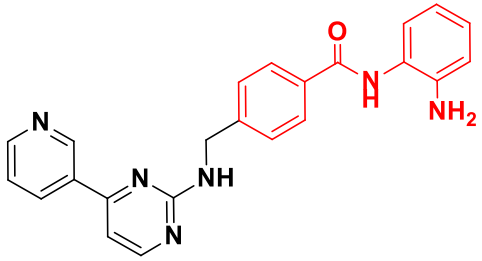
tacedinaline



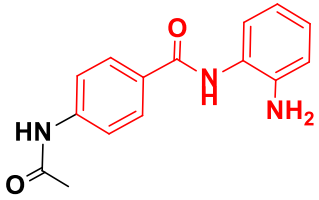
8 DOS compounds co-cluster with known 2 HDAC inhibitors

Cluster contains 2 known chemotherapeutic agents, histone deacetylase inhibitors (HDACi)

mocetinostat



tacedinaline

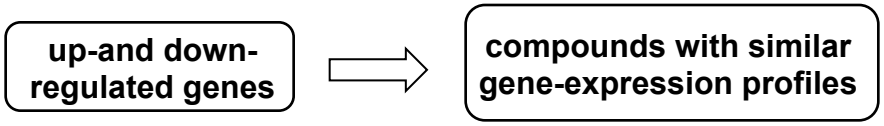


Do the DOS compounds in cluster share this HDACi activity?

HDACi Activity of DOS Compounds: Cmap

Queried the Connectivity Map by using the average L1000 profile of the 8 DOS compounds

Cmap



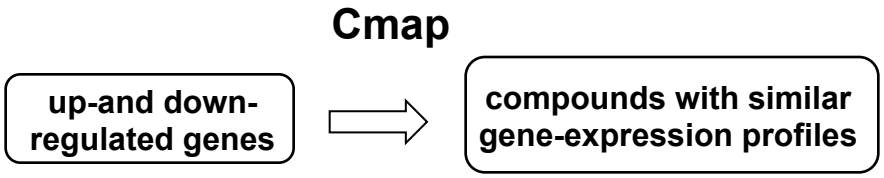
Cmap Results

Rank	Cmap name	Score	HDACi class
1	MS-275	0.98	benzanilide
2	SAHA	0.89	hydroxamic acid
3	TSA	0.78	hydroxamic acid
4	scriptaid	0.54	hydroxamic acid

J.Lamb et al.,Science 313, 1935 (2006)

HDACi Activity of DOS Compounds: Cmap

Queried the Connectivity Map by using the average L1000 profile of the 8 DOS compounds



Cmap Results

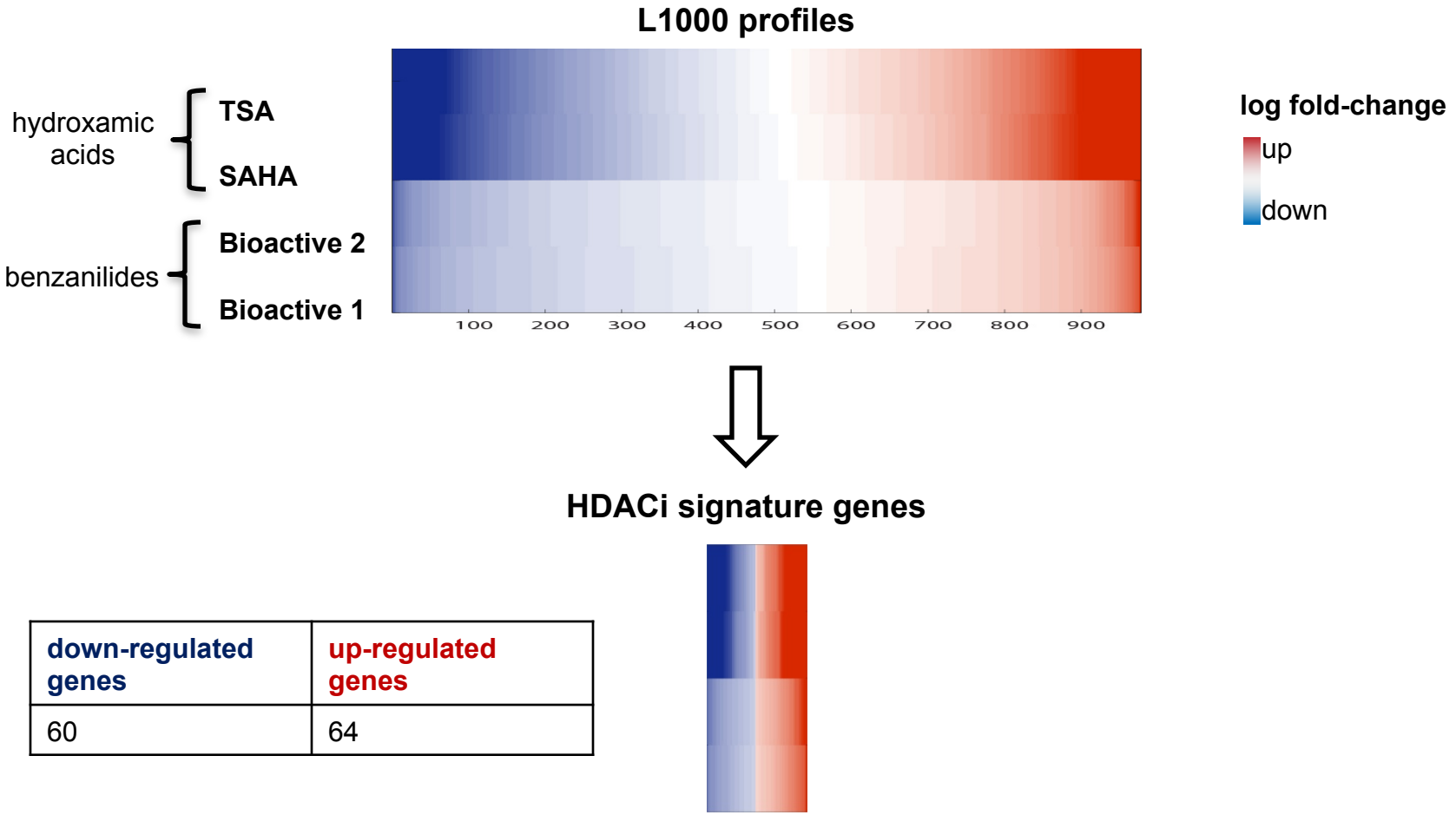
Rank	Cmap name	Score	HDACi class
1	MS-275	0.98	benzanilide
2	SAHA	0.89	hydroxamic acid
3	TSA	0.78	hydroxamic acid
4	scriptaid	0.54	hydroxamic acid

What chemical features from the DOS compounds are important for HDACi activity?

J.Lamb et al.,Science 313, 1935 (2006)

HDACi signature genes

define HDACi activity by genes regulated in same direction by all four HDACis

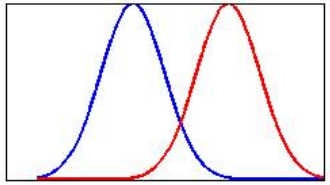


down-regulated genes	up-regulated genes
60	64

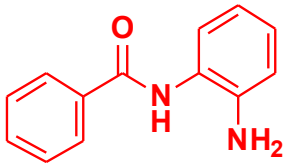
Importance of chemical features

chemical features and combinations

test if means are different



correlation to benzanilide HDACi profile



	appendage 1 (n= 28)	not (appendage 1) (n= 17,532)
average correlation	0.3071	0.0132

$p = 2.37 \times 10^{-9}$

DOS compounds show different patterns of HDACi activity

benzanilide gene-expression

functional annotation of gene sets

ranked list of biological functions

functional term	rank	p-value
acetylation	1	3.9×10^{-13}
vasculature development	4	3.6×10^{-8}
transcription activator activity	17	9.7×10^{-6}
duplication	21	1.2×10^{-5}
regulation of apoptosis	45	7.5×10^{-5}

Nature Protocols 2009; 4(1):44 & Nucleic Acids Res. 2009;37(1):1

DOS compounds show different patterns of HDACi activity

benzanilide gene-expression

functional annotation of gene sets

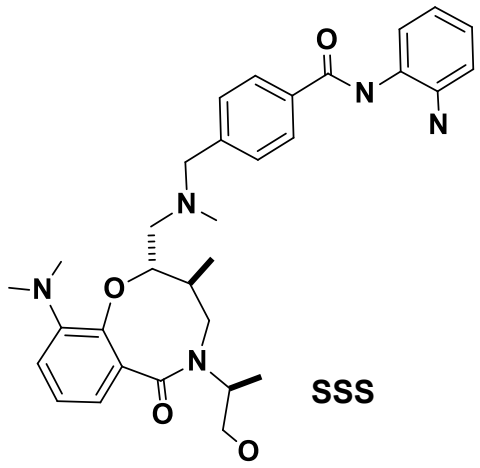
ranked list of biological functions

functional term	rank	p-value
acetylation	1	3.9×10^{-13}
vasculature development	4	3.6×10^{-8}
transcription activator activity	17	9.7×10^{-6}
duplication	21	1.2×10^{-5}
regulation of apoptosis	45	7.5×10^{-5}

ranked list of benzanilide functions (top 50)

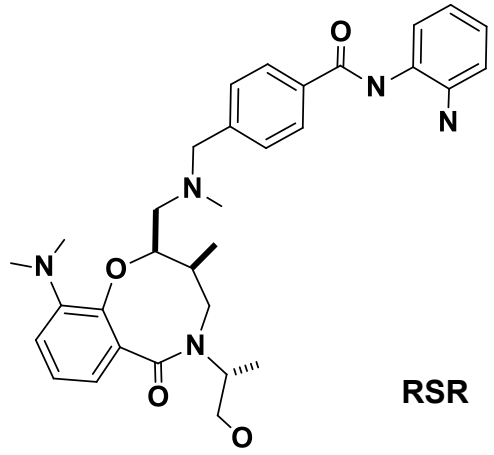
Nature Protocols 2009; 4(1):44 & Nucleic Acids Res. 2009;37(1):1

DOS compounds show different patterns of HDACi activity



SSS

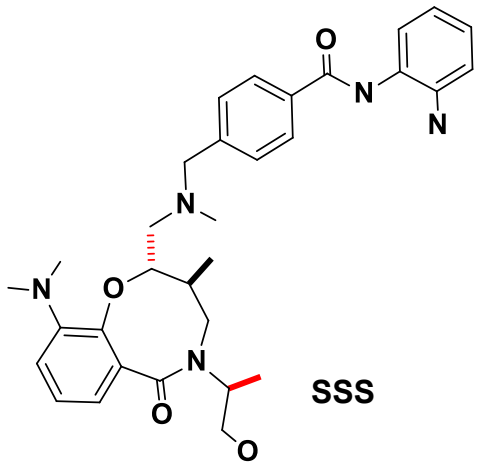
functional term	rank
blood vessel development	1
vasculature development	2
blood vessel morphogenesis	3
LIM domain	4
Zinc finger, LIM-type	5



RSR

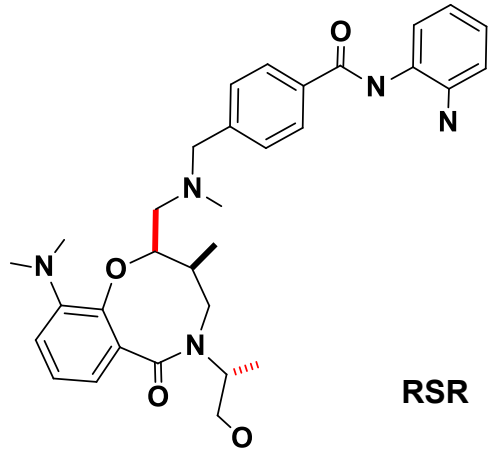
functional term	rank
endopeptidase inhibitor activity	1
peptidase inhibitor activity	2
regulation of apoptosis	3
regulation of programmed cell death	4
oxidation reduction	5

DOS compounds show different patterns of HDACi activity



SSS

functional term	rank
blood vessel development	1
vasculature development	2
blood vessel morphogenesis	3
LIM domain	4
Zinc finger, LIM-type	5



RSR

functional term	rank
endopeptidase inhibitor activity	1
peptidase inhibitor activity	2
regulation of apoptosis	3
regulation of programmed cell death	4
oxidation reduction	5

Future work

- confirm HDACi activity of DOS compounds with an orthogonal method (e.g. mass spectrometry proteomics)

- further explore SARs around identified compounds by synthesizing novel structures

Acknowledgements

Stuart Schreiber

Clemons Group
Amrita Basu, PhD
Nicole E Bodycombe
Vlado Dancik, PhD
Kejie Li, PhD

CDP Team
Sigrun Gustafsdottir
Alkyhan Shamji
Jeremy Duvall
Joshua Bittker

GAP Team

Cmap Team

Diversity Inititatives
Bruce Birren
Eboney Smith
Francie Latour

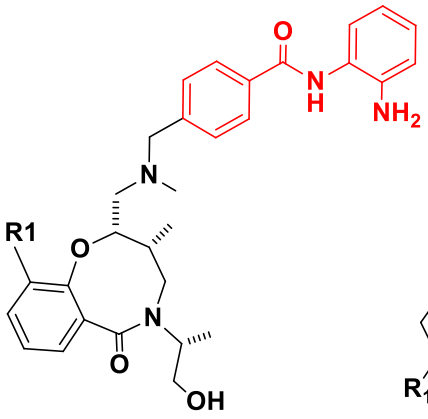
Scientific Communications
Brandon Ogbunugafor

SRPG 2012 Students

BROAD INSTITUTE

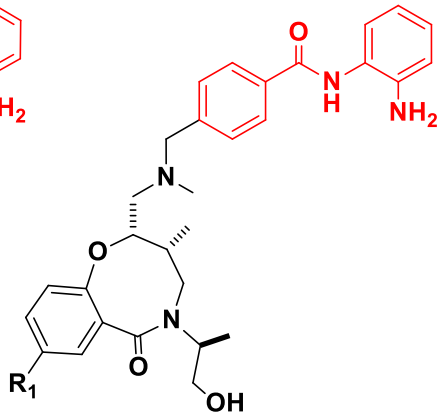
Importance of Cores with appendage 1

Core 1



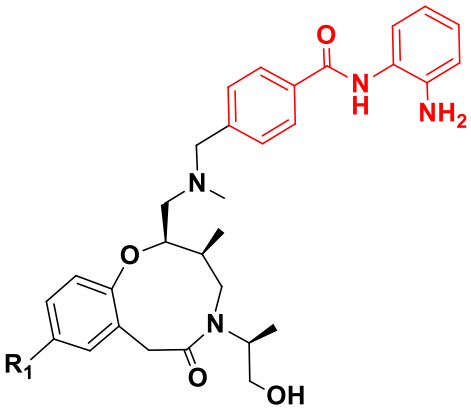
Avg. Correlation: 0.4326

Core 2



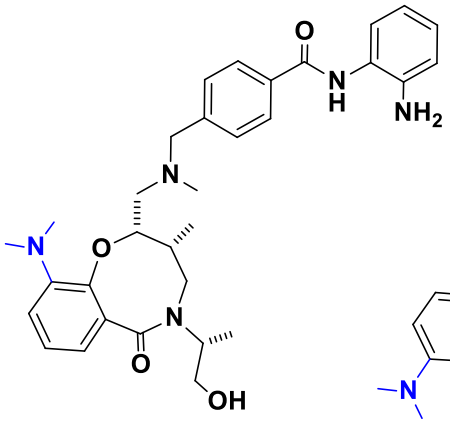
Greater importance for HDACi activity

Core 3



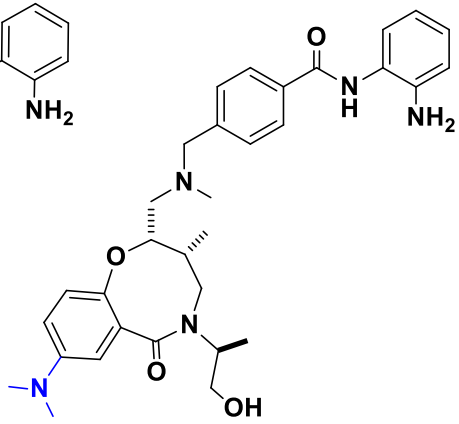
Avg. Correlation: 0.1983

Core 1



Avg Correlation: 0.3823

Core 2

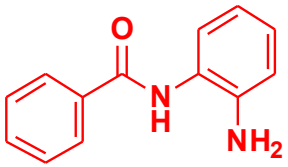


Avg Correlation: 0.5025

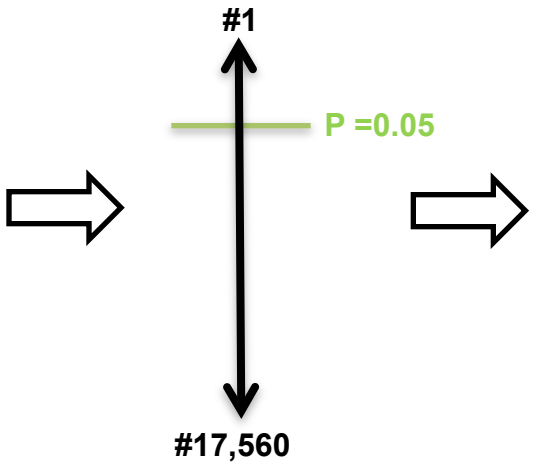
No core is more important than the other ⇨ **Ortho versus Para positioning doesn't matter**

Importance of Appendage 1

How important is appendage 1?



Rank all DOS compounds based on correlation to Benzanilides



20/28 are above a significant threshold

859 DOS compounds are above threshold and don't have appendage 1

Appendage 1 is not necessary for high correlation with HDACi benzanilide activity

Appendage 2 versus Not(appendage 2), p =1

Cores versus Not(Cores), p= 0.37