

Assisted Data Exploration in Chemical Biology

Dennis Moccia
Chemaxon UGM 2012

About Us

- Established in 2012
- Offers Consulting Services
 - Focused on Assay Capture and Data Analysis
 - Custom code, Configuration, Design and Decision Making
- Currently in development of Discovery Analysis Tools
 - Grendel: Web Based Visualizations
 - Fenris: Hit-to-Lead Platform

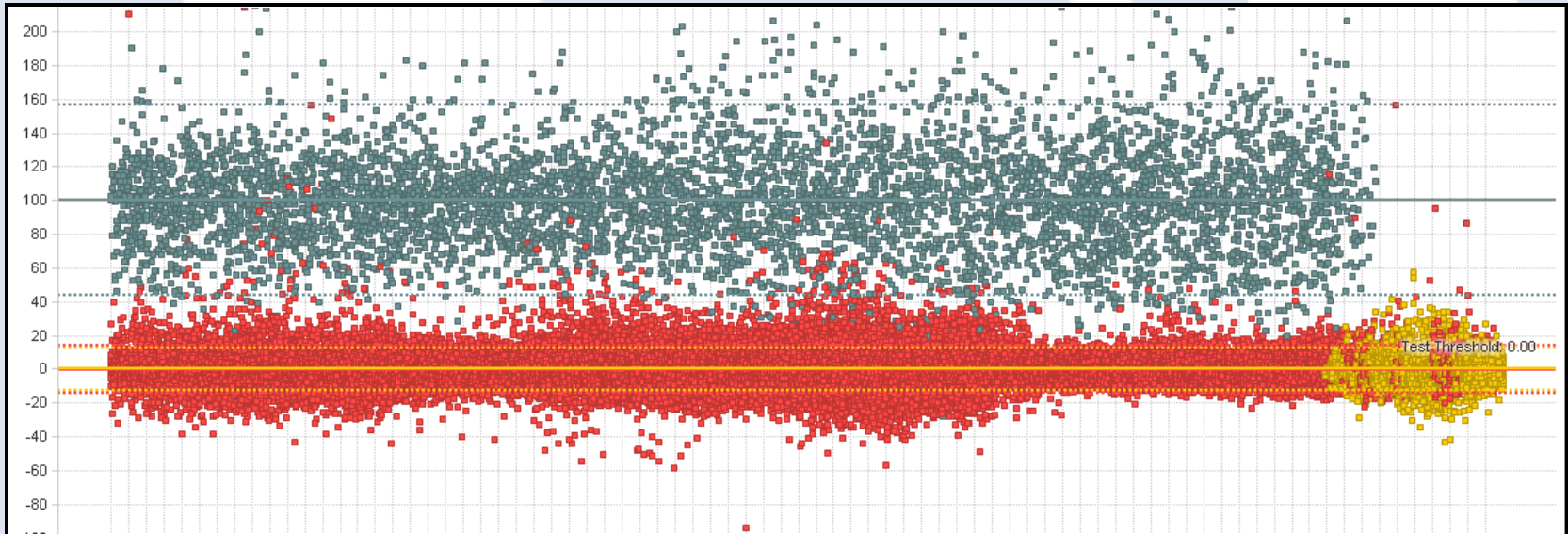
Development Philosophy

- Rapid Development with Flexible Backend
- Java/Groovy/Grails
- Heavy use of Open Source Javascript Libraries
- Oracle/MongoDB
- Leave Power in Users Hands...But Not Too Much!

Overview of Grendel

- Enter Drug Discovery Life Cycle Post Hit Calling
- Trying to Capture Decisions Made During Medchem Program in Structured Way
- Provides Some Guidance to End User
- Scientist Can Explore Their Data

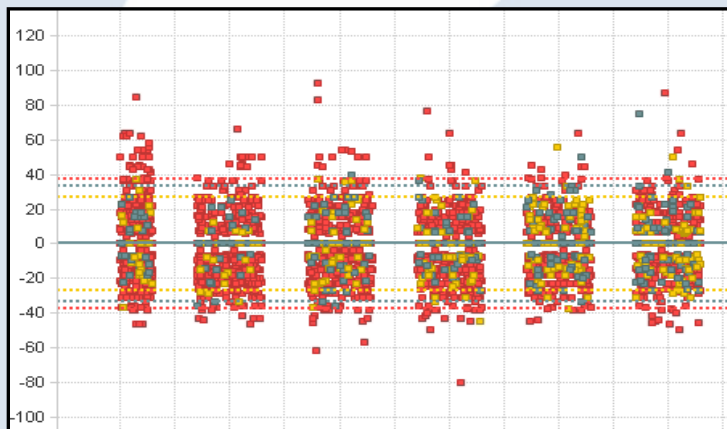
Traditional HTS Follow-up (Cherry Picking)



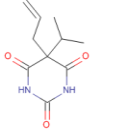

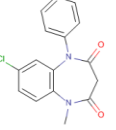

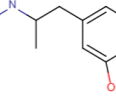

- Take the top 1% of compounds
- Take a predetermined number (2000)
- Determine from visualization/analysis

Modern Approach to Cherry Picking







Review



Represent

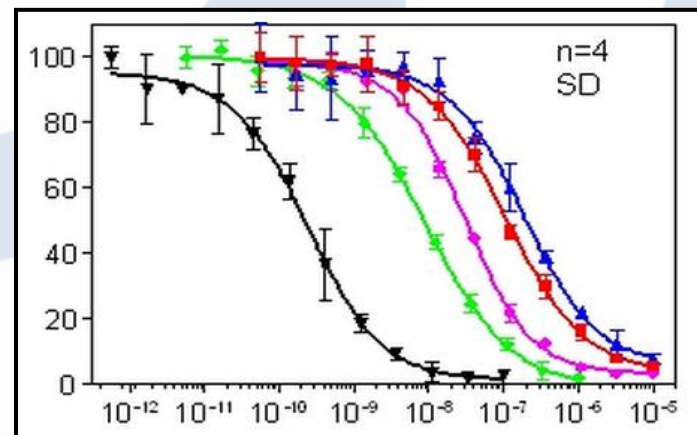
Cluster Center	Cluster	Compounds in Cluster	Average Distance to Centroid	Activity	Total Selected for Cherry Pick
	1	32	0.0125		6
	2	56	0.263		12
	3	47	0.0587		1

Remove

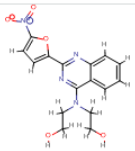
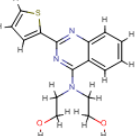
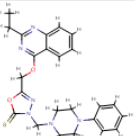
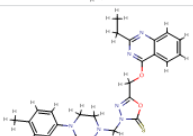
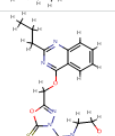
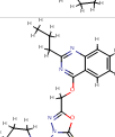
Nitroso	
Aziridine	
Epoxide	
Primary alcohol	
Azide	
Disulfide	

EXACTMASS	107.07	669.39
HBONDACCEPTCOUNT	12	20
HBONDDONORCOUNT	0	18

Re-Screen

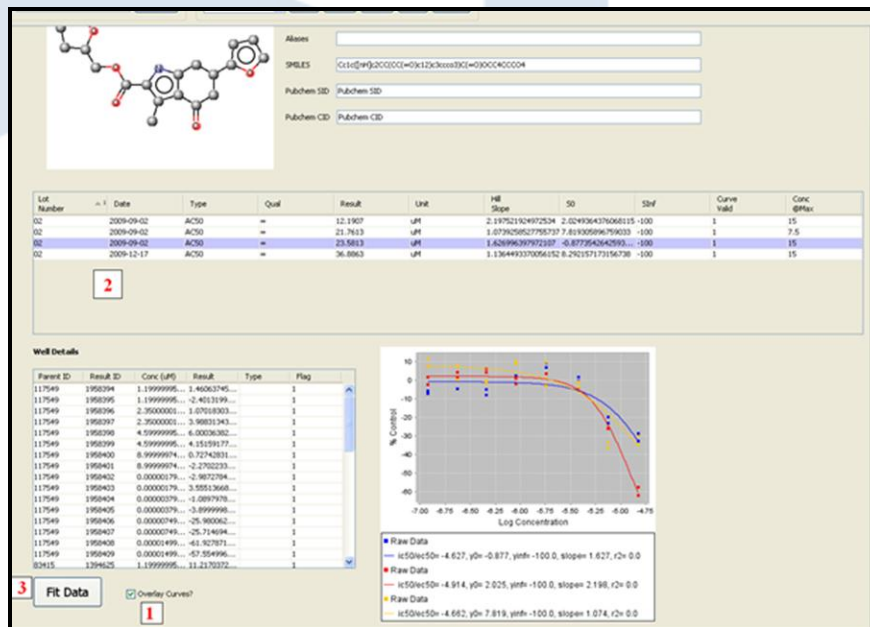
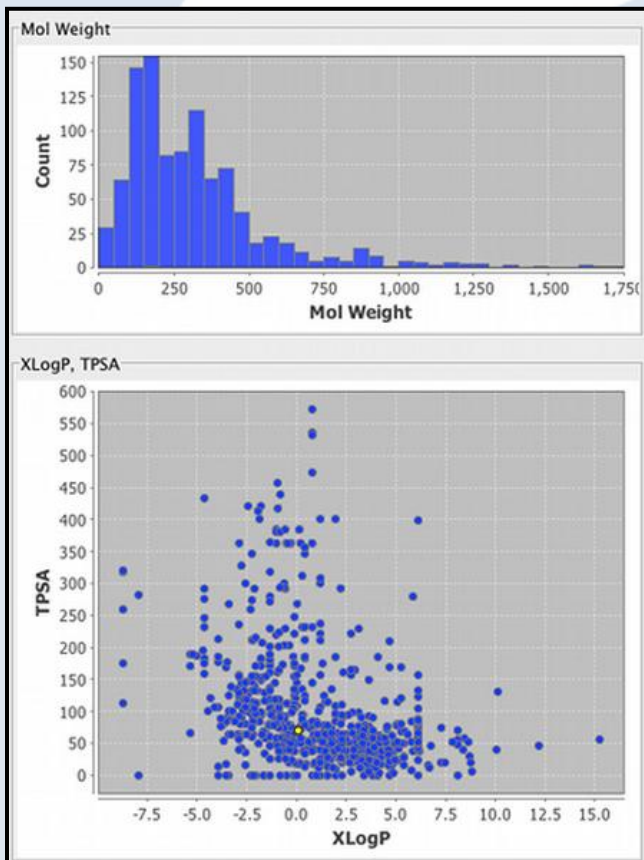


Traditional SAR Tools

Structure	SampleID	HBA	HBD	RB	MASS	Luminescence-based cell-based MC4R Dose	Counterscreen MC4R
	21139	8	2	6	344.11207	27.3	10.12
	42648	6	2	6	315.104148	85.4	12.2
	57098	8	0	7	462.183795	60.2	17.5
	57099	8	0	7	476.199445	95.2	16.33
	57100	8	0	7	401.152161	30.4	95.2
	57101	7	0	7	399.172896	19.11	12.5

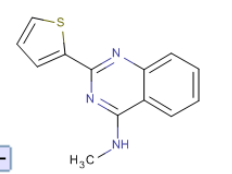
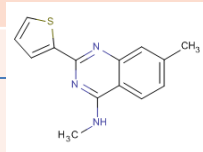
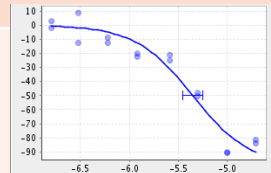
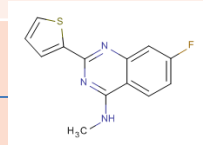
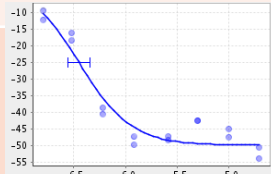
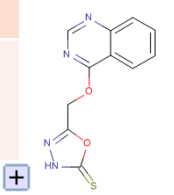
- Sorting
- Heatmaps
- Annotation
- Calculated Fields
- R Group Decomposition

Contemporary SAR Tools



- Visualizations
- External Calculations
- Form Views
- Curves/Images

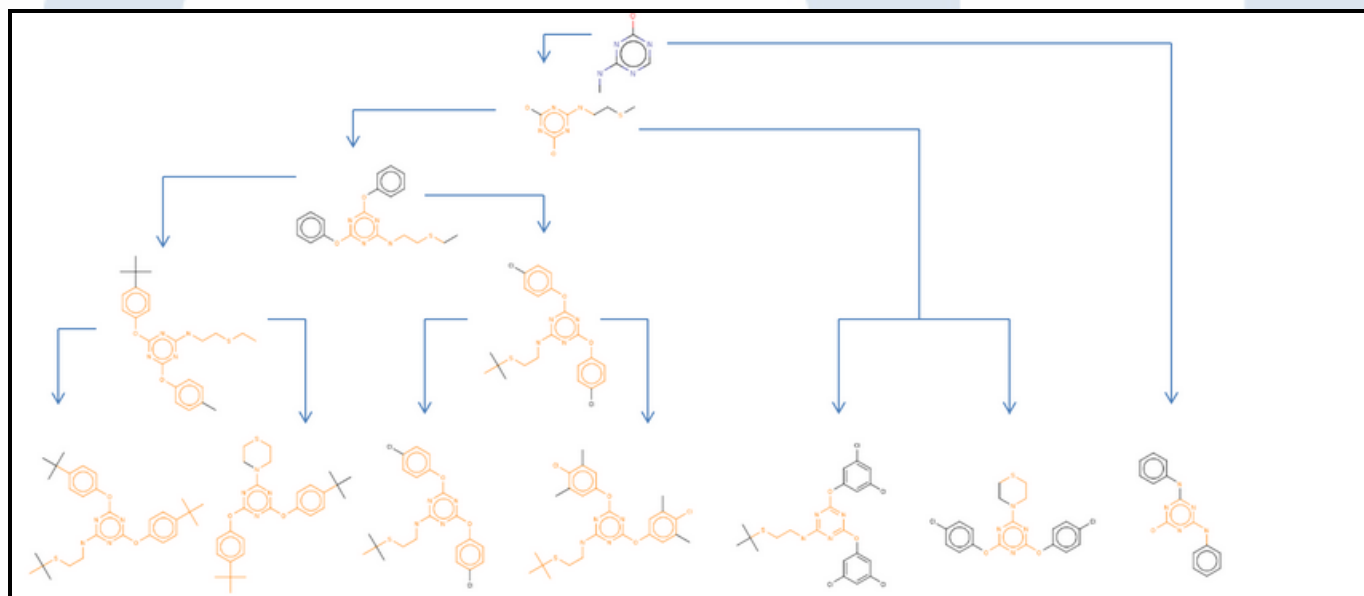
Scaffold/Clustering Techniques

	Cluster	Cluster Count	Luminescence-based cell-based MC4R Dose (IC50 μ M)	Dose Curve	Counterscreen MC4R
	1	2			
			0.68		87.2
			4.92		12.1
	2	14			

Scaffold/Clustering Techniques

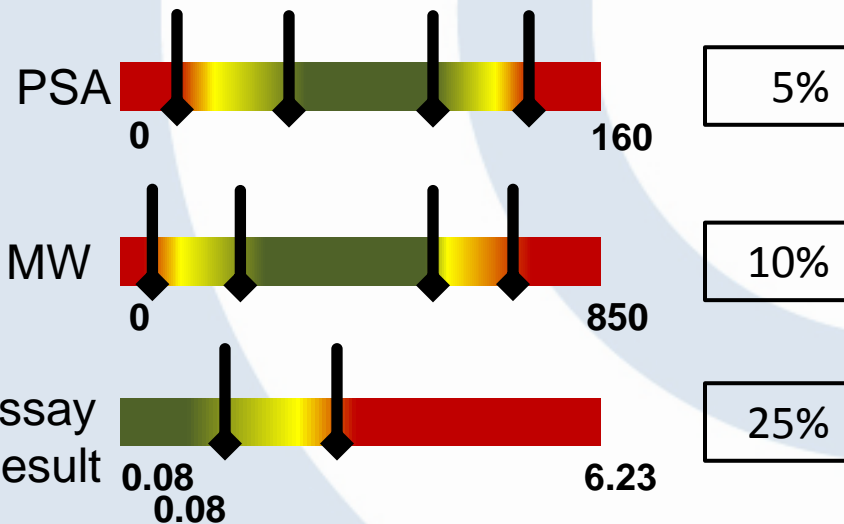
Class Hierarchy

- java.lang.Object
 - chemaxon.clustering.[Common](#)
 - chemaxon.clustering.[Compare](#)
 - chemaxon.clustering.[NearestNeighbors](#)
 - chemaxon.clustering.[JarvisPatrick](#)
 - chemaxon.clustering.[Ward](#)
 - chemaxon.clustering.[LibraryMCS](#) (implements chemaxon.license.Licensable)
 - chemaxon.clustering.[LibraryMCS.ClusterEnumerator](#)



Enhanced Sorting

- Traditional Sorting is “Good Enough”
- Can we enhance it?
- Ranges and Weighting...



better leads??

Medchem Funnel

HTS- 250K Compounds Tested

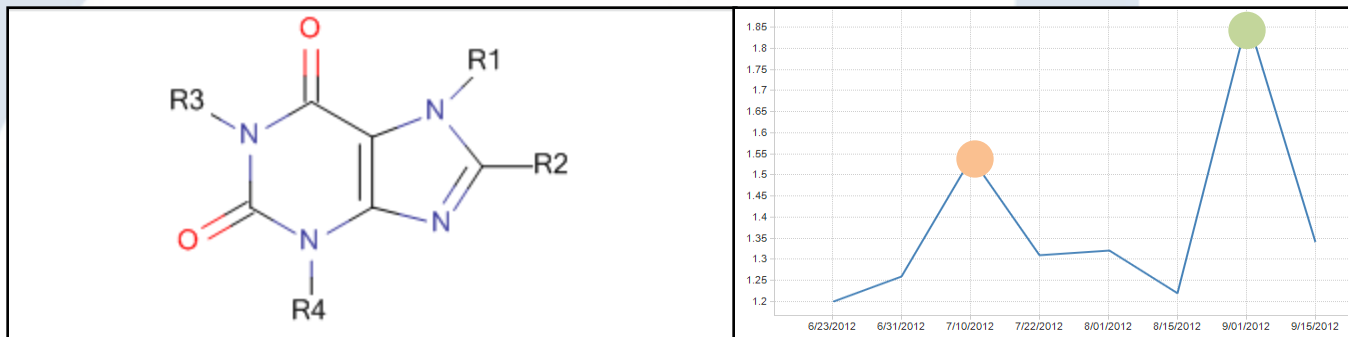
3K 'Active' Compounds

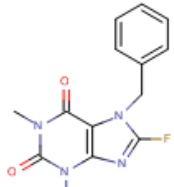
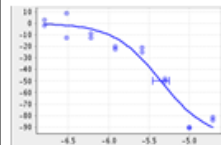
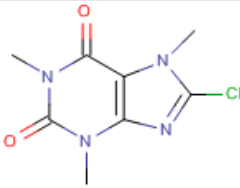
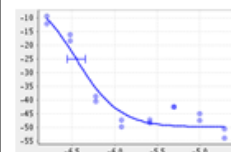
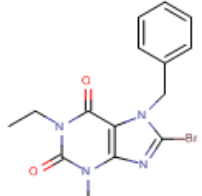
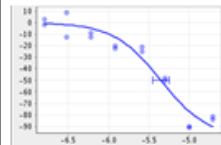
2K Compounds
Selected

750 Active
on Retest

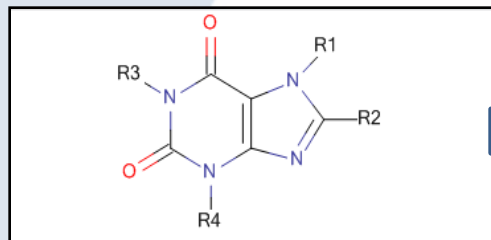
3-4 Lead
Scaffolds

Beyond the Funnel



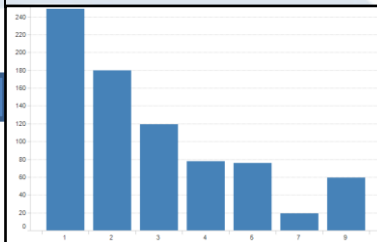
Compound	Date Screened	IC50 uM Calculated	IC50 uM Normalized	Curve Fit	Similarity to Top Performer
	9/1/2012	2.5	2.25		0.82
	7/22/2012	1.5	1.2		0.74
	9/1/2012	0.78	0.67		0.92

Scaffold Timelines

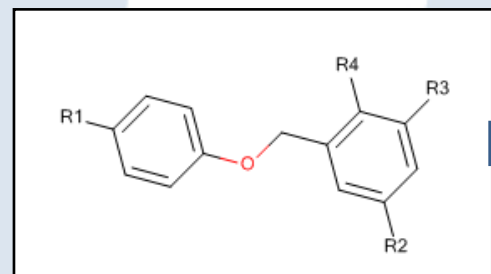


6/23/2011-8/2/2011

39 Compounds

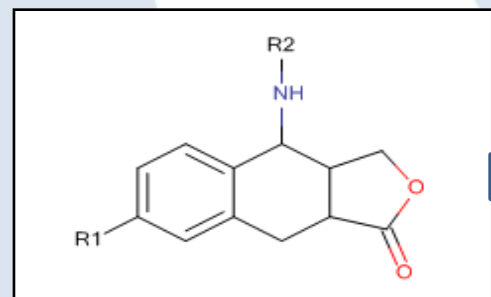
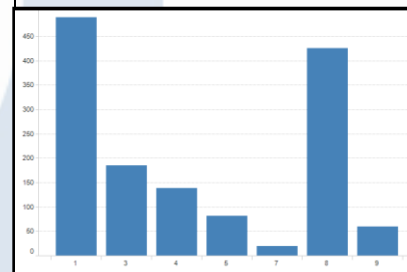


Ran into Solubility issues



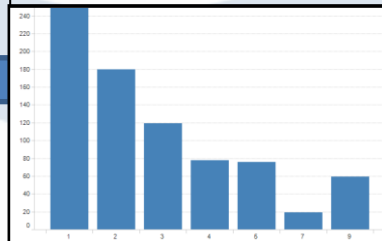
6/23/2011-11/2/2011

87 Compounds



6/23/2011-7/22/2011

19 Compounds



Activity not increasing with R group substitutions

Visualizations

- Will support Scatterplot, Histogram, Line, Heatmap, Pie, Dendogram Initially
- SVG - Javascript – Data is Bound to Chart
 - Interactive Graphs
 - Can add Animation (such as Time Element
- Will Support Live Filtering
- Publication Quality Graphs

Special Thanks

- Chemaxon
 - Jon Patterson, Douglas Drake and Alex Allardyce
 - Chemaxon Support Team
- Broad Institute
 - Informatics Team

