COLLABORATIVE DRUG DISCOVERY

A Collaborative Platform to Drive Drug Repurposing

Barry Bunin, Ph.D.

CEO and President

Collaborative Drug Discovery, Burlingame, CA.
www.collaborativedrug.com

CDD
COLLABORATIVE DRUG DISCOVERY

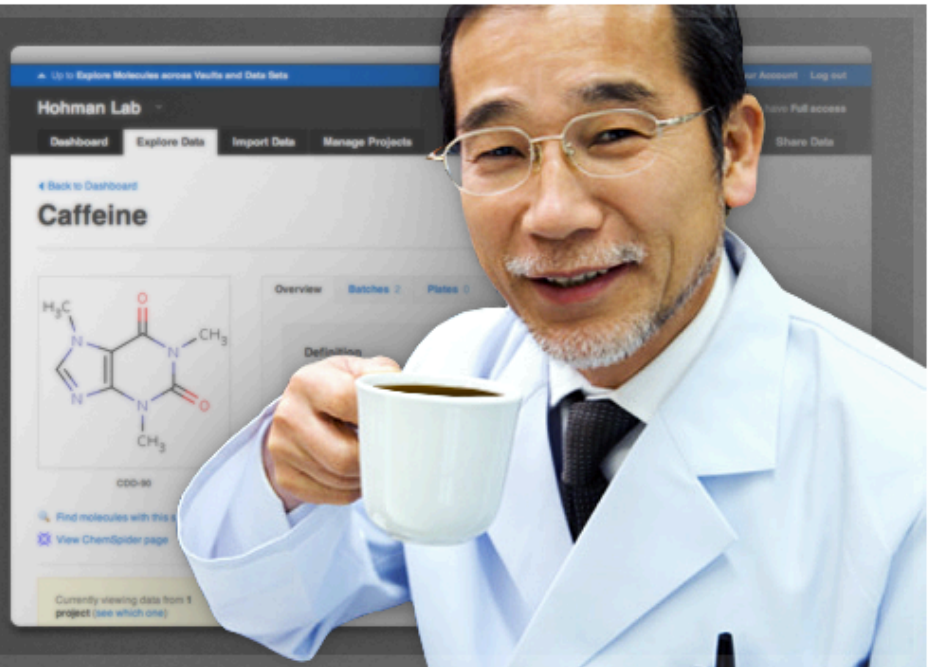
Log in to CDD Vault

Home | Features | News & Events | About Us | **Sign Up**

Streamline drug discovery with CDD's collaborative, web-based software.

CDD Vault is a web application for intelligent data management and secure collaboration. CDD makes drug discovery easier and more efficient for thousands of industry and academic scientists around the world.

Learn more | Sign up now



Simple and secure data management

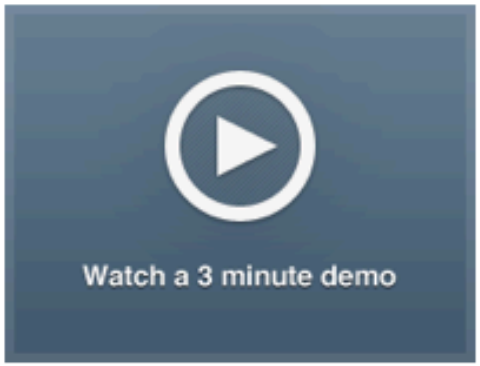
Securely store and easily mine experimental data, including bioassays and chemical structures, in a private data vault that CDD hosts for your group.

Learn more about managing your data with CDD Vault

Collaborate securely with your partners

Not everyone wants to share data, but if you do, you retain full control over who can access which data

Learn more about collaborating with CDD Vault



"One of the biggest barriers for academic drug discovery is the poor access to

Our Users

CDD has announced collaborations with the leading pharmas, biotech companies, and academic institutions.

Industry/Nonprofit Groups

AstraZeneca plc
GlaxoSmithKline
Novartis
Pfizer, Inc.
Sanofi-Aventis S.A.
Acelot, Inc.
Acetylon Pharmaceuticals, Inc.
Algomedix, Inc.
ASINEX
AsisChem, Inc.
BioSeek, Inc.
Cedars-Sinai Medical Center
ChemBridge Corporation
Drugs for Neglected Diseases initiative
Fred Hutchinson Cancer Research Center
Indel Therapeutics, Inc.
LITMUS, LLC
Longevica Pharmaceuticals, Inc.
Marine Biology Laboratory
Melior Discovery, Inc.
Myelin Repair Foundation
Numerate, Inc.
PharmSelex
San Francisco VA Medical Center

Academic Groups

Broad Institute
Columbia University
Consejo Superior de Investigaciones Cientificas (CSIC), Instituto de Parasitologia y Biomedicina “Lopez-Neyra”, Spain
Cornell University
Harvard University
Indiana University
Indiana University-Purdue University Indianapolis
Institut Pasteur de Montevideo, Uruguay
Johns Hopkins University
Louisiana State University-Health Sciences Center
Massachusetts Institute of Technology
MOLISA GmbH, Germany
NIMH Psychoactive Drug Screening Program
Northeastern University
Northwestern University
Ohio State University
Public Library of Science
Purdue University
The Rockefeller University
Scripps Research Institute
Stanford University
UCSF General Hospital
University of California, Berkeley
University of California, Davis
University of California, Los Angeles
University of California, San Diego
University of California, San Francisco
University of California, Santa Cruz

CDD Vault

CDD COLLABORATIVE DRUG DISCOVERY

Search – Collaborative Drug Discovery

collaborativedrug.com https://app.collaborativedrug.com/vaults/256/search

Google

Getting Started | Latest Headlines | Collaborative Drug ... | CDD Google Apps | Collaborative Drug ... | Salesforce.com – sal... | Apple | Yahoo! | Google Maps | YouTube | Wikipedi

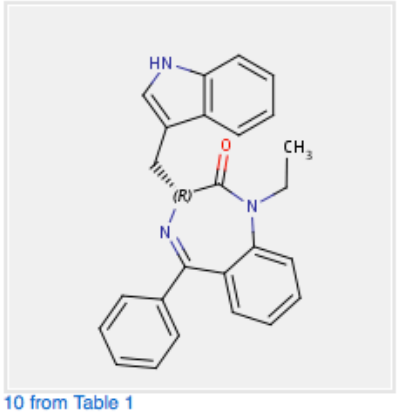
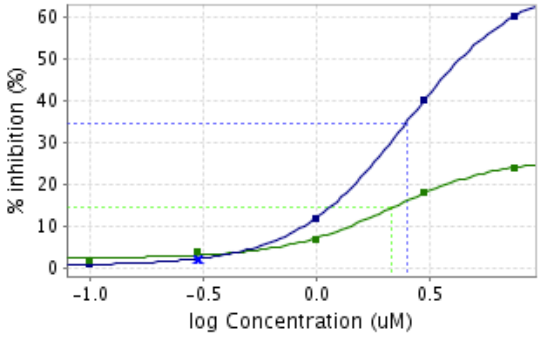
h – Collaborative Drug Disco...

17 structures · Show 100 | Change display options | Plot results | Export results | Add results to project

File format:

- (•) XLS - Excel
- ( ) CSV - Comma-Separated Values
- ( ) SDF - Structure-Data File

Export or cancel

Molecule	CCK-A Dose Response Curve			
	IC50 (uM)	Hill slope	R squared	Dose-re...e Plot
flag outliers  10 from Table 1 Barry Bunin's Vault	2.15	1.82	0.997	 Mark outliers
	2.51	1.76	1.0	

b.collaborativedrug.com/vaults/256/searches/98034#

Dial +1 (636) 277-0135

Collaborative Capabilities is the Prerequisite for More Efficient Drug Discovery

Individual Pain Points	Organizational Need	Technical Solution
Keep scientist's data private until ready to share (i.e. with the rest of their groups)	Collaboration is as much about more efficient work within a single organization as between organizations.	Temporal data access controls
Allow scientists to partition or sequester data to enable selective data sharing	Control of chemical or biological data access based on Project-specific permissions	Spatial data partitioning
Allow collaboration in scientist's natural workflow (i.e. without uploading data multiple times in multiple places)	Control based on type of data – especially given a marketplace with specialization in synthesis, screening, discovery, development	Sequester by data classification

COLLABORATIVE DRUG DISCOVERY

"Off the Shelf R&D"

DAMIEN HIRST
B.1965
PHARMACEUTICALS

✓All pharmas have assets on shelf that reached clinic

✓Get others to help in repurposing / repositioning assets

✓How can CDD help?

- Support Collaborations

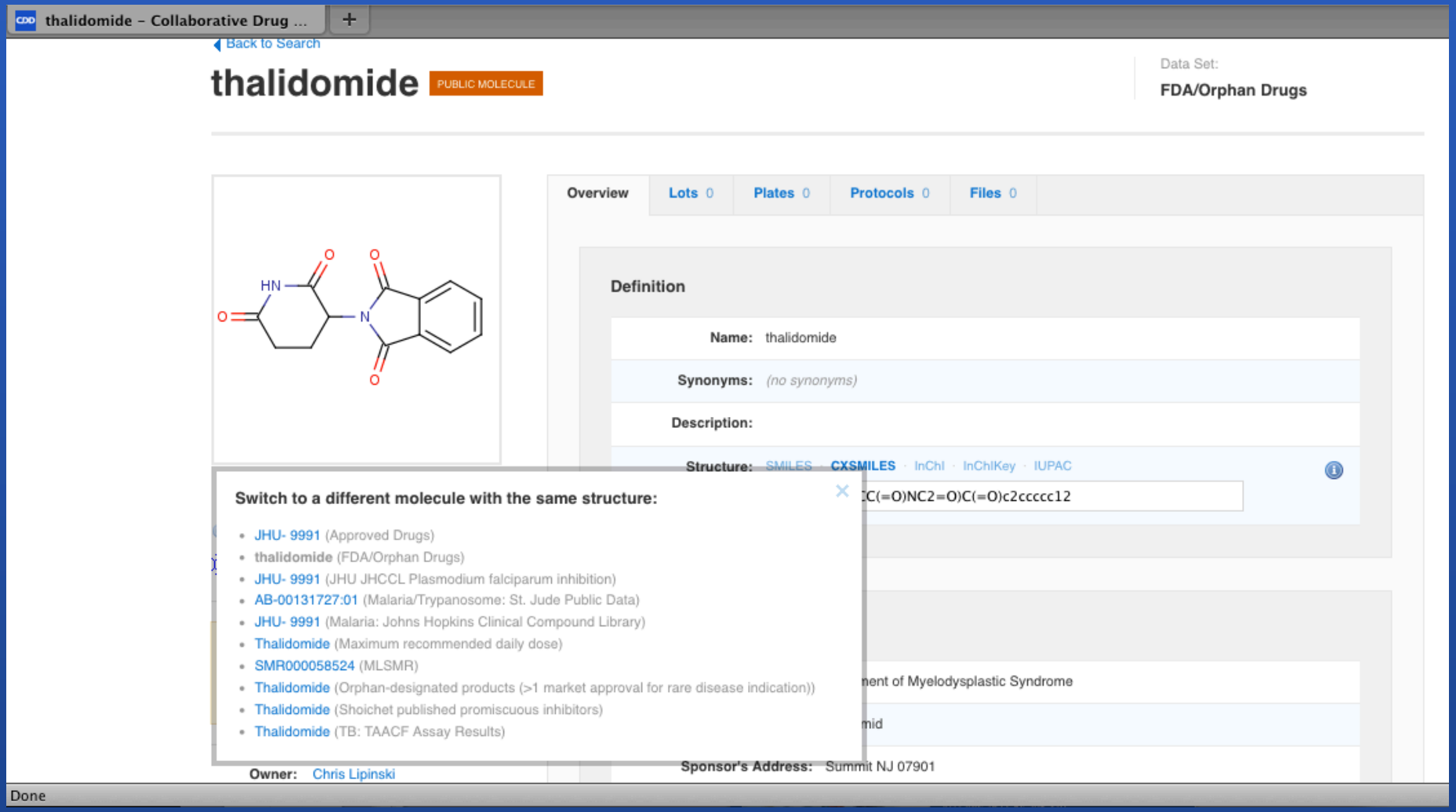
- Provide informatics tools

- Data storage securely on cloud

- Collaborators with appropriate access permissions

Reference: "*In silico* repositioning of approved drugs for rare and neglected diseases" Ekins S, Williams AJ, Krasowski MD, Freundlich JS. *Drug Discovery Today.* 2011 Apr; 16 (7-8): 298-310.

Thalidomide in CDD Public



See: https://www.collaborativedrug.com/blog

ELSEVIER

In silico repositioning of approved drugs for rare and neglected diseases

Sean Ekins[1,2,3,4], **Antony J. Williams**[5], **Matthew D. Krasowski**[6] and **Joel S. Freundlich**[7]

[1] Collaborations in Chemistry, 601 Runnymede Avenue, Jenkintown, PA 19046, USA
[2] Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010, USA
[3] Department of Pharmaceutical Sciences, University of Maryland, College Park, MD 21201, USA
[4] Department of Pharmacology, University of Medicine & Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854, USA
[5] Royal Society of Chemistry, 904 Tamaras Circle, Wake Forest, NC 27587, USA
[6] Department of Pathology, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, USA

Sean Ekins
Sean Ekins is Principal Consultant for Collaborations in Chemistry and Collaborations Director at Collaborative Drug Discovery Inc. He has written over 160 papers and book chapters on topics including drug–drug interaction screening, computational absorption, distribution, metabolism, and excretion (ADME)/Tox, collaborative computational technologies and neglected

Experimental Validation of Drug Repositioning Predictions

DDT 2011 repurposing in silico.pdf

TABLE 1

Examples of approved drug molecules identified using low-throughput screening methods as having effects against diseases other than the original target[a]

Molecule	Original use	New use	Method of discovery	Refs
Aprepitant	Nausea: NK-1 receptor antagonist	Drug-resistant HIV-1 infection: downregulates CCR5 in macrophages Cryptosporidiosis in immunosuppressed hosts	Initial hypothesis tested with another NK-1 receptor antagonist *in vitro* Tested *in vivo* in immunosuppressed mice infected with *Cryptosporidium parvum*; decreased substance P levels	[99,100] [101]
Amiodarone	Class III anti-arrhythmic	Chagas disease: blocks ergosterol biosynthesis	Literature search	[102]
Glybenclamide	Antidiabetic	Antithrombotic activity in mouse models IC_{50} 9.6 μM	Common pharmacophore with an experimental TP receptor antagonist SQ29,548	[103]
Tamoxifen	Antiestrogen	Anti-protozoal: *Leishmania amazonensis* IC_{50} 11.1–16.4 μM	Focused screening to test hypothesis and *in vivo* mice studies	[104,105]
Trimetrexate	Antifolate used in *Pneumocystis carinii* infection in patients with AIDS	Inhibitor of *Trypanosoma cruzi* DHFR IC_{50} 6.6 nM	Enzyme activity and antiparasite activity assays for one compound	[106]
Riluzole	Amyotrophic lateral sclerosis: inhibits glutamate release and reuptake	Currently in clinical trials for treating melanoma, but might have activity against other cancers	Treatment of GRM1-positive human melanoma cells reduced levels of released glutamate, suppressed melanoma cell growth and also suppressed tumor growth in xenograft model; induced cell cycle arrest, leading to apoptosis	[107]
Sertraline	Antidepressant (selective serotonin reuptake inhibitor)	Neuroprotective, prolongs survival, improves motor performance and ameliorates brain atrophy in the R6/2 HD model	Previously shown that another SSRI was neuroprotective	[108]

Experimental Validation of Drug Repositioning Predictions

DDT 2011 repurposing in silico.pdf

Examples of approved drug molecules identified using HTS or *in silico* screening methods as having effects against diseases other than original target[a]

Molecule	Original use	New use	Method of discovery	Refs
Itraconazole	Antifungal: lanosterol 14α-demethylase inhibitor	Inhibition of angiogenesis by inhibiting human lanosterol 14α-demethylase; IC_{50} 160 nM	*In vitro* HUVEC proliferation screen against FDA-approved drugs (JHCCL)	[109]
Astemizole	Non-sedating antihistamine (removed from US market by FDA in 1999)	Antimalarial IC_{50} 227 nM against *Plasmodium falciparum* 3D7	*In vitro* screen for *P. falciparum* growth of 1937 FDA-approved drugs (JHCCL)	[110]
Mycophenolic acid	Immunosuppressive drug: inhibits guanine nucleotide biosynthesis	Inhibition of angiogenesis by targeting type 1 inosine monophosphate dehydrogenase; IC_{50} 99.2 nM	*In vitro* HUVEC proliferation screen of 2450 FDA- and foreign-approved drugs (JHCCL)	[111]
Entacapone and tolcapone	Parkinson's Disease: catechol-*O*-methyltransferase inhibitors	Antitubercular: entacapone inhibits InhA; IC_{50} 80 μM	Used a chemical systems biology approach	[77]
Nitazoxanide	Infections caused by *Giardia* and *Cryptosporidium* spp.	Antitubercular: multiple potential targets	Screens against replicating and non-replicating Mtb	[112]
(±)-2-amino-3-phosphonopropionic acid	Human metabolite, mGluR agonist	Antimalarial: inhibits HSP-90; IC_{50} 0.06 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Acrisorcin	Antifungal	Antimalarial: inhibits HSP-90; IC_{50} 0.05 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Harmine	Anticancer	Antimalarial: inhibits HSP-90; IC_{50} 0.05 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Acetophenazine, fluphenazine and periciazine	Antipsychotics–D2 and 5-HT_2 inhibitors	Human androgen receptor antagonists acetophenazine (K_i 0.8 μM), fluphenazine(K_i 0.8 μM), periciazine (K_i 3.0 μM)	Docking of known drugs into androgen receptor followed by *in vitro* screening	[96]
Levofloxacin, gatifloxacin,	DNA gyrase	Active against ATCC17978; inactive	Screening of 1040 drugs from	[114]

Repositioning Using CDD + Public Data

DDT 2011 repurposing in silico.pdf

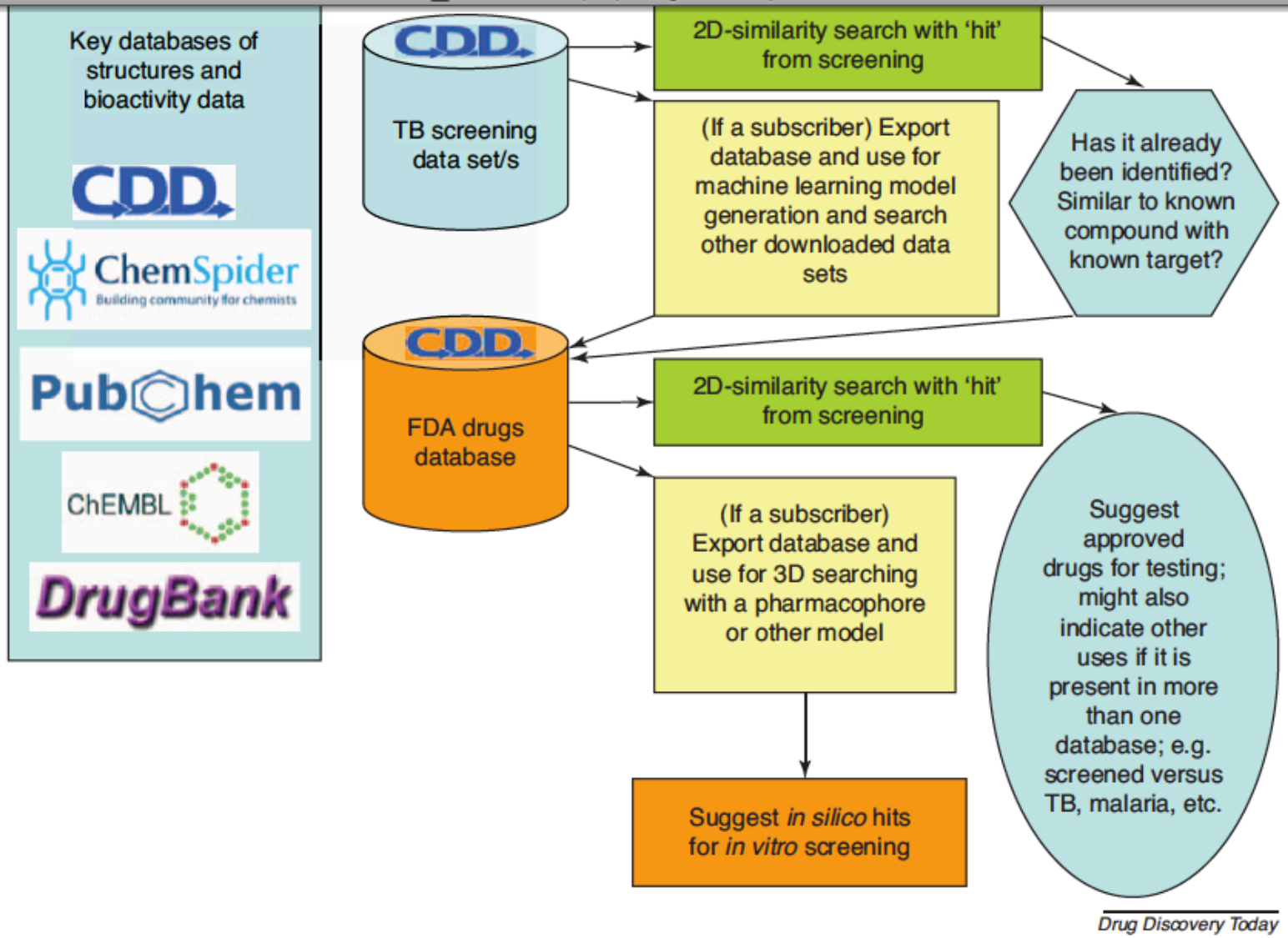


FIG. 3

A repositioning strategy using the CDD (http://www.collaborativedrug.com/register), ChemSpider (http://www.chemspider.com) or other databases (bioactivity dat

CDD Public Data Sets for Drug Repositioning

CDD COLLABORATIVE DRUG DISCOVERY

SHOW HIDE	**Enamine Representative Diverse Screening Library**	Enamine	Dmytro Mykytenko	200000	4/12/2011

Original 200K diverse screening library was generated especially for Collaborative Drug Discovery users from the world's largest stock of commercially available screening compounds (over 1.7 M species). The library features exclusive drug-like compounds with refined ADME properties. Our high quality compounds can be cherry-picked and supplied immediately in different formats.

SHOW HIDE	**FDA Drugs Repurposed using HTS methods**	In vitro repurposing		109	3/18/2011

Drugs identified with new uses using HTS methods. This table greatly extends a previously published version "Ekins S, Williams AJ, Krasowski MD, Freundlich JS. In silico repositioning of approved drugs for rare and neglected diseases Drug Discov Today. 2011 Mar 1. PMID: 21376136 doi:10.1016/j.drudis.2011.02.016 The table lists molecules, Old use / target, new use/ target, how discovered and references. Abbreviations: CCR5, Chemokine receptor 5; DHFR, Dihydrofolate reductase; DOA, Drugs of abuse, FDA, Food and Drug Administration; GLT1, Glutamate transporter 1; HSP-90, Heat shock protein 90; JHCCL, John Hopkins Clinical Compound Library; Mtb, Mycobacterium tuberculosis; NK-1, neurokinin- 1 receptor; OCTN2.

SHOW HIDE	**Orphan-designated products (>1 market approval for rare disease indication)**	Rare dise...urposing		76	3/18/2011

FDA Table 2 - from the FDA resource, the rare disease research database (RDRD), which lists Orphan-designated products (http://www.fda.gov /ForIndustry/DevelopingProductsforRareDiseasesConditions/HowtoapplyforOrphanProductDesignation/ucm216147.htm) with at least one marketing approval for a rare disease indication. This data was analyzed by Ekins and Williams (paper submitted).

SHOW HIDE	**Orphan-designated products (>1 market approval for common disease indication)**	Rare dise...urposing		105	3/18/2011

FDA Table 1 - from the FDA resource, the rare disease research database (RDRD), which lists Orphan-designated products (http://www.fda.gov /ForIndustry/DevelopingProductsforRareDiseasesConditions/HowtoapplyforOrphanProductDesignation/ucm216147.htm) with at least one marketing approval for a common disease indication. The FDA did not associate the data with molecule structures.

	Colorado Center for Drug Discovery	C2D2 Public	Greg Miknis, Associate	2240	3/7/2011

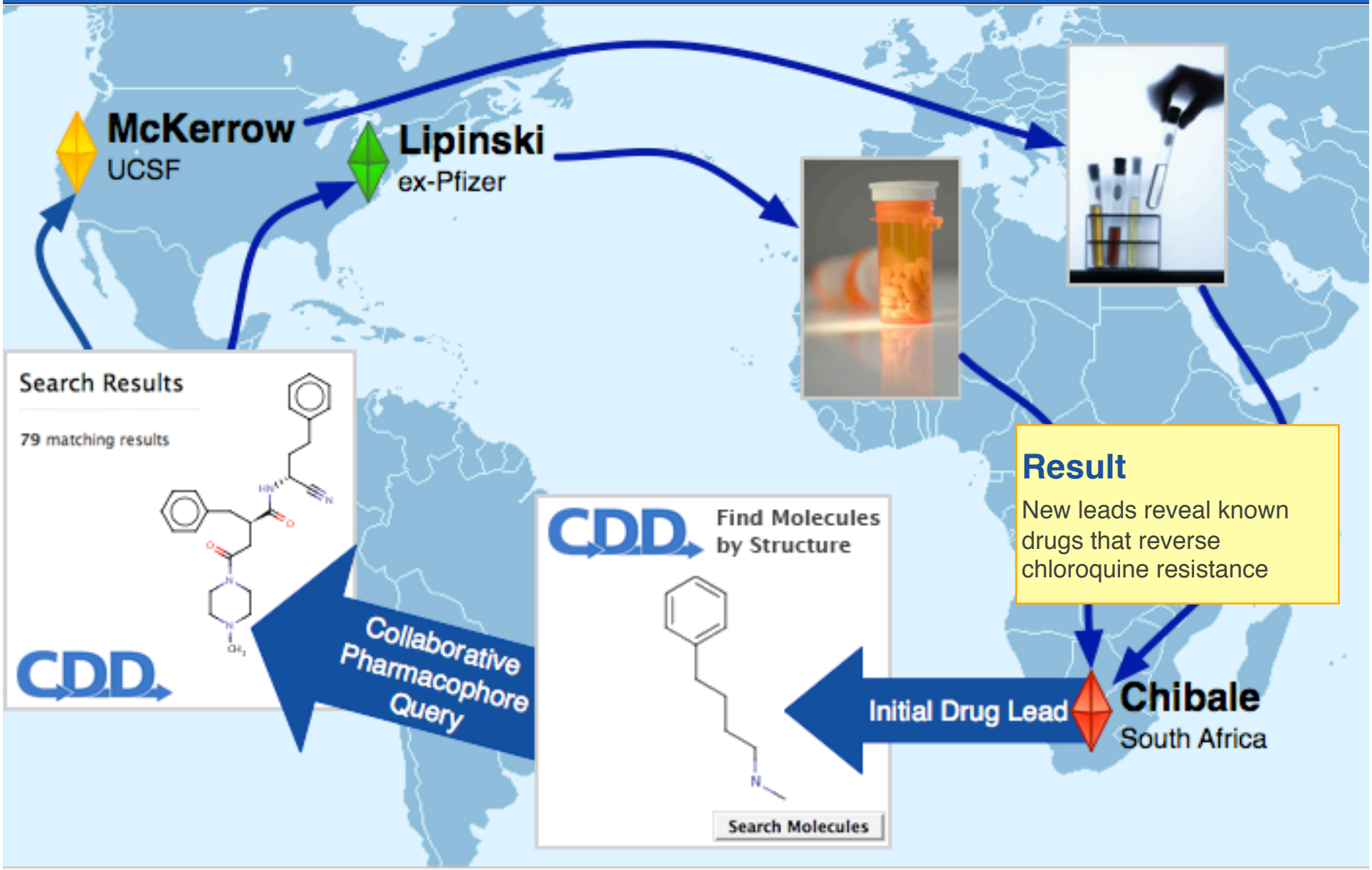
Unique Collaborative Platform for Drug Repositioning

- Network
  - Traction: thousands of leading researchers log into CDD today:
  - Academic customers: Harvard, Columbia, Johns Hopkins, UCSF (new assays)
  - Pharmas relationships: Pfizer, GSK, Novartis, Lilly (commercial partners)
  - Startups
  - Research institutes, Non profits NIH, BMGF, MM4TB etc

- Neutral
  - Trusted for >7 years in the cloud
  - Moral high-ground due to years dedicated to neglected disease
  - Credible position

- IP
  - CDD handles data corresponding to composition of matter & utility patents
  - Templates for rapid web-based transactions (IP corresponding to data)
  - CDD does not own IP

CDD Drug Repurposing 2009 Case Study



CDD 2011 EU PPP Case Study

COLLABORATIVE DRUG DISCOVERY

refox File Edit View History Bookmarks Tools Window Help (Charged) Fri 1:32 PM

Welcome to the temporary website of MM4TB

http://www.mm4tb.org/

Google

Getting Started Latest Headlines CDD Google Apps Collaborative Drug ... Salesforce.com – sal... Apple Yahoo! Google Maps YouTube Wikipedia Recent Changes – C... News

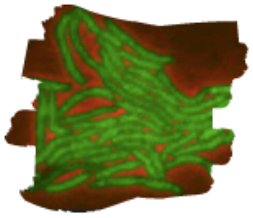
ne to the temporary websit...

MM4TB

More Medicines for Tuberculosis

Welcome to the mm4tb website

This is a preliminary version of the website, where you will find basic information about our project.
The full version of the website will be ready soon, please visit us again in the near future to learn more!



The More Medicines for Tuberculosis (MM4TB) consortium evolved from the highly successful FP6 project, New Medicines for TB (NM4TB), that delivered a candidate drug for clinical development two years ahead of schedule. Building on these firm foundations and exploiting its proprietary pharmacophores, MM4TB will continue to develop new drugs for TB treatment. An integrated approach will be implemented by a multidisciplinary team that combines some of Europe's leading academic TB researchers with two major pharmaceutical companies and four SMEs, all strongly committed to the discovery of

Project Partners

- École Polytechnique Fédérale de Lausanne, Switzerland
- Uppsala University, Sweden
- University of Cambridge, United Kingdom
- Institut Pasteur, France
- Università degli Studi di Padova, Italy
- Vichem Chemie Research Ltd., Hungary
- Indian Institute of Science, India
- Università degli Studi Piemonte Orientale "A. Avogadro", Italy
- Tydock Pharma, Italy
- Eidgenössische Technische Hochschule Zürich, Switzerland
- Astra Zeneca R&D, India
- Università degli Studi di ... Italy
- Barts & the London Queen Mary's School of Medicine & Dentistry, University of London, United Kingdom
- A. N. Bakh Institute of Biochemistry of the Russian Academy of Science, Russian Federation
- Comenius University, Slovakia
- John Innes Centre, Norwick, United Kingdom
- Cellworks, Bangalore, India
- Collaborative Drug Discovery, USA
- Universidad del Pais Vasco/ Euskal Herriko Unibertsitatea,

Email: bbunin@collaborativedrug.com

Drug Toxicity Alerts

CDD COLLABORATIVE DRUG DISCOVERY

Caffeine – Collaborative Drug Discovery

file:///Users/moses/code/design-svn/templates/molecule.html

Google

CDD · Lipinski Group

Chris Lipinski: Your Account · Help · Log out

Dashboard | Group Data | Import Data | Share Data

Currently using **Lipinski Group** data and **6 shared data sets** Choose data sets ...

Back to Dashboard

CAS 957217-65-1 PUBLIC MOLECULE

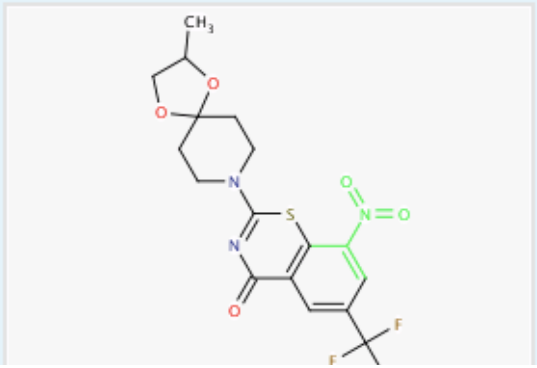
Available in 3 data sets. Now viewing:

TB MIC Prathipati NIAID – 37532

Overview | Batches 2 | Plates 1 | Protocols 3 | Associated Files 2 | Alerts 2

2 Alerts

Structure Match Preview



Alert

Abbott Alarm NMR failed

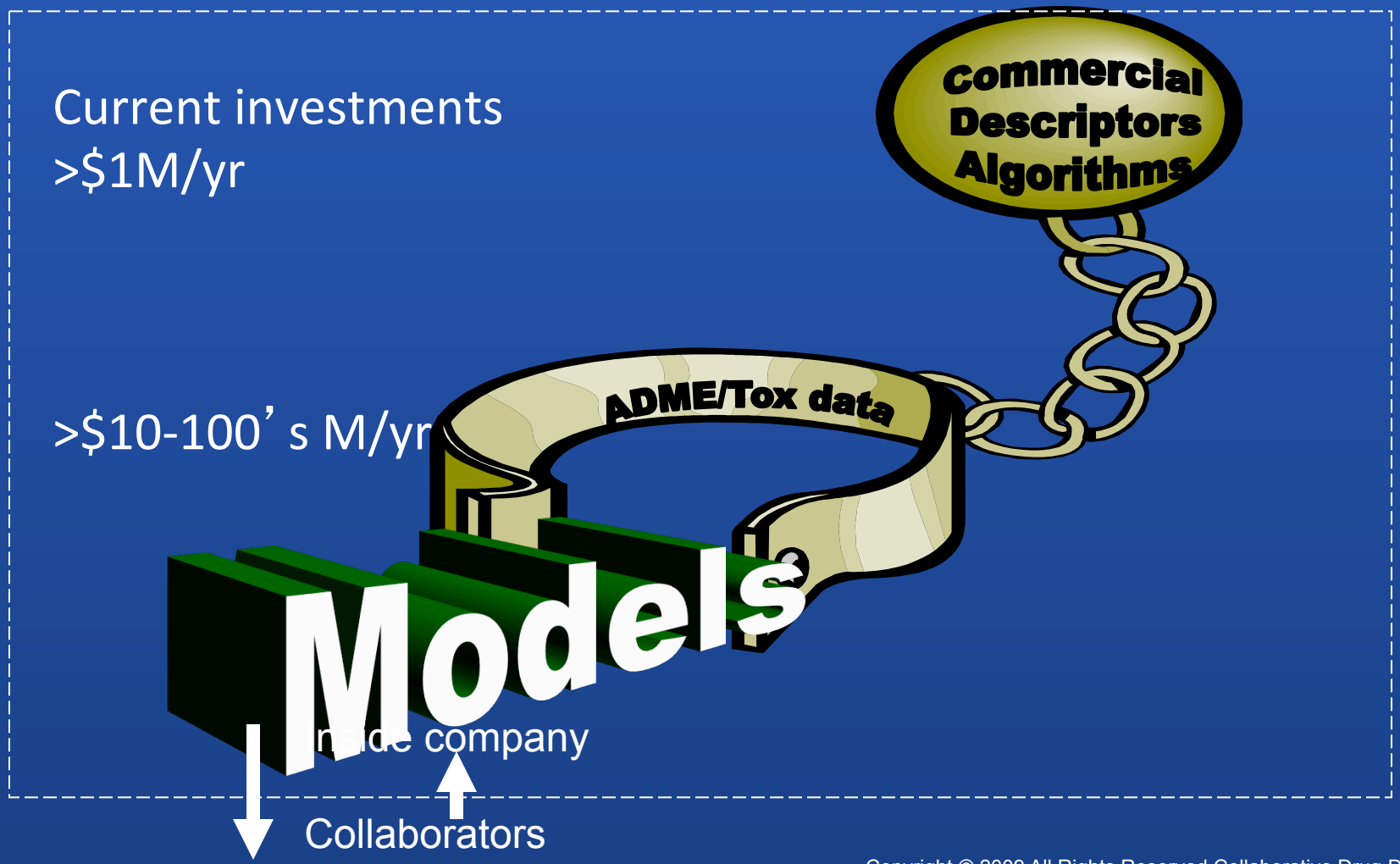
- csc
- Nc1cccc2aaaaa12

Pfizer Lint failed

- O~N(=O)-c(:*):* (aromatic NO2)

What if you could…

1. Spend only 20% on descriptors and algorithms
2. Selectively share your models with collaborators and control access
3. Have someone else host the models / predictions



Yes we can…predictions now can be shared without revealing structures!

Using Open Source Descriptors and Algorithms for Modeling ADME Properties

Rishi R. Gupta†, Eric M. Gifford†, Ted Liston†, Chris L. Waller†, Moses Hohman*, Barry A. Bunin* and Sean Ekins*

†Pfizer Global Research and Development, Eastern Point Road, Groton, CT 06340

**Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010*

ABSTRACT

Aim: Computational models could be more readily shared with collaborators if they were generated with open source descriptors (e.g. Chemistry development kit, CDK) and modeling algorithms.

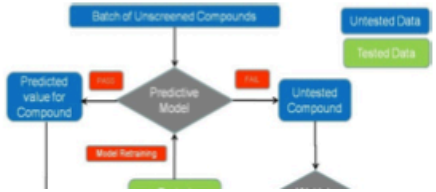
Method and Results: We evaluated open source descriptors and model building algorithms using a training set of ~50K molecules and a test set of ~25K molecules with human liver microsomal metabolic stability (HLM) data. A C5.0 decision tree model demonstrated that open CDK+SMARTS keys (Kappa = 0.43, sensitivity = 0.57, specificity 0.91, positive predicted value (PPV) = 0.64) are equivalent to models built with commercial MOE2D+SMARTS keys (Kappa = 0.43, sensitivity = 0.58, specificity 0.91, PPV = 0.63). Extending the dataset to ~ 200K molecules confirmed this observation. The same combination of descriptor set and modeling method was applied to a variety of other ADME endpoints such as solubility etc. and the results were encouraging.

Conclusion: Open source descriptors and algorithms demonstrated comparable results to commercial descriptors with cost savings.

INTRODUCTION

Evaluation of lead compounds for drug-like properties (e.g. ADME/TOX) very early on in the discovery process using computational methods can alert for late stage failures. The Active Learning paradigm where computational models can be utilized to make decisions to screen every compound or not (Figure 1), can be used. What is missing are ways to share computational models between collaborators.

Figure 1. This schematic describes the "*in silico* screening" methodology where a predictive *in silico* model can be used to pre-screen the compounds before they can go through the actual *in vitro* screening.



METHOD

Datasets:

Human Liver Microsomal Stability (HLM) data on ~200K compounds.

Compounds were synthesized and tested in the HLM assay at Pfizer.

Datasets were binned as per the guidance provided by experts in the Pharmacokinetics, Dynamics and Metabolism (PDM) business unit (Table1).

Datasets:

- Data was binned in 3 bins as shown in Table 1.
- A 3 bin classification model as well as a continuous model on the full dataset was built.
- The distribution of the data in each class in this and the other datasets is shown in Table 2.

Table 1. Classification bins for HLM assay

	Low Risk (Low)	Moderate Risk (Moderate)	High Risk (High)
HLM	$CL_{int} < 9.2$	$9.2 < CL_{int} < 48$	$CL_{int} > 48$

Table 2. Dataset size and data distribution for the HLM assay that were used to build the categorical models is shown in this table.

Descriptors:

- Pfizer modified Molecular Operating Environment 2D set (MOE2D, 2008) (463 descriptors).
- CDK (http://cdk.sourceforge.net/) (Steinbeck et al., 2006) fingerprints (195 descriptors).
- SMARTS keys (166 Keys)

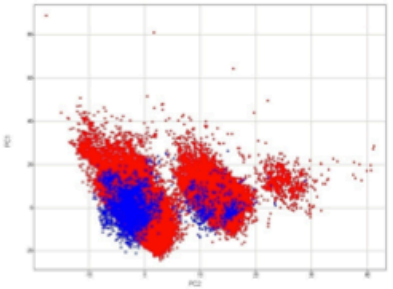
Table 3. Details on the HLM data modeling grid showing performance of various modeling methods versus descriptors.

	SVM	RP Forest Uni Class	RP Forest	Rulequest Cubist C5.0
CDK	Kappa = 0.14 Sensitivity = 0.11 Specificity = 0.96 PPV = 0.43	Kappa = 0.16 Sensitivity = 0.54 Specificity = 0.70 PPV = 0.33	Kappa = 0.11 Sensitivity = 0.85 Specificity = 0.33 PPV = 0.25	Kappa = 0.39 Sensitivity = 0.54 Specificity = 0.91 PPV = 0.61
MOE2D and SMARTS Keys	Not evaluated	Not evaluated	Not evaluated	Kappa = 0.43 Sensitivity = 0.58 Specificity = 0.91 PPV = 0.63 (Baseline)
CDK and SMARTS Keys	Not evaluated	Not evaluated	Not evaluated	Kappa = 0.43 Sensitivity = 0.58 Specificity = 0.91 PPV = 0.63

Table 4. Summary of results for C5.0 HLM models generated with very large training sets and different molecular descriptors. PPV = positive predicted value.

HLM Model with CDK and SMARTS Keys	HLM Model with MOE2D and SMARTS Keys
# of Descriptors: 578	# of Descriptors: 818
# of Training Set compounds: 193,650	# of Training Set compounds: 193,930
XValidation Results: 38,730 compounds	XValidation Results: 38,786 compounds
Training R^2: **0.79**	Training R^2 : **0.77**
20% Test Set R^2: **0.69**	20% Test Set R^2: **0.69**
Blind Data Set (2310 compounds):	**Blind Data Set (2310 compounds):**
R^2 = **0.53**	R^2 = **0.53**

Figure 3. The chemical space of the HLM training dataset ~193,000 compounds (red circles) and the test set of 2300 compounds (blue circles). First we calculated the PC's for the training set i.e. a matrix of and 579 descriptors and then we calculated the PC's for the test set. The total variance explained by the 3 PCs for the training set was 0.251. The total variance explained by the 3 PCs for the test set was 0.309.



CONCLUSIONS

- We have presented the largest validated metabolic stability model to date.
- We obtained similar results for permeability and efflux *in silico* models built on over 50,000 compounds.
- A new approach using widely available open descriptors and algorithms is suggested which enable sharing of models that do not require commercial software.

REFERENCES

Gupta, R.R., Gifford, E.M., Liston, T., Waller, C.L., Hohman, M., Bunin, B.A. and Ekins, S.; Using Open Source Computational Tools for Predicting Human Metabolic Stability and Additional ADME/Tox Properties, Drug Metabolism and Disposition, 2010 (in press).

Archive, Mine, Collaborate

COLLABORATIVE DRUG DISCOVERY

Summary

- What is implemented in software guides common business rules and creates efficiencies.
- CDD has lowered the barriers to archive and collaborate – instead of frustration after deciding what “makes sense”, collaboration occurs naturally and instantaneously.
- Collaborative informatics provides cohesiveness when scaling collaborative efforts.
- CDD’s unique technology allows groups with diverse IP/data requirements to work together, as if one.