Challenges and Recommendations for Obtaining Chemical Structures of Industry-Provided Repurposing Candidates

Christopher Southan [1], Antony J. Williams[2] and Sean Ekins[3,4]

[1] TW2Informatics, Göteborg, Sweden
[2] Royal Society of Chemistry, Wake Forest, NC.
[3] Collaborations in Chemistry, Fuquay Varina, NC.
[4]Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010, USA.

Why old drugs ? Why repurposing?

Drug Discovery Today • Volume 00, Number 00 • March 2011 REVIEWS

ELSEVIER

Reviews • KEYNOTE REVIEW

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]

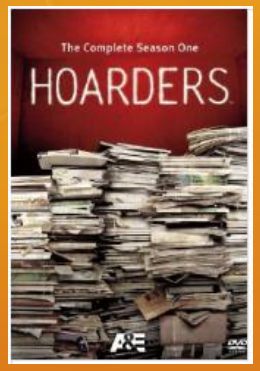
[1] Collaborations in Chemistry, 601 Runnymede Avenue, Jenkintown, PA 19046, USA
[2] Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010, USA
[3] Department of Pharmaceutical Sciences, University of Maryland, College Park, MD 21201, USA
[4] Department of Pharmacology, 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
[7] Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

One approach to speed up drug discovery is to examine new uses for existing approved drugs, so-called 'drug repositioning' or 'drug repurposing', which has become increasingly popular in recent years. Analysis of the literature reveals many examples of US Food and Drug Administration-approved drugs that are active against multiple targets (also termed promiscuity) that can also be used to therapeutic advantage for repositioning for other neglected and rare diseases. Using proof-of-principle examples, we suggest here that with current *in silico* technologies and databases of the structures and biological activities of chemical

- More cost effective R&D?
- Repurposing/ repositioning - Quicker to bring to market?
- Recent focus on neglected & rare diseases
- Over 7000 diseases affecting less than 200,000
- 1000's of diseases with no treatments
- >300 orphan drugs approved since 1983

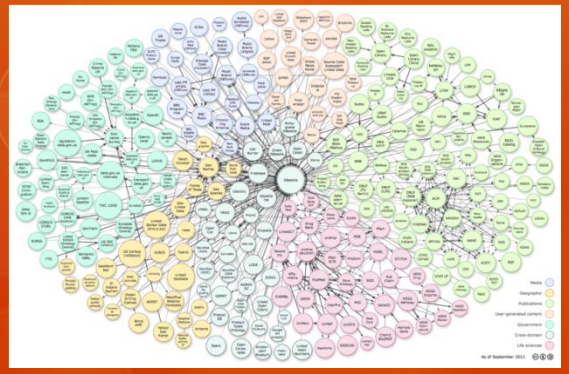
Ekins S, Williams AJ, Krasowski MD and Freundlich JS, Drug Disc Today, 16: 298-310, 201

From data hoarding to open data



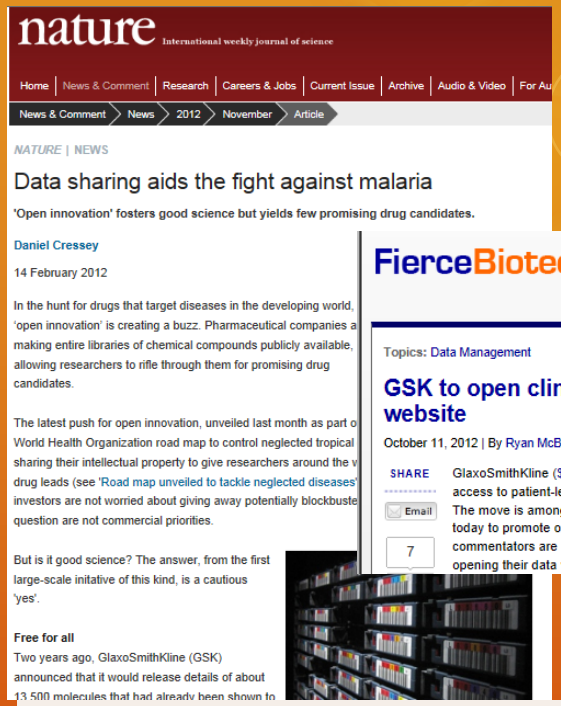
(IMDB)

Me



Linked Open data cloud 2011 (Wikipedia)

Pharma company data hoarding - to open data


nature
international weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Au

News & Comment › News › 2012 › November › Article

NATURE | NEWS

Data sharing aids the fight against malaria

'Open innovation' fosters good science but yields few promising drug candidates.

Daniel Cressey

14 February 2012

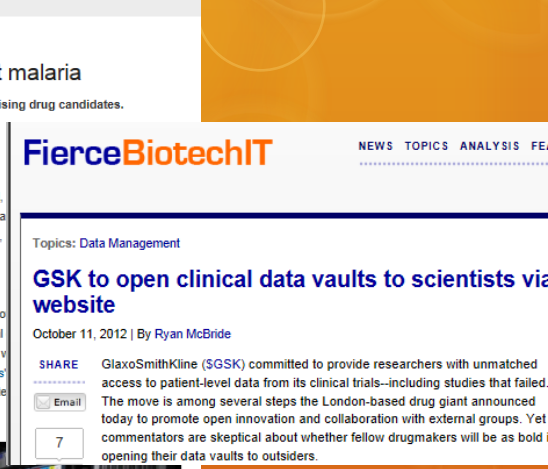
In the hunt for drugs that target diseases in the developing world, 'open innovation' is creating a buzz. Pharmaceutical companies a making entire libraries of chemical compounds publicly available, allowing researchers to rifle through them for promising drug candidates.

The latest push for open innovation, unveiled last month as part o World Health Organization road map to control neglected tropical sharing their intellectual property to give researchers around the v drug leads (see 'Road map unveiled to tackle neglected diseases' investors are not worried about giving away potentially blockbuste question are not commercial priorities.

But is it good science? The answer, from the first large-scale initative of this kind, is a cautious 'yes'.

Free for all

Two years ago, GlaxoSmithKline (GSK) announced that it would release details of about 13,500 molecules that had already been shown to


FierceBiotechIT

NEWS TOPICS ANALYSIS FE

Topics: Data Management

GSK to open clinical data vaults to scientists via website

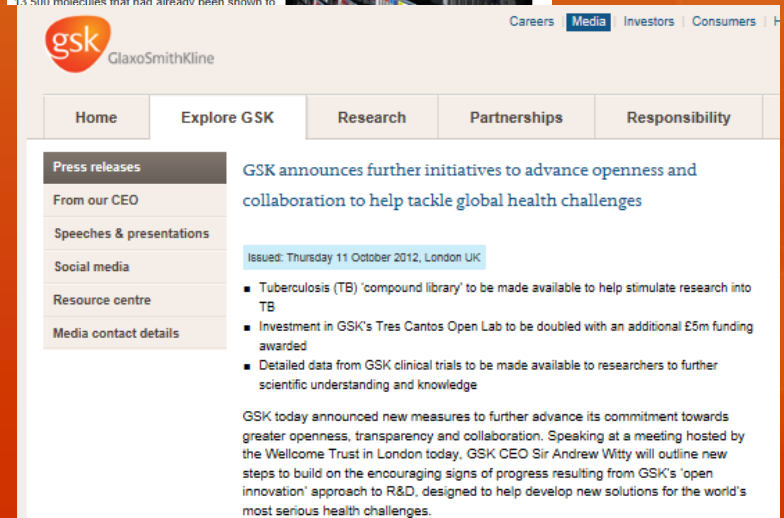
October 11, 2012 | By Ryan McBride

SHARE

Email

7

GlaxoSmithKline ($GSK) committed to provide researchers with unmatched access to patient-level data from its clinical trials--including studies that failed. The move is among several steps the London-based drug giant announced today to promote open innovation and collaboration with external groups. Yet commentators are skeptical about whether fellow drugmakers will be as bold opening their data vaults to outsiders.


gsk GlaxoSmithKline

Careers | Media | Investors | Consumers

Home | Explore GSK | Research | Partnerships | Responsibility

Press releases

From our CEO

Speeches & presentations

Social media

Resource centre

Media contact details

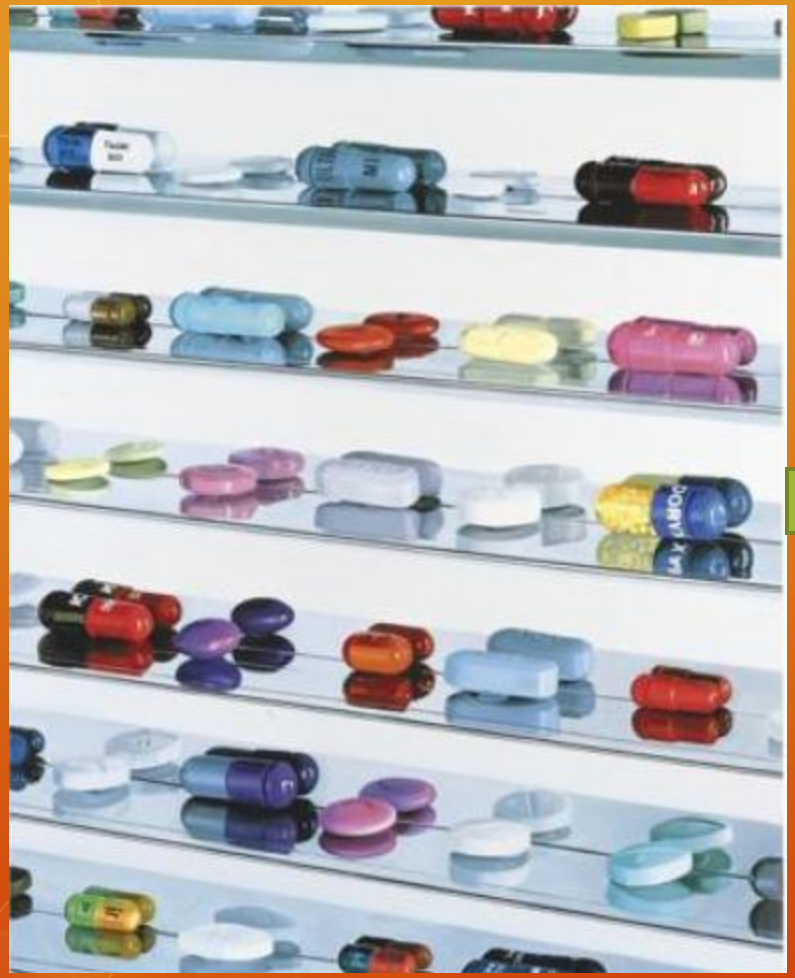
GSK announces further initiatives to advance openness and collaboration to help tackle global health challenges

Issued: Thursday 11 October 2012, London UK

- Tuberculosis (TB) 'compound library' to be made available to help stimulate research into TB
- Investment in GSK's Tres Cantos Open Lab to be doubled with an additional £5m funding awarded
- Detailed data from GSK clinical trials to be made available to researchers to further scientific understanding and knowledge

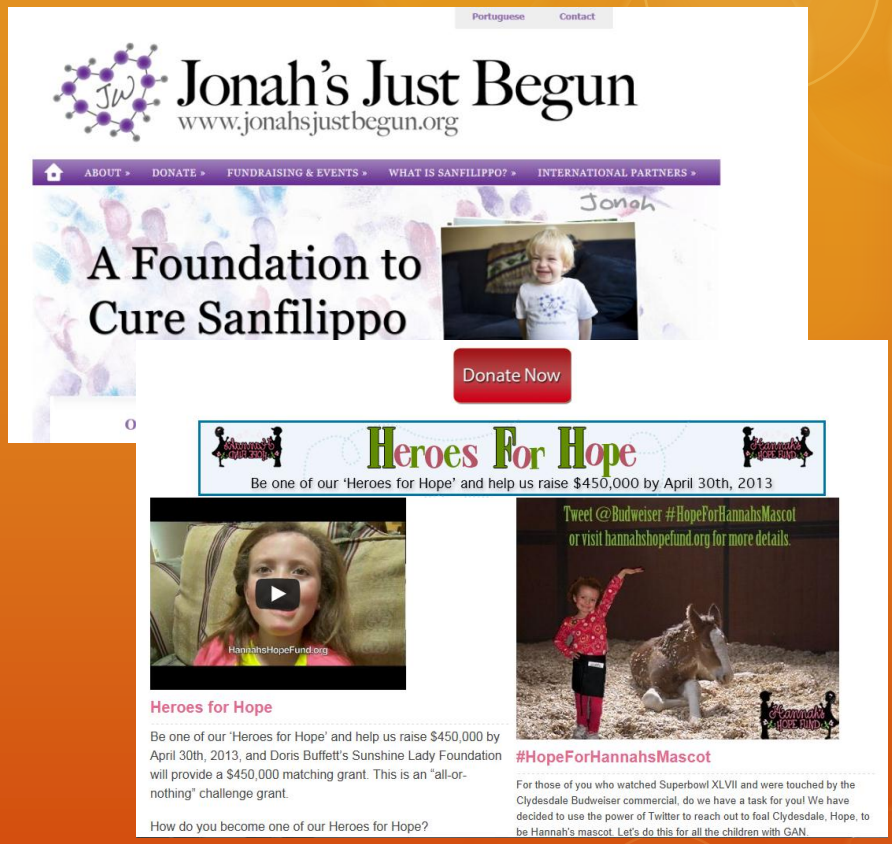
GSK today announced new measures to further advance its commitment towards greater openness, transparency and collaboration. Speaking at a meeting hosted by the Wellcome Trust in London today, GSK CEO Sir Andrew Witty will outline new steps to build on the encouraging signs of progress resulting from GSK's 'open innovation' approach to R&D, designed to help develop new solutions for the world's most serious health challenges.

“Off the Shelf R&D”



DAMIEN HIRST
B.1965
PHARMACEUTICALS

Rare diseases need cures



All pharmas have assets on shelf that reached clinic

Find new uses for these molecules

Get to the patient faster

Finding Promiscuous Old Drugs for New Uses

Antivir Chem Chemother. 2012 Feb 28. doi: 10.3851/IMP2080. [Epub ahead of print]

Inhibition of Influenza A Virus Replication by Antagonism of a PI3K-AKT-mTOR Pathway Member Identified by Gene Trap Insertional Mutagenesis.

Murray JL, McDonald NJ, Sheng J, Shaw MW, Hodge TW, Rubin DH, O'Brien WA, Smee DF.

everolimus

5-fluorouracil

Cancer Cell. 2011 Sep 13;20(3):384-99. doi: 10.1016/j.ccr.2011.08.013.

An integrated in vitro and in vivo high-throughput screen identifies treatment leads for ependymoma.

Atkinson JM, Shelat AA, Carcaboso AM, Kranenburg TA, Arnold LA, Boulos N, Wright K, Johnson RA, Poppleton H, Mohankumar KM, Féau C, Phoenix T, Gibson P, Zhu L, Tong Y, Eden C, Ellison DW, Priebe W, Koul D, Yung WK, Gajjar A, Stewart CF, Guy RK, Gilbertson RJ.

J Invest Dermatol. 2011 Dec;131(12):2467-76. doi: 10.1038/jid.2011.300. Epub 2011 Sep 22.

Rapamycin suppresses self-renewal and vasculogenic potential of stem cells isolated from infantile hemangioma.

Greenberger S, Yuan S, Walsh LA, Boscolo E, Kang KT, Matthews B, Mulliken JB, Bischoff J.

ceftriaxone

Neuro Oncol. 2011 Sep;13(9):974-82. Epub 2011 Jul 15.

Antiparasitic mebendazole shows survival benefit in 2 preclinical models of glioblastoma multiforme.

Bai RY, Staedtke V, Aprhys CM, Gallia GL, Riggins GJ.

J Cell Physiol. 2011 Oct;226(10):2484-93. doi: 10.1002/jcp.22609.

Role of excitatory amino acid transporter-2 (EAAT2) and glutamate in neurodegeneration: opportunities for developing novel therapeutics.

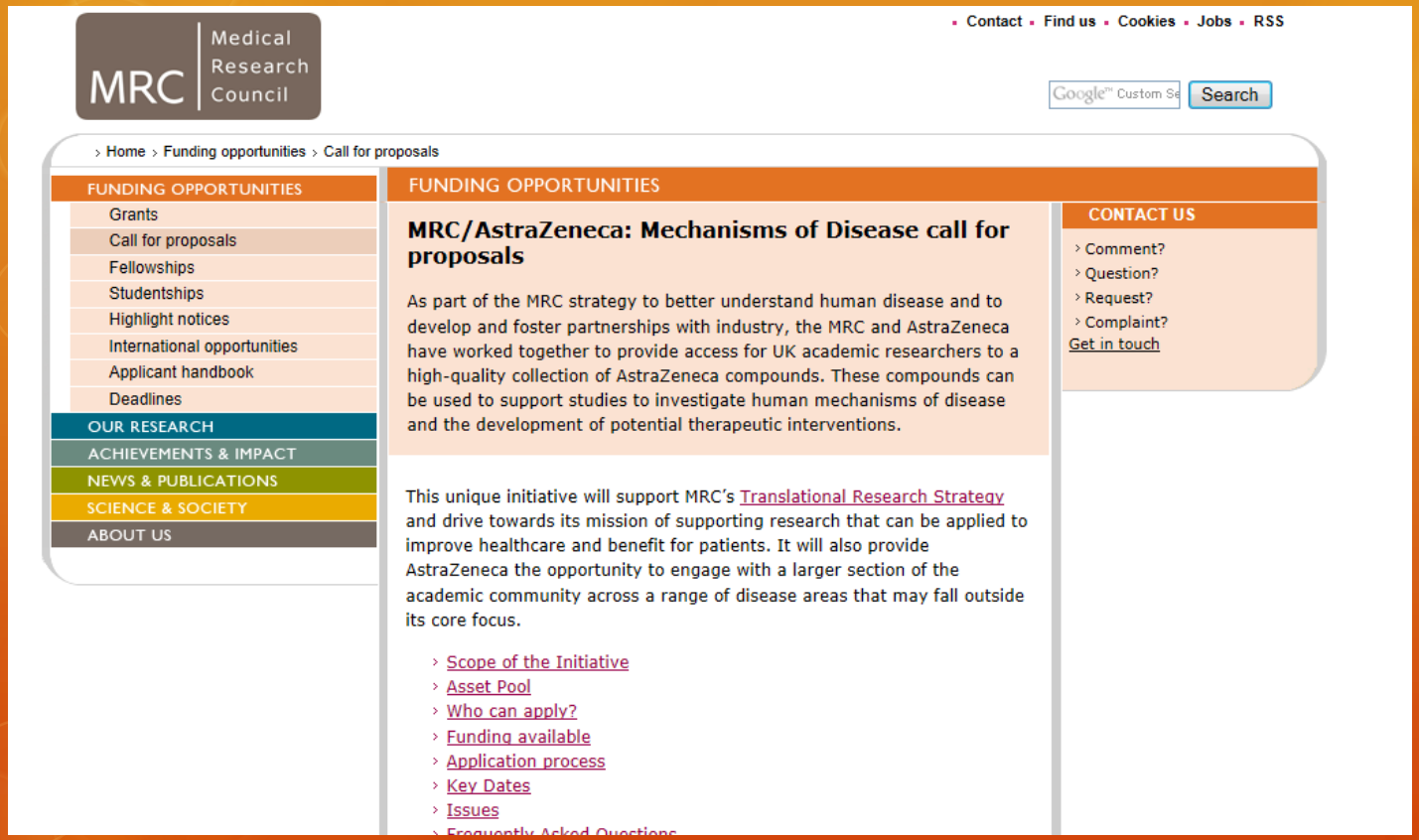
Kim K, Lee SG, Kegelman TP, Su ZZ, Das SK, Dash R, Dasgupta S, Barral PM, Hedvat M, Diaz P, Reed JC, Stebbins JL, Pellecchia M, Sarkar D, Fisher PB.

Could In silico / in vitro repositioning find leads-drugs quicker?

Approximate small-molecule drug and proto-drug numbers

Historical development entry	**~35,000**
Approached regulatory entry (INNs)	**~7,000**
Between clinical phases	**~15,000**
In active trials	**~1,500**
FDA approved	**~1,400**
INNs issued per year	**~150**
Discontinued (post approval)	**~50**
New approvals per year	**~15**

MRC/AstraZeneca: Mechanisms of Disease Call assets



Launched Dec 2011

http://goo.gl/R8trN

http://www.mrc.ac.uk/Fundingopportunities/Calls/MoD/MRC008389

MRC/AstraZeneca: Mechanisms of Disease Call assets

Full list of compounds:

Please click the AZ code for further compound information.

CONTACT

- Comment?
- Question?
- Request?
- Complaint?

Get in touch

Back To Top

Compound details

Pre-Development code	AZ code & further information	Mechanism of Action	Original development indication	Type of proposals invited
AZ10353926	**AZD0530** (Saracatinib)	SRC Tyrosine Kinase Inhibitor	Solid tumour	Pre-clinical & Clinical
AZ12272852	**AZD1236**	Matrix Metallopeptidase (MMP) 9|12 Inhibitor	Chronic Obstructive Pulmonary Disease	Pre-clinical & Clinical
AZ12501796	**AZD1656**	Glucokinase Activator	Diabetes and Obesity	Pre-clinical & Clinical
AZ12472520	**AZD2624**	Neurokinin Receptor NK3 Antagonist	Schizophrenia	Pre-clinical & Clinical
AZ11941831	**AZD3355**	GABABR1 Receptor Agonist	Gastroesophageal Reflux Disease	Pre-clinical & Clinical

22 molecules from AstraZeneca

NCATS

May 2012 the National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) launched the 'Discovering New Therapeutic Uses for Existing Molecules' program.

http://goo.gl/FWchw

http://www.ncats.nih.gov/research/reengineering/rescue-repurpose/therapeutic-uses/directory.html

Library of Industry Provided Reagents

Table of Compounds and Biologics*

Code Number & Link to More Information	Mechanism of Action	Original Development Indication(s)	Route of Administration Formulation Available (CNS Penetrant+)
AVE5530 canosimibe	Acyl-coenzyme A:cholesterol O -acyltransferase (ACAT) inhibitor Cholesterol absorption inhibitor	Hypercholesterolemia	Oral
SSR149744C celivarone	Anti-arrhythmic, Vaughan Williams Class I to IV	Maintenance of sinus rhythm in atrial fibrillation patients Prevention of shocks and major clinical outcomes in patients with implanted cardiac defibrillator	Oral
PF-05416266 senicapoc (ICA-17043)	Calcium-activated potassium channel blocker (KCa3.1), intermediate-conductance	Sickle cell disease Asthma	Oral
ABT-639	Calcium channel, voltage-gated (Cav3.2, T-type) blocker	Pain	Oral (Yes)
CP-945598 otenabant	Cannabinoid receptor 1 (CB1) antagonist	Obesity	Oral (Yes)
LY2828360	Cannabinoid receptor 2 (CB2) agonist	Osteoarthritis pain	Oral (Yes)
AZD1981	Chemoattractant receptor-homologous molecule expressed on Th2 cells (CRTh2)/prostaglandin D2	Asthma Chronic obstructive pulmonary disease	Oral

Consists of 58 'parked' therapeutic agents

contributed by eight drug companies,

To be evaluated by academics as a repurposing pilot effort funded by NCATS

No structures provided on website!

It started with a blog!

Jul 18

Petition NCATS to release structures for library of Industry Provided Reagents

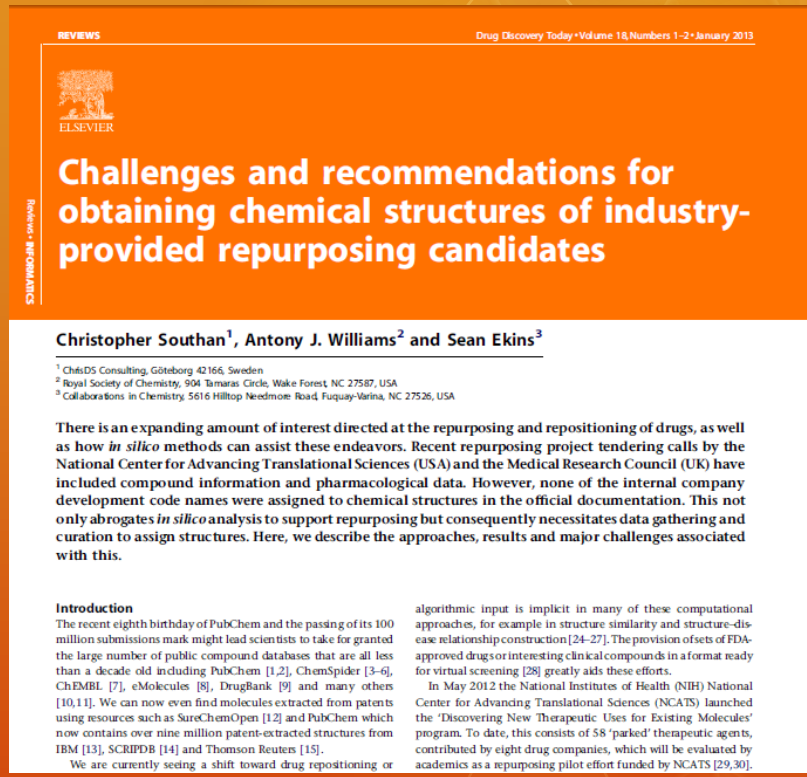
Uncategorized by sean

Going to stick my little neck out again. I am totally behind the idea of drug companies making their compounds available for researchers to find new uses for rare diseases. But I found out last week that NCATS had put a list of compounds and biologics on their website without releasing structures. This is absolutely nutty (by all means correct me). They release a lot of other data but no structures. If you have commercial databases you probably will be able to find the compounds after a bit of digging but WHY? I ask you.

Why might we want to see the structures? – how about using in silico tools to screen the compounds for new activities. Apparently this paper was available to the folks at NIH when they had the April 2011 roundtable..but I guess they forgot it. So by not providing the structures anyone who wants to do in silico screening is precluded from doing this for these compounds until they can spend a significant time digging them out of the literature.

We might also want to analyze their physicochemical properties to predict if the compounds even stand a chance of finding other activities,or targets, for example lots of work on similarity analysis seems to be going on nowadays.

So Please! I call on everyone to suggest, petition, tweet, blog, whatever to NCATS (#NCATS) that if they do not make the compounds available I am sure other ways will be found to make them accessible. This might just help the rare and neglected disease community besides a lucky few people who get grants to do yet more in vitro and in vivo testing. Wake up – in silico is cheaper and faster and just might find something valuable..but only if structures are provided.



REVIEWS Drug Discovery Today • Volume 18, Numbers 1–2 • January 2013

ELSEVIER

Reviews • INFORMATICS

Challenges and recommendations for obtaining chemical structures of industry-provided repurposing candidates

Christopher Southan[1], Antony J. Williams[2] and Sean Ekins[3]

[1] ChrisDS Consulting, Göteborg 42166, Sweden
[2] Royal Society of Chemistry, 904 Tamaras Circle, Wake Forest, NC 27587, USA
[3] Collaborations in Chemistry, 5616 Hilltop Needmore Road, Fuquay-Varina, NC 27526, USA

There is an expanding amount of interest directed at the repurposing and repositioning of drugs, as well as how *in silico* methods can assist these endeavors. Recent repurposing project tendering calls by the National Center for Advancing Translational Sciences (USA) and the Medical Research Council (UK) have included compound information and pharmacological data. However, none of the internal company development code names were assigned to chemical structures in the official documentation. This not only abrogates *in silico* analysis to support repurposing but consequently necessitates data gathering and curation to assign structures. Here, we describe the approaches, results and major challenges associated with this.

Introduction

The recent eighth birthday of PubChem and the passing of its 100 million submissions mark might lead scientists to take for granted the large number of public compound databases that are all less than a decade old including PubChem [1,2], ChemSpider [3–6], ChEMBL [7], eMolecules [8], DrugBank [9] and many others [10,11]. We can now even find molecules extracted from patents using resources such as SureChemOpen [12] and PubChem which now contains over nine million patent-extracted structures from IBM [13], SCRIPDB [14] and Thomson Reuters [15].

We are currently seeing a shift toward drug repositioning or algorithmic input is implicit in many of these computational approaches, for example in structure similarity and structure-disease relationship construction [24–27]. The provision of sets of FDA-approved drugs or interesting clinical compounds in a format ready for virtual screening [28] greatly aids these efforts.

In May 2012 the National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) launched the 'Discovering New Therapeutic Uses for Existing Molecules' program. To date, this consists of 58 'parked' therapeutic agents, contributed by eight drug companies, which will be evaluated by academics as a repurposing pilot effort funded by NCATS [29,30].

http://goo.gl/uTswV

Southan et al., DDT, 18: 58-70 (2013)

Pooling resources

Tuesday, 10 July 2012

NCATS repurposing candiates, names-to-structures

24 of August: There is now a follow-up post; NCATS repurposing compounds in PubChem: Pa
structures.
29 July: there are some updates, comments, and new IDs described at the bottom of this pos
resolved to CIDs is now 30 out of 56. Checking has revealed source ambiguity for one record.
that preceded this one but picks up interesting and different angles (petition-ncats-to-release-

****

For the newly announced Discovering New Therapeutic Uses for Existing Molecules initiative a
Accelerating Translational Science (NCATS) eight companies have agreed to make 58 agents
(56 small-molecules plus one antibody and one plasmid). These have undergone pre-clinical a
open for applications from academic groups to explore new mechanisms and therapeutic oppo
seems like a good thing I feel no need to add to the punditry on this subject that includes the
repurposing programmes get lift off ".

However, I will pick up on the important aspect of blinding of the name-to-structure (n>s) link
be equable many companies do publish papers on clinical candidates wherein this link is mad
essentially *ad hoc* process by which of several routes an n>s linkage may transfer to the majo
structures have usually been exemplified in published patents some companies persist in exte
code name for the drug candidate. There is some background on the first few slides of this pr
discussed this in a a recent article on dug name mapping in clinicaltrials.org The issues arou
include the following:

1) In the NCATS context discovering new therapeutic uses for existing molecules is difficult
2) Not being able to do any chemoinformatics, including virtual screening, open tox resource
make proposals for taking compounds forward.
3) Even if IP holders for these structures disclose them for applicants under CDA this still c
to public surfacing (e.g. no external database searches ?).
4) It seems increasingly anachronistic that clinical candidates are allowed to stay blinded ri
5) Both journals andclinical trials databases are complicit in not mandating n>s for publicatio

http://goo.gl/FW6Bl

Aug 21

Collaborations to get the NCATS Library of Industry provided reagents

repurposing by sean

Its seems a while since I blogged on the absolutely bizarre posting of 58 molecules as the 'library of industry provided reagents' to be used as a starting point for repurposing – without posting structures. Since my last Blog I have become aware of at least 3 specific groups trying to collate the molecules, Tudor Oprea and collaborators at UNM and elsewhere..Chris Lipinski and Chris Southan. At the recent ACS in Philly. Chris Lipinski presented his results and I thank him for sharing the data and molecules (he included the Oprea results)..

Chris Lipinski was able to find 38 small molecules and 2 biologics usng CAS SciFinder, Thomson Reuters Integrity, various web postings

Tudor Oprea et al. was able to find 41 small molecules and 2 biologics using US Patents database (IBM), Google, publications

Chris Southan described his approach on his blog and found 30 compounds and put them in PubChem

We have looked at the molecules Chris Lipinski found and could not find a significant difference in a few molecular properties to differentiate those discontinued and those still in clinical trials.

What has not been done so far is look at overlap across all 3 groups above. How can we bring all these efforts together? Are there other efforts to do the same out there? e.g. have NCATS tried to do this?

But the question still resonates WHY?

Why do these 3 groups 'have to' collate the molecules?

Why could the NCATS initiative have not posted the molecules on the website or linked to them in PubChem in the first place (would have taken no effort for each company to provide a structure)?

Why did they get groups to propose repurposing the molecules without disclosing molecule structures?

Why is computational analysis not at the forefront of the repurposing efforts before spending experimental resources?

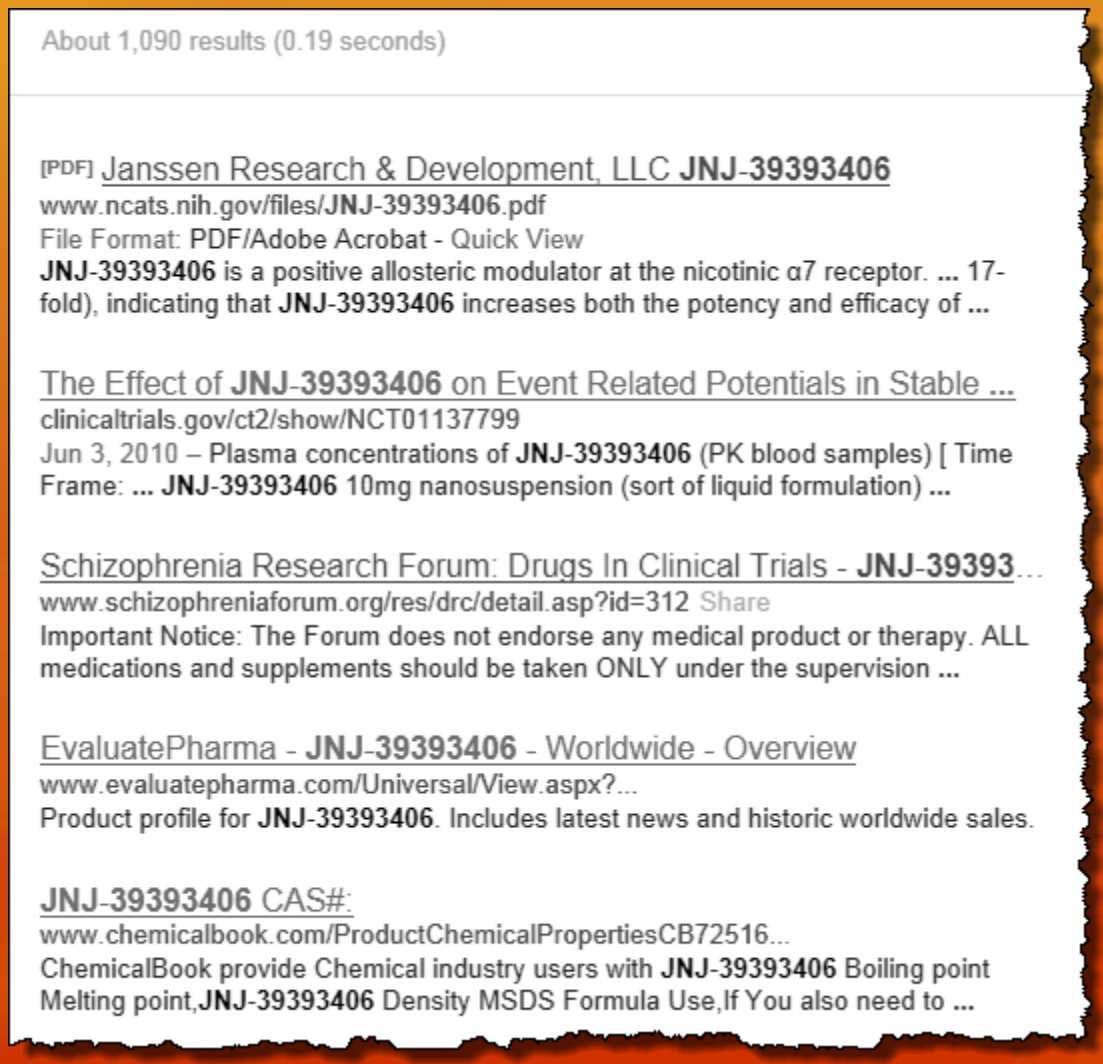
Why oh why did someone not think of this, or did they?

Its not like people have not proposed how to do this kind of thing before.

http://goo.gl/NXPVs

Searching for one member of the NCATS list

JNJ-39393406, using a standard Google search, provides >1000 hits (some are highlighted here)



About 1,090 results (0.19 seconds)

[PDF] Janssen Research & Development, LLC **JNJ-39393406**
www.ncats.nih.gov/files/**JNJ-39393406**.pdf
File Format: PDF/Adobe Acrobat - Quick View
JNJ-39393406 is a positive allosteric modulator at the nicotinic α7 receptor. ... 17-fold), indicating that **JNJ-39393406** increases both the potency and efficacy of ...

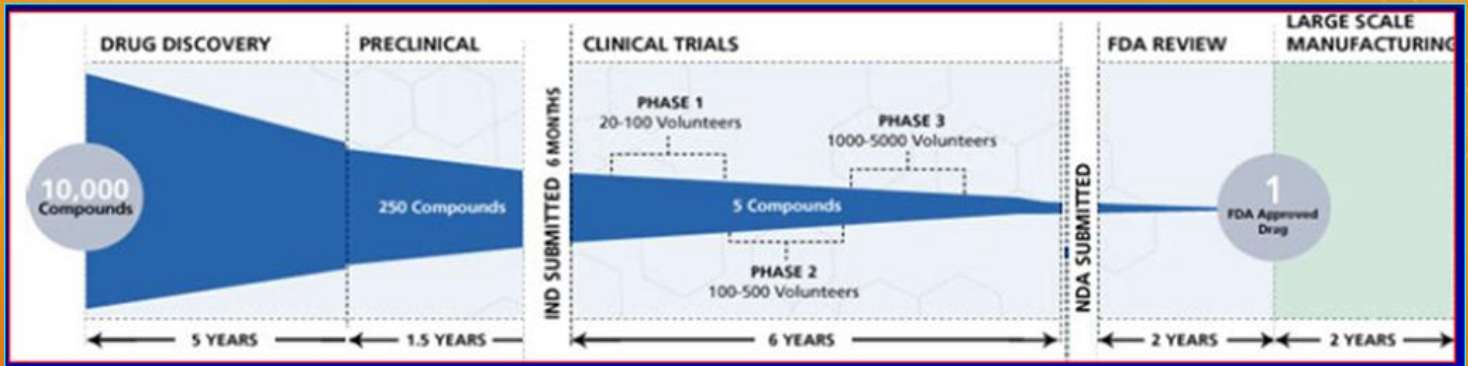
The Effect of **JNJ-39393406** on Event Related Potentials in Stable ...
clinicaltrials.gov/ct2/show/NCT01137799
Jun 3, 2010 – Plasma concentrations of **JNJ-39393406** (PK blood samples) [Time Frame: ... **JNJ-39393406** 10mg nanosuspension (sort of liquid formulation) ...

Schizophrenia Research Forum: Drugs In Clinical Trials - **JNJ-39393**...
www.schizophreniaforum.org/res/drc/detail.asp?id=312 Share
Important Notice: The Forum does not endorse any medical product or therapy. ALL medications and supplements should be taken ONLY under the supervision ...

EvaluatePharma - **JNJ-39393406** - Worldwide - Overview
www.evaluatepharma.com/Universal/View.aspx?...
Product profile for **JNJ-39393406**. Includes latest news and historic worldwide sales.

JNJ-39393406 CAS#:
www.chemicalbook.com/ProductChemicalPropertiesCB72516...
ChemicalBook provide Chemical industry users with **JNJ-39393406** Boiling point Melting point,**JNJ-39393406** Density MSDS Formula Use,If You also need to ...

Patterns of information disclosures



Approximate timelines

[cpd registration system structure and ID----------]

[patent IUPAC or image----------]

[internal code name(s) externally blinded-------]

[code name(s) > structure declared externally -----]

[journal papers ----------]

[International Non-proprietary name INN]

[INN indexed in MeSH----------]

[USAN, BAN, JAN ----------]

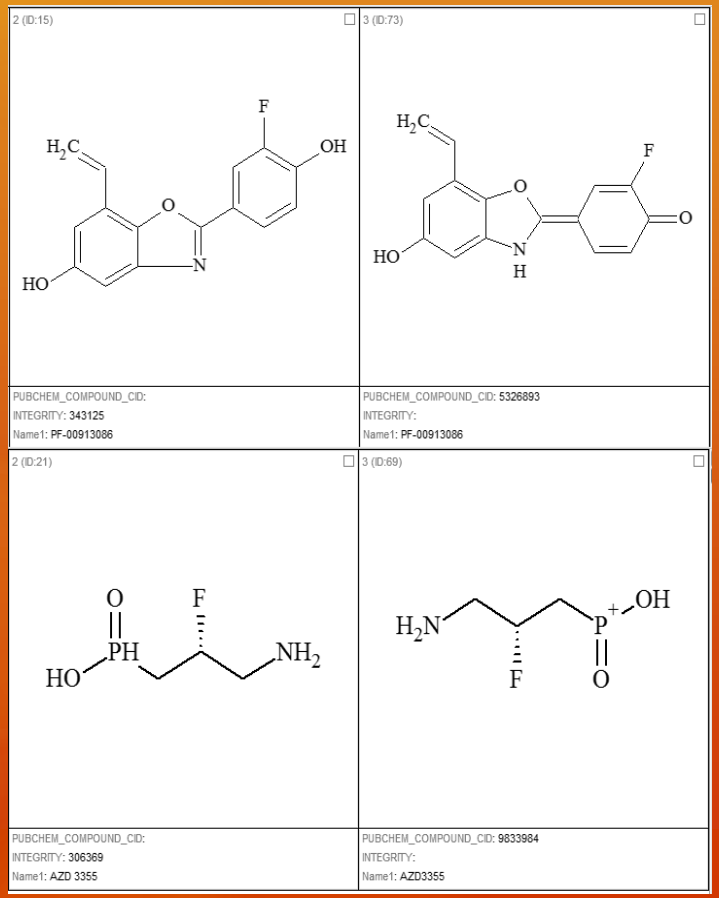
[brand name(s)----------]

Southan et al., DDT, 18: 58-70 (2013)

Different forms of compounds

provided by the analyses of Southan and Lipinski (protonation state and tautomerization).

Found by AJW - manual review of the ChemFOlder file

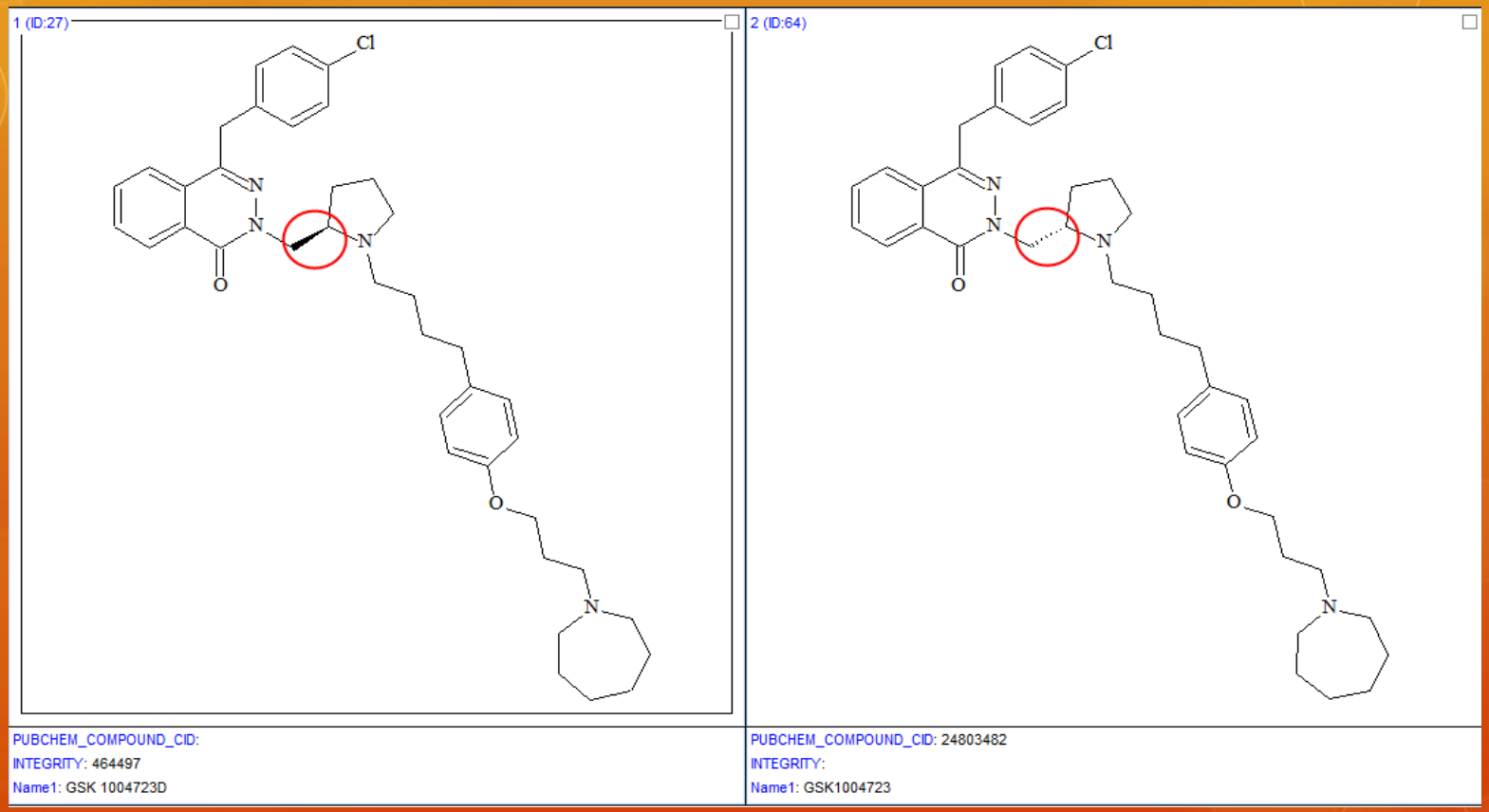


http://goo.gl/yIcVy

Southan et al., DDT, 18: 58-70 (2013)

Single stereocenter inversion difference

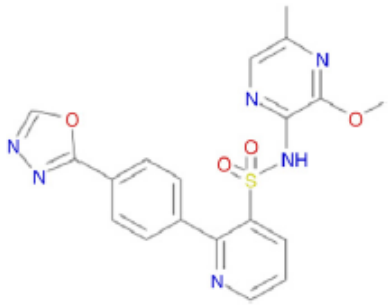
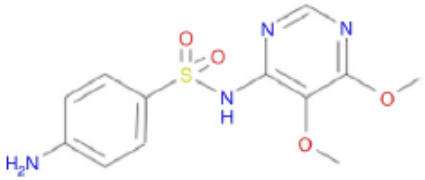
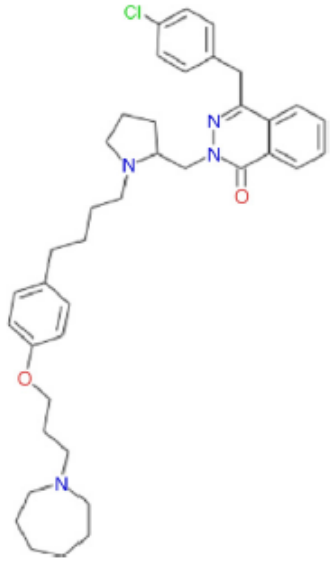
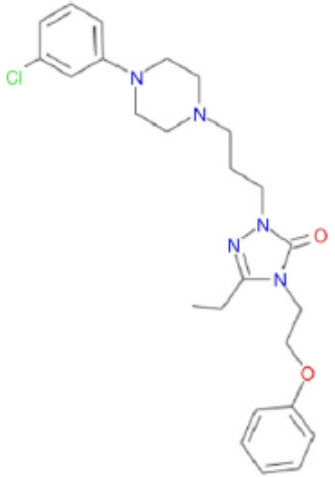
detected during the analyses of Southan and Lipinski for GSK1004723.



Southan et al., DDT, 18: 58-70 (2013)

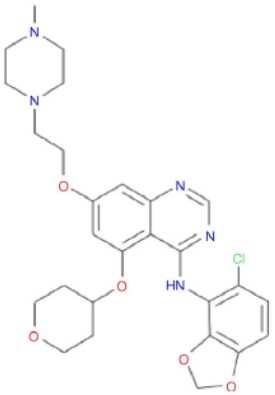
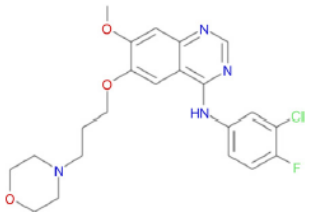
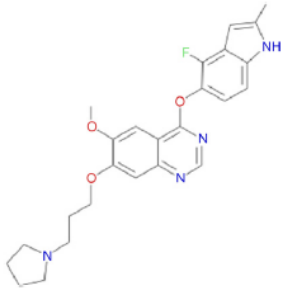
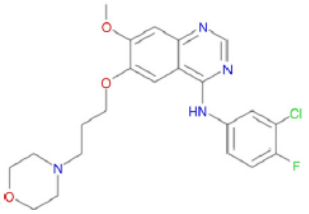
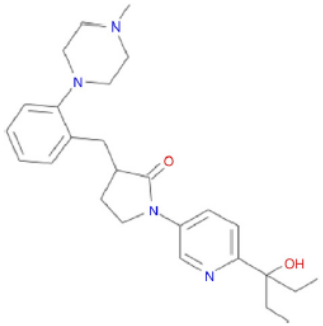
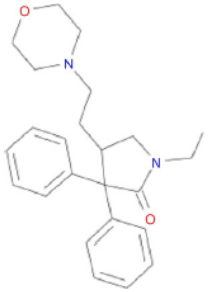
Repurposing by similarity to known drugs

NCATS molecules with similarity to approved drugs using molecules from a public database [28].

NCATS ID Target and mechanism information [87]	NCATS structure	Closest structure in FDA drugs set (Tanimoto similarity using MDL keys) Target and mechanism from Wikipedia	FDA drug structure
ZD4054 Oncology, pulmonary artery hypertension Endothelin receptor A (ET_A) antagonist		Sulfadoxine (0.85) Antimalarial Sulfadoxine competitively inhibits dihydropteroate synthase, interfering with folate synthesis	
GSK 1004723D Allergic rhinitis Histamine H1/H3 receptor antagonist		Nefazodone (0.81) Antidepressant targets 5-HT_{2A} receptors and other GPCRs	

Southan et al., DDT, 18: 58-70 (2013)

Repurposing by similarity to known drugs

AZD 0530 Oncology Src tyrosine kinase inhibitor		Gefitinib (0.85) Anticancer, selective inhibitor of epidermal growth factor receptor (EGFR) tyrosine kinase domain	
AZD 2171 Oncology Vascular endothelial growth factor receptor (VEGFR)1, 2 and 3 tyrosine kinase inhibitor		Gefitinib (0.81) Anticancer, selective inhibitor of epidermal growth factor receptor (EGFR) tyrosine kinase domain	
CE-210666 Depression 5-Hydroxytryptamine 1B receptor (5-HT_{1B}) antagonist		Doxapram (0.78) A respiratory stimulant acts on chemoreceptors in the carotid bodies	

Southan et al., DDT, 18: 58-70 (2013)

Summary of NCATS and MRC results (CL = Chris Lipinski, CS = Chris Southan). The aggregate collaborative result for the NCATS set was 41. Of these, 37 had PubChem CID assignments plus four SMILES with no exact match in PubChem.

Classification	Count
NCATS: listed reagents	58
NCATS: biologics	2
NCATS: small molecules	56
Out of 56: INNs and/or USANS (and code names)	15
Out of 56: code-names-only (no INN)	41
Out of 56: code names in ChemSpider	8
Out of 56: CL found	36
Out of 56: CS found	37
Out of 56: CS or CL	41
Out of 56:CL but not CS	4
Out of 56:CS but not CL	10
Out of 56:CS and CL	30
From 41 NCATS code-names-only: PubChem matches	3
From 41 NCATS structures: PubChem CIDs	37
From 41 NCATS structures: SMILES-only	4
NCATS code names still blinded	15
MRC–AZ list	22
MRC AZD codes	21
MRC AZD codes: with INNs	3
MRC AZD codes: in NCATS	7
MRC AZD codes: mapped to CIDs	12
MRC AZD codes: still blinded	10
MRC + NCATS: structures	46
MRC + NCATS: ChemSpider CS IDs	37
MRC + NCATS: PubChem CIDs	42
MRC + NCATS 42 CIDs: active in bioAssays (ChEMBL)	20
MRC + NCATS 42 CIDs: 'same connectivity' isomer expansion	139
MRC + NCATS 42 CIDs: vendor SIDs	15
MRC + NCATS 42 CIDs: SureChem patent matches	42

Cutting to the chase

Several months led to structures

12 of 22 MRC cpds

41 of 56 NCATS cpds

Also ran predictions with TB and malaria Bayesian models (data on request)

Southan et al., DDT, 18: 58-70 (2013)

Making molecules public

Oct 26

NCATS Repurposing – paper in press and molecules tweeted

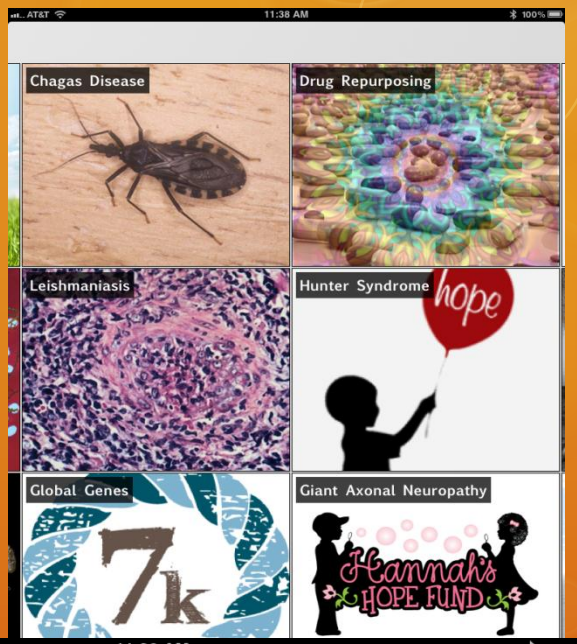
collaborative science and computers, databases, Open Data, Uncategorized

by sean

From past blogs I have detailed The NCATS library of industry Provided reagents.

Through the heroic collaborative efforts of Dr. Chris Southan and Dr. Antony Williams we now have the structures for most of the NCATS and MRC molecules described in a paper in press. In addition the molecules sourced to date have just been tweeted by myself @collabchem with the hash tag #ODDT so you can go and use them for analysis.

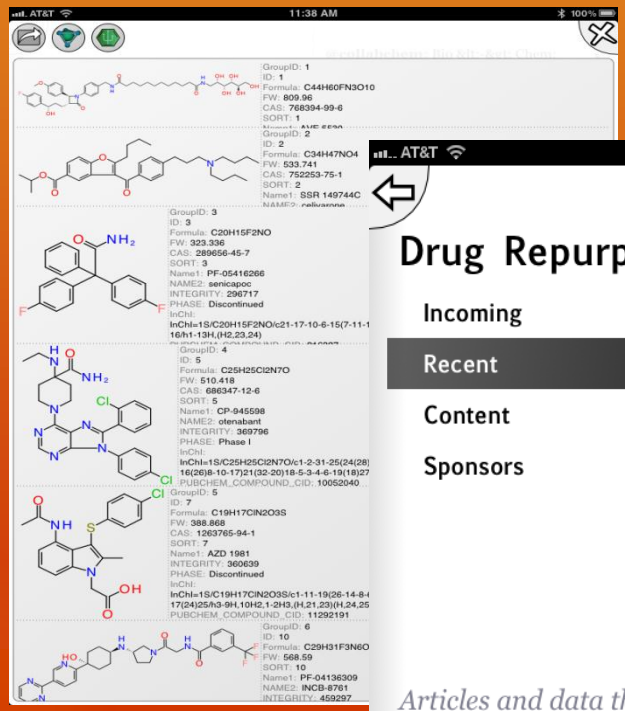
Big thanks to Dr. Alex Clark with helping to get them out through mobile apps..I tweeted from his App MMDS. A historical moment?

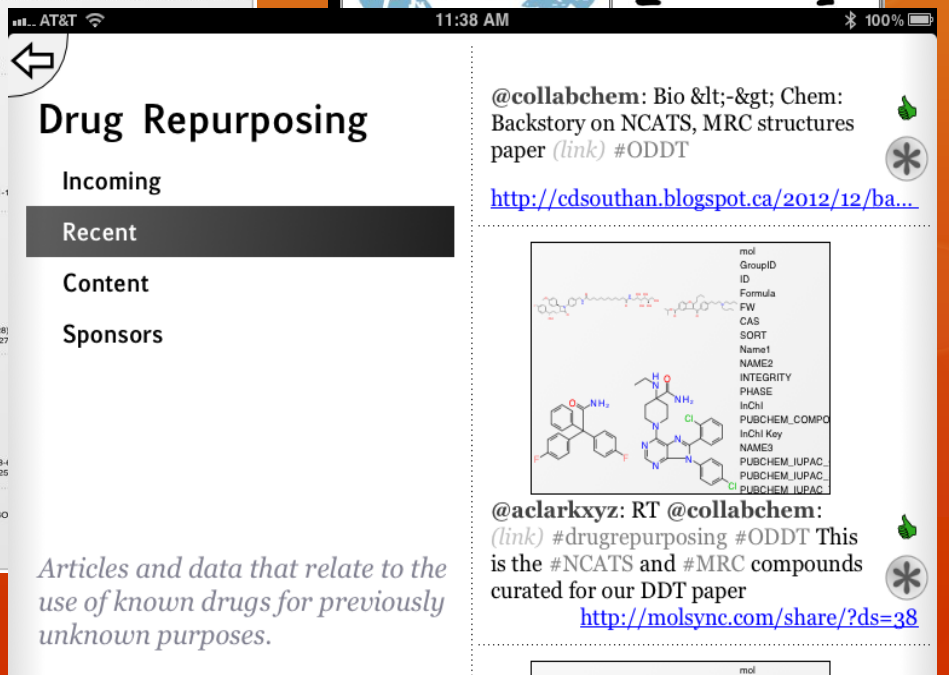


Drew molecules in MMDS app

Tweeted them with #drugrepurposing #oddt

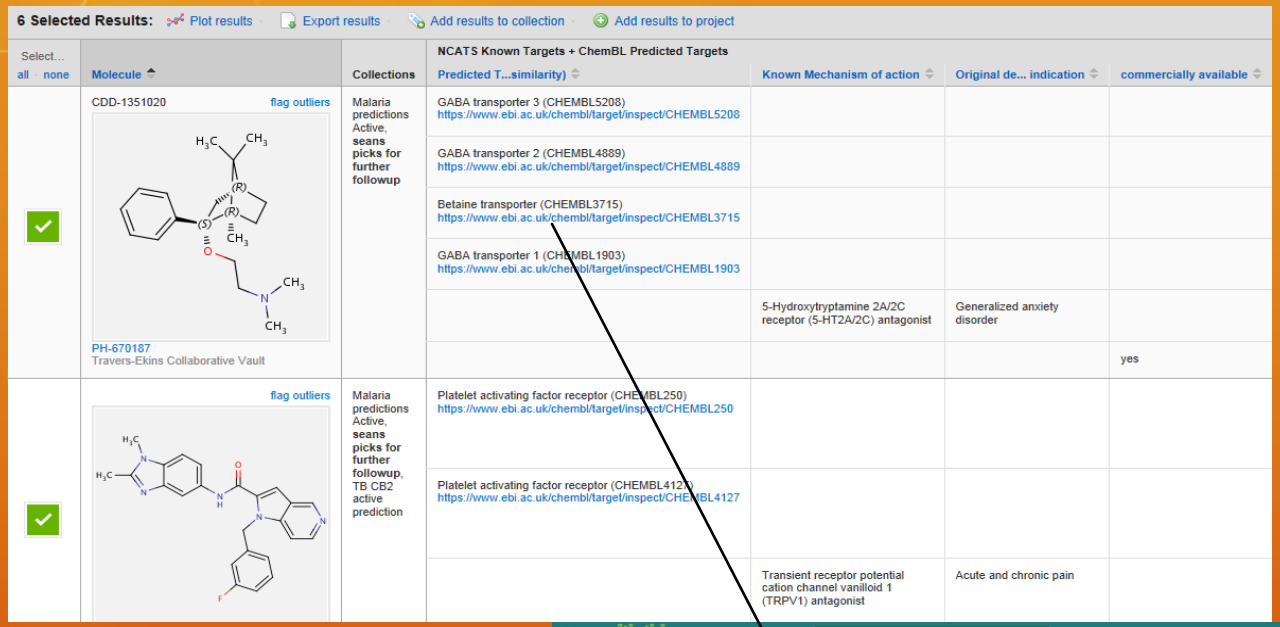
Visible and downloadable in Open Drug Discovery Teams Mobile app (free)



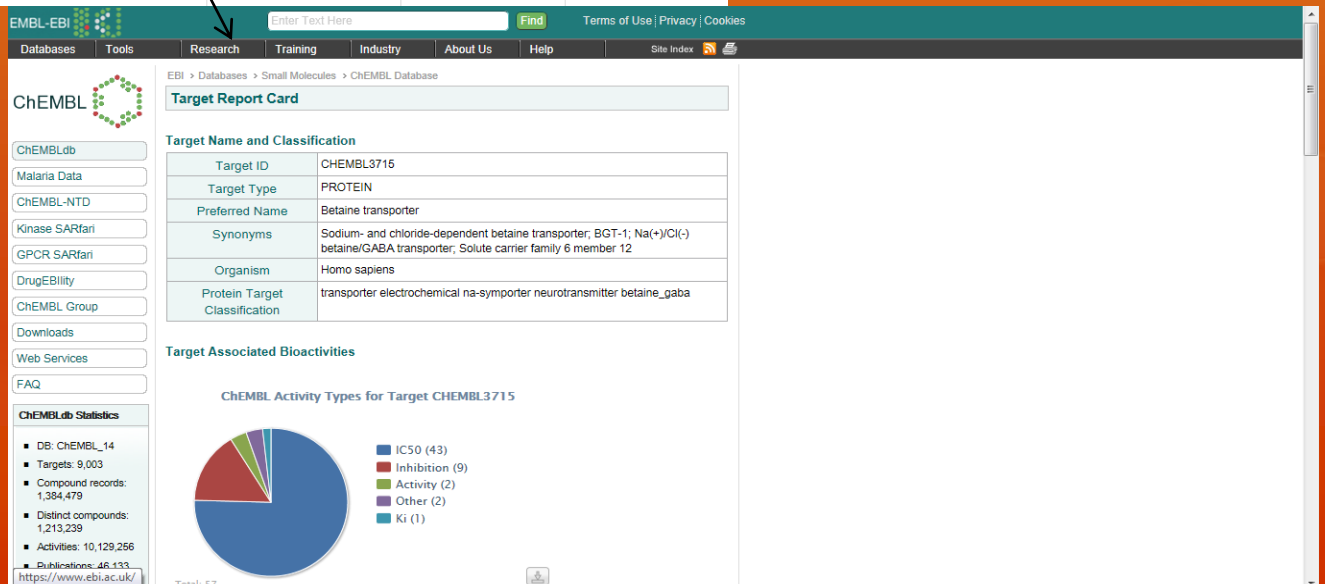


Mol Informatics, 31: 585-597, 2012

Adding links to most similar targets in CHEMBL using CDD



- Data currently in a private vault
- Potential target inference by ligand similarity
- Will make public
- Validation



Challenges

Despite the comprehensive interrogation of data sources, some representations remain equivocal

Organizations allocating the company codes have the primary provenance for the fidelity of the relationship between database electronic structure representations and their own results in vitro, in vivo and in the clinic.

Avoid duplication of effort by others

Companies interested in verifying or correcting our assignments, or even surfacing de novo their hitherto blinded mappings, are welcome to contact us.

Recommendations

Authors to ensure their drafted PubMed abstract encapsulates the code name-to-structure mapping by having the code name and IUPAC juxtaposed in the abstract text (in the title is even better)

Inclusion in the abstract of the Human Genome Organization (HUGO) Gene Nomenclature Committee (HGNC) gene symbol for the primary target (or other major protein database identifier for non-human targets)

Before resource-intensive experimental testing it would therefore be valuable to run a battery of in silico methods to predict, prospectively, potential new targets and new uses and reduce off-target or safety risks

Provide repurposing opportunities that were not envisaged in the initial proposal calls.

Acknowledgments

- Chris Lipinski
- Dr Jeremy Yang and colleagues (University of New Mexico) for kindly providing access to the Smartsfilter web
- application
- Alex Clark (MMI)
- Steve Carney & Reviewers at DDT
- Mike Travers
- Barry Bunin

Disclaimer

We accept no responsibility for the correctness of the structures

You can find me @... CDD Booth 205

PAPER ID: 13433
PAPER TITLE: “Dispensing processes profoundly impact biological assays and computational and statistical analyses”
April 8th 8.35am Room 349

PAPER ID: 14750
PAPER TITLE: “Enhancing High Throughput Screening For Mycobacterium tuberculosis Drug Discovery Using Bayesian Models”
April 9th 1.30pm Room 353
PAPER ID: 21524

PAPER TITLE: “Navigating between patents, papers, abstracts and databases using public sources and tools”
April 9th 3.50pm Room 350
PAPER ID: 13358

PAPER TITLE: “TB Mobile: Appifying Data on Anti-tuberculosis Molecule Targets”
April 10th 8.30am Room 357

PAPER ID: 13382
PAPER TITLE: “Challenges and recommendations for obtaining chemical structures of industry-provided repurposing candidates”
April 10th 10.20am Room 350

PAPER ID: 13438
PAPER TITLE: “Dual-event machine learning models to accelerate drug discovery”
April 10th 3.05 pm Room 350