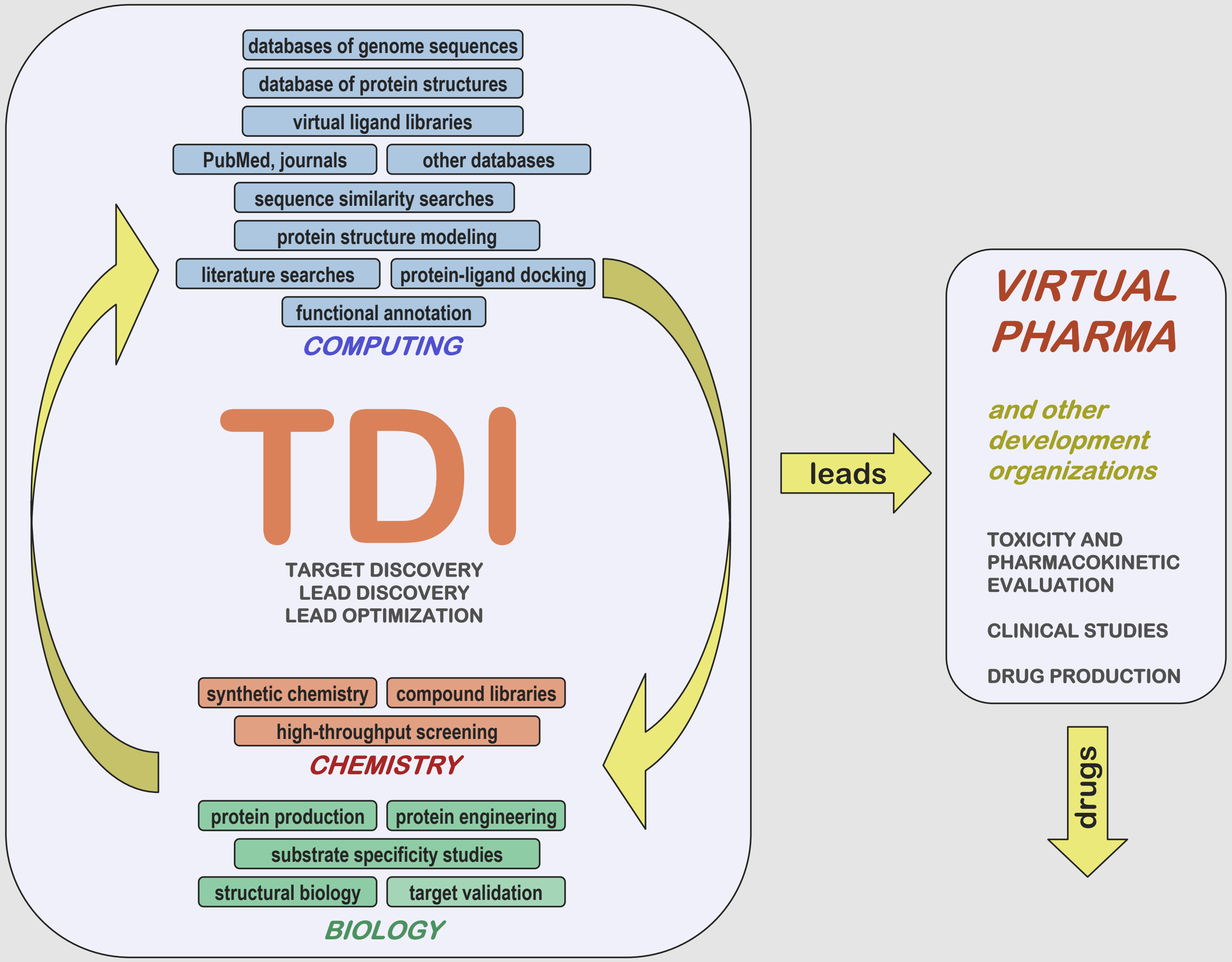
A “kernel” for the Tropical Disease Initiative

An open source approach to drug discovery



databases of genome sequences

database of protein structures

virtual ligand libraries

PubMed, journals

other databases

sequence similarity searches

protein structure modeling

literature searches

protein-ligand docking

functional annotation

COMPUTING

TDI

TARGET DISCOVERY
LEAD DISCOVERY
LEAD OPTIMIZATION

synthetic chemistry

compound libraries

high-throughput screening

CHEMISTRY

protein production

protein engineering

substrate specificity studies

structural biology

target validation

BIOLOGY

leads

VIRTUAL PHARMA

and other development organizations

TOXICITY AND PHARMACOKINETIC EVALUATION

CLINICAL STUDIES

DRUG PRODUCTION

drugs

Marc A. Marti-Renom
http://sgu.bioinfo.cipf.es

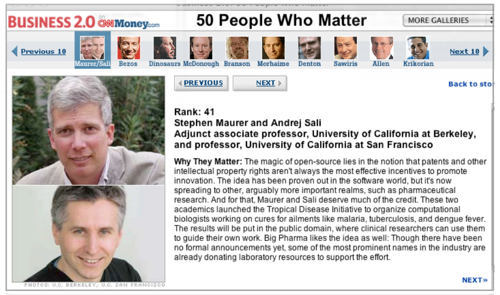
Structural Genomics Unit
Bioinformatics Department
Prince Felipe Resarch Center (CIPF), Valencia, Spain

TDI *a story*

the Tropical Disease Initiative

In brief...

Kernel 1.0



BUSINESS 2.0 on CNNMoney.com

50 People Who Matter

Rank: 41
Stephen Maurer and Andrej Sali
Adjunct associate professor, University of California at Berkeley, and professor, University of California at San Francisco

Why They Matter: The magic of open-source lies in the notion that patents and other intellectual property rights aren't always the most effective incentives to promote innovation. The idea has been proven out in the software world, but it's now spreading to other, arguably more important realms, such as pharmaceutical research. And for that, Maurer and Sali deserve much of the credit. These two academics launched the Tropical Disease Initiative to organize computational biologists working on cures for ailments like malaria, tuberculosis, and dengue fever. The results will be put in the public domain, where clinical researchers can use them to guide their own work. Big Pharma likes the idea as well: Though there have been no formal announcements yet, some of the most prominent names in the industry are already donating laboratory resources to support the effort.

PHOTOS: U.C. BERKELEY, U.C. SAN FRANCISCO

The Synaptic Leap

Open Source Biomedical Research

Navigation

IP Guidelines

Welcome

Open Source Biomedical Research for the 21st Century

Roadmap

User login

Recent posts

2004

.Steve Maurer (Berkeley) and Arti Rai (Duke)

.PLoS Medicine, Dec. 2004. Vol 1(3):e56

2005

.TDI web site http://TropicalDisease.org

.Ginger Taylor and The Synaptic Leap

2006

.TSL web site http://TheSynapticLeap.org

.Maurer and Sali 41th in “50 Who Matter”

2008

.TDI kernel http://TropicalDisease.org/kernel

2009

.TDI kernel papers published

Initial feed-back...

14 Mar 2005

I think TDI is a unique and very interesting project. I would like so much to make something for it...

So, where are we going? What's happening? What can we do?

I still trust in open source drug discovery. :-))

Luca Brivio

any, the bottlenecks are?
eas and potential avenues to explore,
tion Plan!

Regards,
Jacob Lester

I am interested in beginning resea
disease for underserved populati
however, confused.
If someone will tell me where te
begin on, I'd be greatful.

Thank you kindly,
Adam Huber

9 Mar 2005
I'm a programmer, not a
the list active :)

GNU started with RMS.
Linux started with Linu
You need someone gre
sending patches...

I know this is chicken-egg, but someone needs to point this out, since I haven't seen this brought up in the papers or the website.

And you might consider merging into the bios.net effort mentioned already. Together, you just might reach the critical mass for things to take off. Consider this like when people jumped off the HURD project to come together and make linux work.

Daniel Amelang

stic that the rest

Stephen Mark Maurer

Initial feed-back...

14 Mar 2005

I think TDI is a unique and very interesting project. I w
it...

So, where are we going? What's happening? Wha

I still trust in open source drug discovery. :-))

Luca Brivio

16 Feb 2005

Hi,

It would be interesting to know what, if any, the bottlenecks are?
The Wiki site contains many interesting ideas and potential avenues to explore, but from what I can see it is **lacking an Action Plan!**

Regards,
Jacob Lester

I am interested in beginning rese
disease for underserved populati
however, confused.
If someone will tell me where te
begin on, I'd be greatful.

Thank you kindly,
Adam Huber

9 Mar 2005
I'm a programmer, not a
the list active :)

GNU started with RMS.
Linux started with Linu
You need someone gre
sending patches...

I know this is chicken-egg, but someone needs to point this out, since I haven't seen this brought up in the papers or the website.

And you might consider merging into the bios.net effort mentioned already. Together, you just might reach the critical mass for things to take off. Consider this like when people jumped off the HURD project to come together and make linux work.

Daniel Amelang

stic that the rest

Stephen Mark Maurer

Initial feed-back...

14 Mar 2005

I think TDI is a unique and very interesting project. I w
it...

So, where are we going? What's happening? Wha

I still trust in open s

Luca Brivio

16 Feb 2005

Hi,

bottlenecks are?
potential avenues to explore,
n!

10 Feb 2005

Hello,
My name is Adam Huber and I am a medical student at UNSW in Sydney Australia. I am interested in beginning research focused on tropical and infectious disease for underserved populations (A mission that seemingly matches TDI). I am, however, confused.
If someone will tell me where to sign up and give me some research topics to begin on, I'd be greatful.

Thank you kindly,
Adam Huber

9 Mar 2005
I'm a programmer, not a
the list active :)

GNU started with RMS.
Linux started with Linu
You need someone gre
sending patches...

I know this is chicken-egg, but someone needs to point this out, since I haven't seen this brought up in the papers or the website.

And you might consider merging into the bios.net effort mentioned already. Together, you just might reach the critical mass for things to take off. Consider this like when people jumped off the HURD project to come together and make linux work.

Daniel Amelang

stic that the rest

Stephen Mark Maurer

Initial feed-back...

14 Mar 2005

I think TDI is a unique and very interesting project. I w
it...

So, where are we going? What's happening? Wh

I still trust in open s

Luca Brivio

16 Feb 2005

Hi,

bottlenecks are?
potential avenues to explore,
n!

10 Feb 2005

Hello,
My name is Adam Huber and I am a medical student at UNSW in Sydney Australia.
I am interested in beginning research focused on tropical and infectious
disease for underserved populations (A mission that seemingly matches TDI). I am,

9 Mar 2005
I'm a programmer, not a bioinformatician, but I stumbled across your site and thought I'd say something to keep the list active :)

GNU started with RMS. He gave us programming/administration tools to play with.
Linux started with Linus. He released an operating system for us to play with.
You need someone great in the field to release something for everyone to 'play with'. Then people start sending patches...

I know this is chicken-egg, but someone needs to point this out, since I haven't seen this brought up in the papers or the website.

And you might consider merging into the bios.net effort mentioned already. Together, you just might reach the critical mass for things to take off. Consider this like when people jumped off the HURD project to come together and make linux work.

Daniel Amelang

stic that the rest

Stephen Mark Maurer

Initial feed-back...

14 Mar 2005

I think TDI is a unique and very interesting project. I w
it...

So, where are we

I still trust in open s

Luca Brivio

16 Feb 2005

Hi,

bottlenecks are?
otential avenues to explore,
n!

10 Feb 2005

Hello,
My name is Adam Huber and I am a medical student at UNSW in Sydney Australia.
I am interested in beginning research focused on tropical and infectious disease for underserved populations (A mission that seemingly matches TDI). I am,

9 Mar 2005
I'm a programmer, not a bioinformatician, but I stumbled across your site and thought I'd say something to keep the list active :)

GNU started with RMS. He gave us programming/administration tools to play with.
Linux started with Linus. He released an operating system for us to play with.
You need someone great in the field to release something for everyone to 'play with'. Then people start sending patches...

I know this is chicken-egg, but someone need
papers or the website.

And you might consider merging into the bios
critical mass for things to take off. Consider th
together and make linux work.

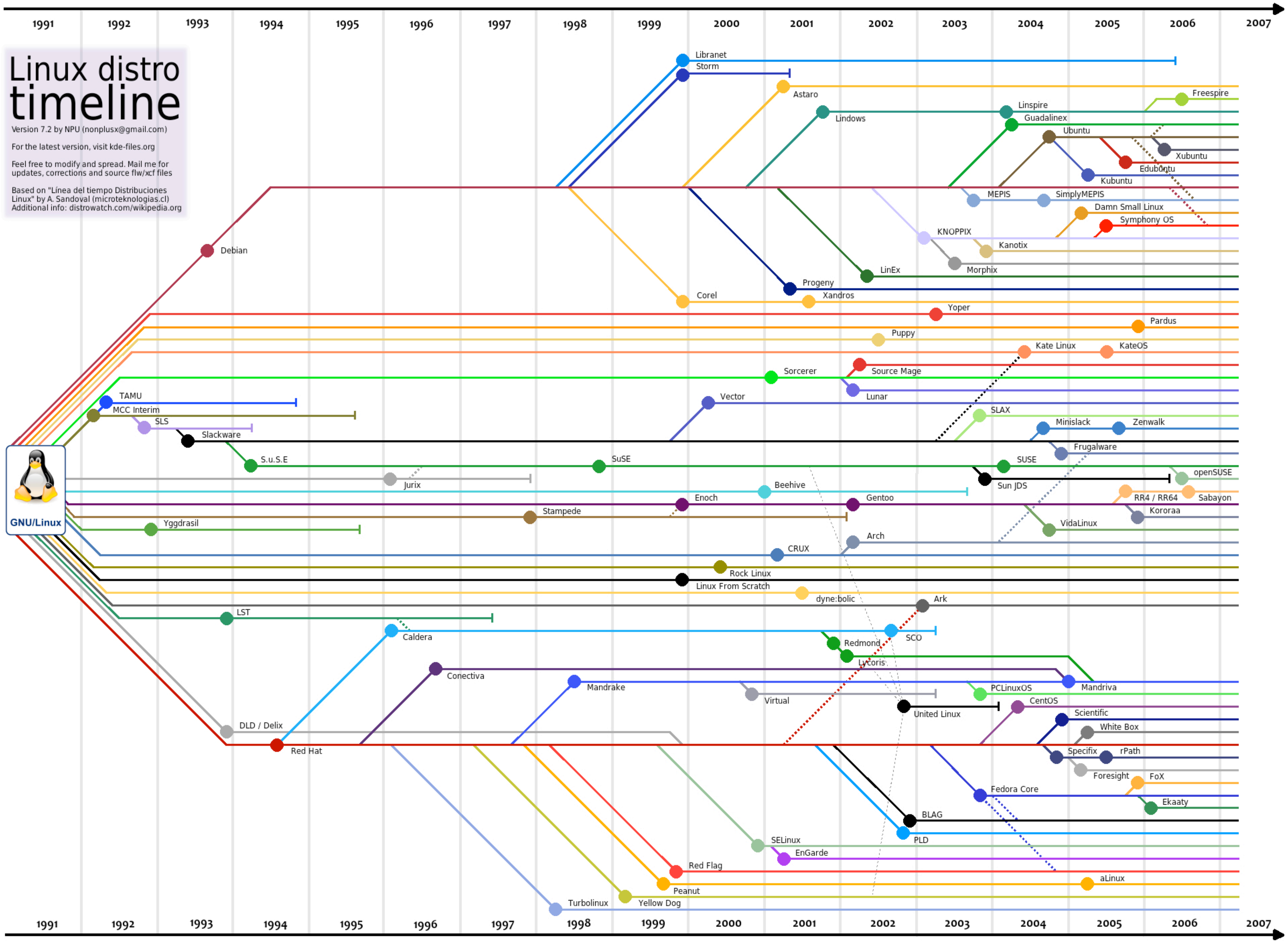
Daniel Amelang

19 Jan 2005

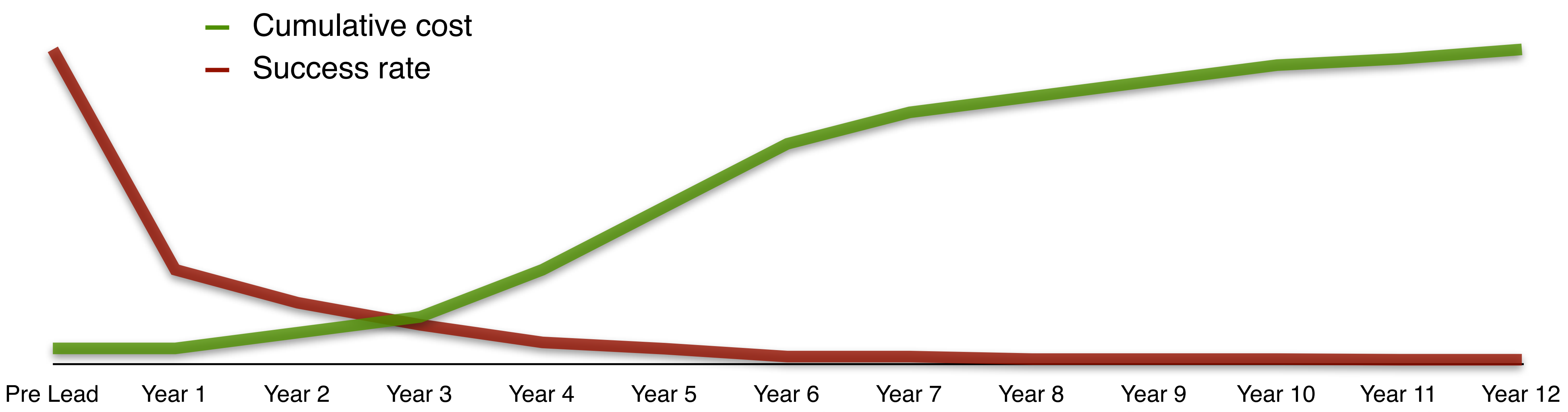
If we do the science well, I'm optimistic that the rest of TDI will fall into place.

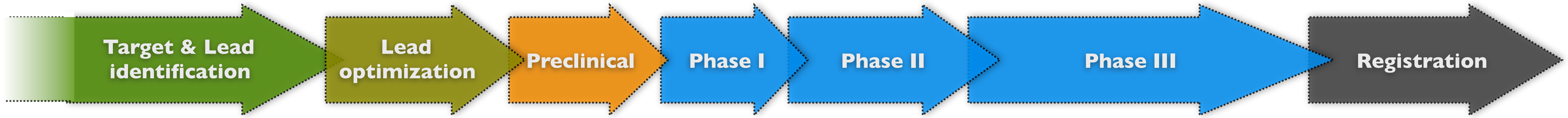
Stephen Mark Maurer

Open Source without a Kernel?

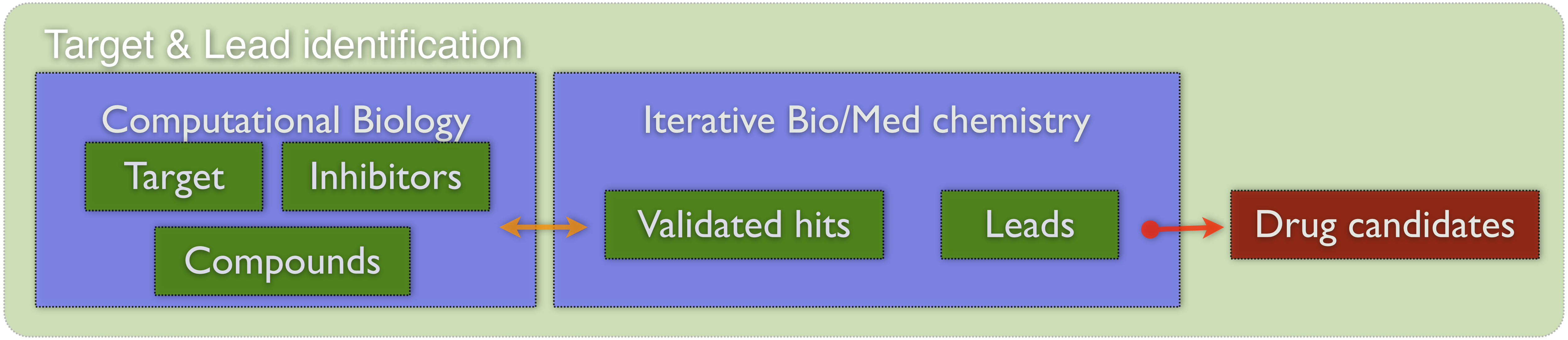


Drug Discovery pipeline





Target & Lead identification → Lead optimization → Preclinical → Phase I → Phase II → Phase III → Registration



Target & Lead identification

Computational Biology

Target

Inhibitors

Compounds

Iterative Bio/Med chemistry

Validated hits

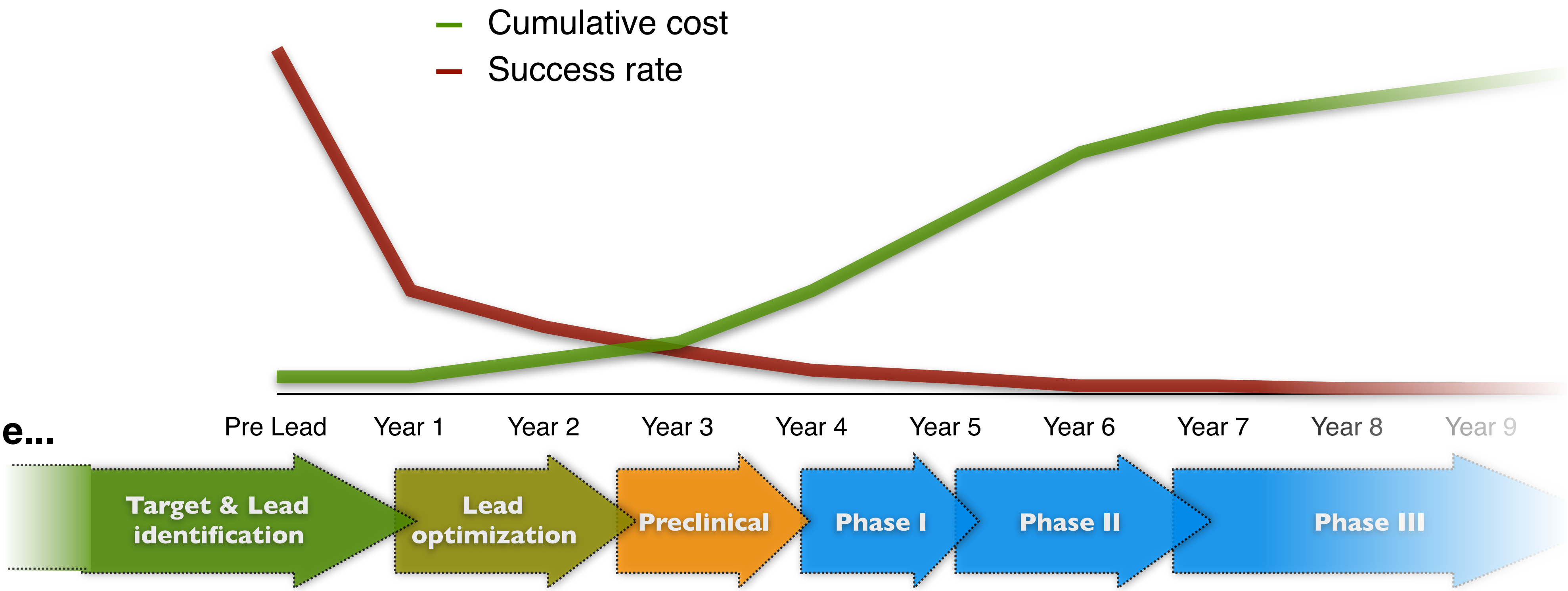
Leads

Drug candidates

Adapted from: - Nwaka & Ridley. (2003) *Nature Reviews. Drug Discovery*. **2**:919
- Austin, Brady, Insel & collins. (2004) *Science*. **306**:1138

Drug Discovery pipeline

shorter time...

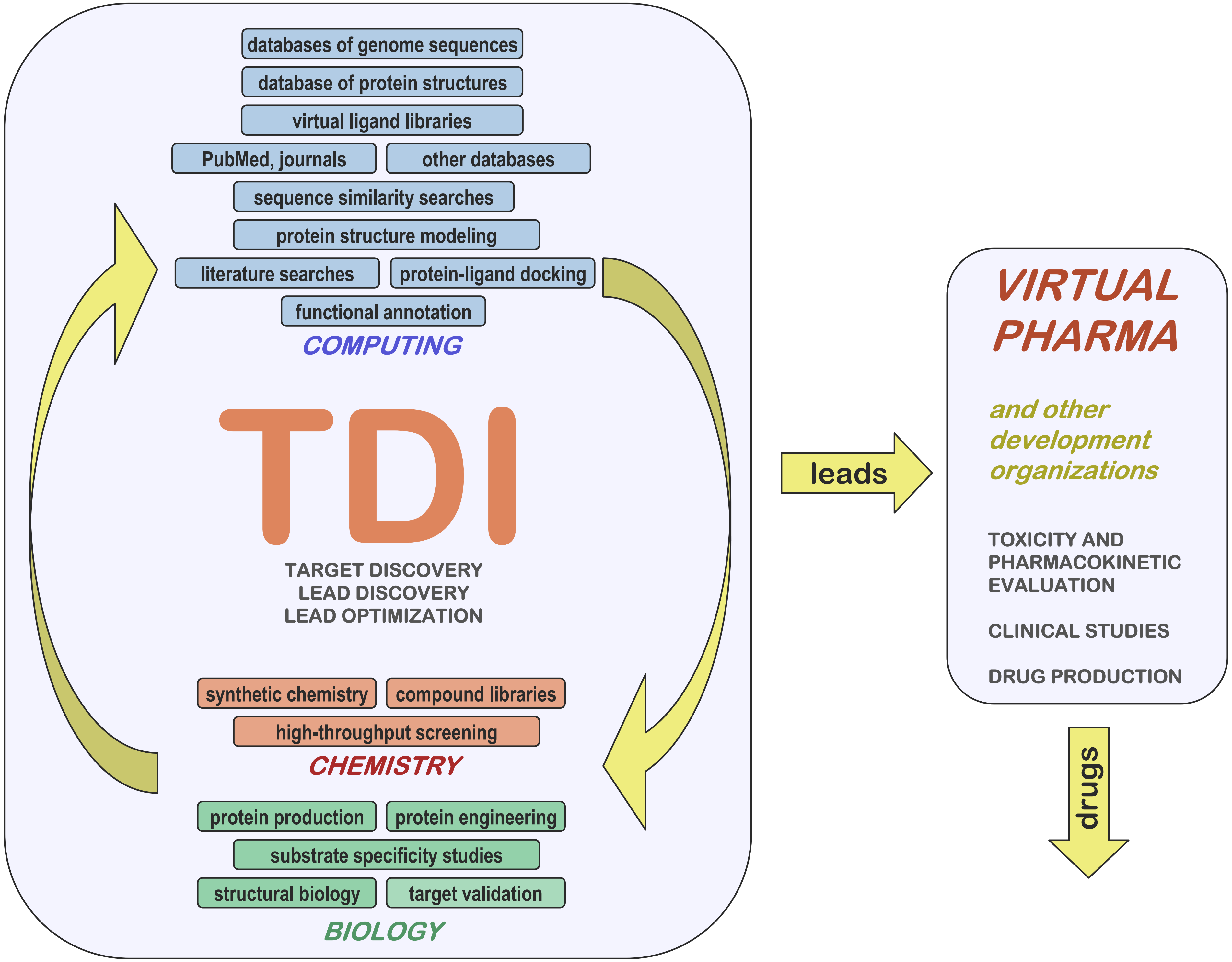


Target & Lead identification → Lead optimization → Preclinical → Phase I → Phase II → Phase III

TDI

+ Completeness of genome projects (*eg*, Malaria)
+ New and more complete biological databases
+ New software and computers (cheaper and faster)
+ Internet == more people == less cost

TDI flowchart



databases of genome sequences

database of protein structures

virtual ligand libraries

PubMed, journals

other databases

sequence similarity searches

protein structure modeling

literature searches

protein-ligand docking

functional annotation

COMPUTING

TDI

TARGET DISCOVERY
LEAD DISCOVERY
LEAD OPTIMIZATION

synthetic chemistry

compound libraries

high-throughput screening

CHEMISTRY

protein production

protein engineering

substrate specificity studies

structural biology

target validation

BIOLOGY

leads

VIRTUAL PHARMA

and other development organizations

TOXICITY AND
PHARMACOKINETIC
EVALUATION

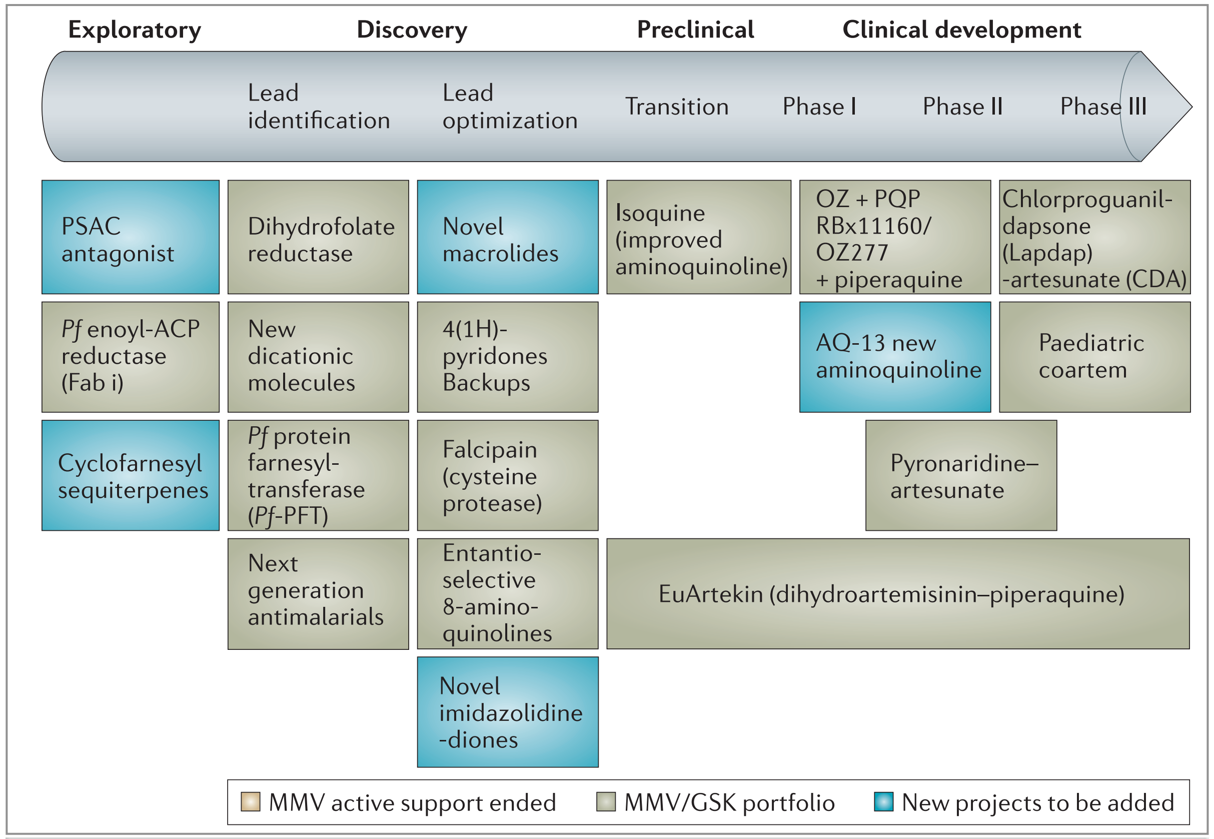
CLINICAL STUDIES

DRUG PRODUCTION

drugs

Non-Profit organizations

Open-Source + Out-Source = low cost business model



Exploratory | Discovery | Preclinical | Clinical development

Lead identification | Lead optimization | Transition | Phase I | Phase II | Phase III

PSAC antagonist

Dihydrofolate reductase

Novel macrolides

Isoquine (improved aminoquinoline)

OZ + PQP RBx11160/ OZ277 + piperaquine

Chlorproguanil-dapsone (Lapdap) -artesunate (CDA)

Pf enoyl-ACP reductase (Fab i)

New dicationic molecules

4(1H)-pyridones Backups

AQ-13 new aminoquinoline

Paediatric coartem

Cyclofarnesyl sequiterpenes

Pf protein farnesyl-transferase (*Pf*-PFT)

Falcipain (cysteine protease)

Pyronaridine–artesunate

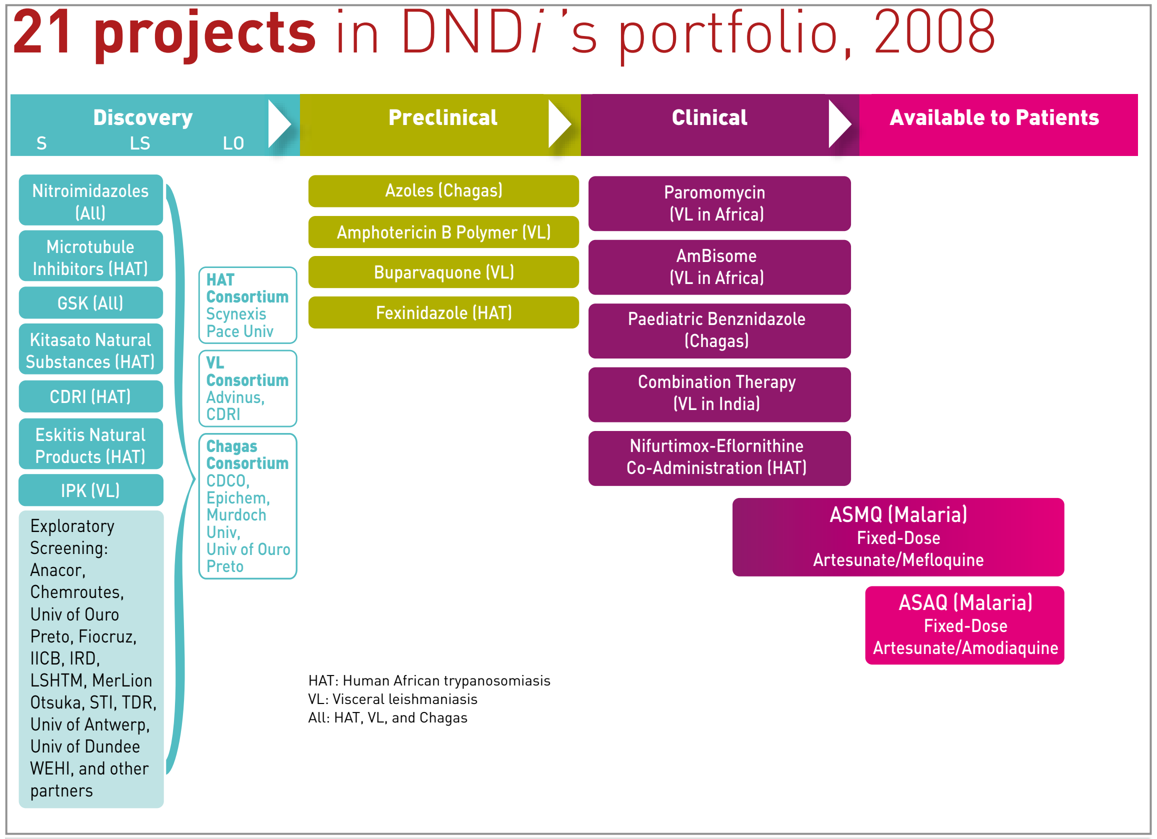
Next generation antimalarials

Entantio-selective 8-amino-quinolines

EuArtekin (dihydroartemisinin–piperaquine)

Novel imidazolidine -diones

□ MMV active support ended ▨ MMV/GSK portfolio 🟦 New projects to be added



21 projects in DNDi's portfolio, 2008

Discovery S LS LO | **Preclinical** | **Clinical** | **Available to Patients**

Nitroimidazoles (All)

Microtubule Inhibitors (HAT)

GSK (All)

Kitasato Natural Substances (HAT)

CDRI (HAT)

Eskitis Natural Products (HAT)

IPK (VL)

Exploratory Screening: Anacor, Chemroutes, Univ of Ouro Preto, Fiocruz, IICB, IRD, LSHTM, MerLion Otsuka, STI, TDR, Univ of Antwerp, Univ of Dundee WEHI, and other partners

HAT Consortium Scynexis Pace Univ

VL Consortium Advinus, CDRI

Chagas Consortium CDCO, Epichem, Murdoch Univ, Univ of Ouro Preto

Azoles (Chagas)

Amphotericin B Polymer (VL)

Buparvaquone (VL)

Fexinidazole (HAT)

Paromomycin (VL in Africa)

AmBisome (VL in Africa)

Paediatric Benznidazole (Chagas)

Combination Therapy (VL in India)

Nifurtimox-Eflornithine Co-Administration (HAT)

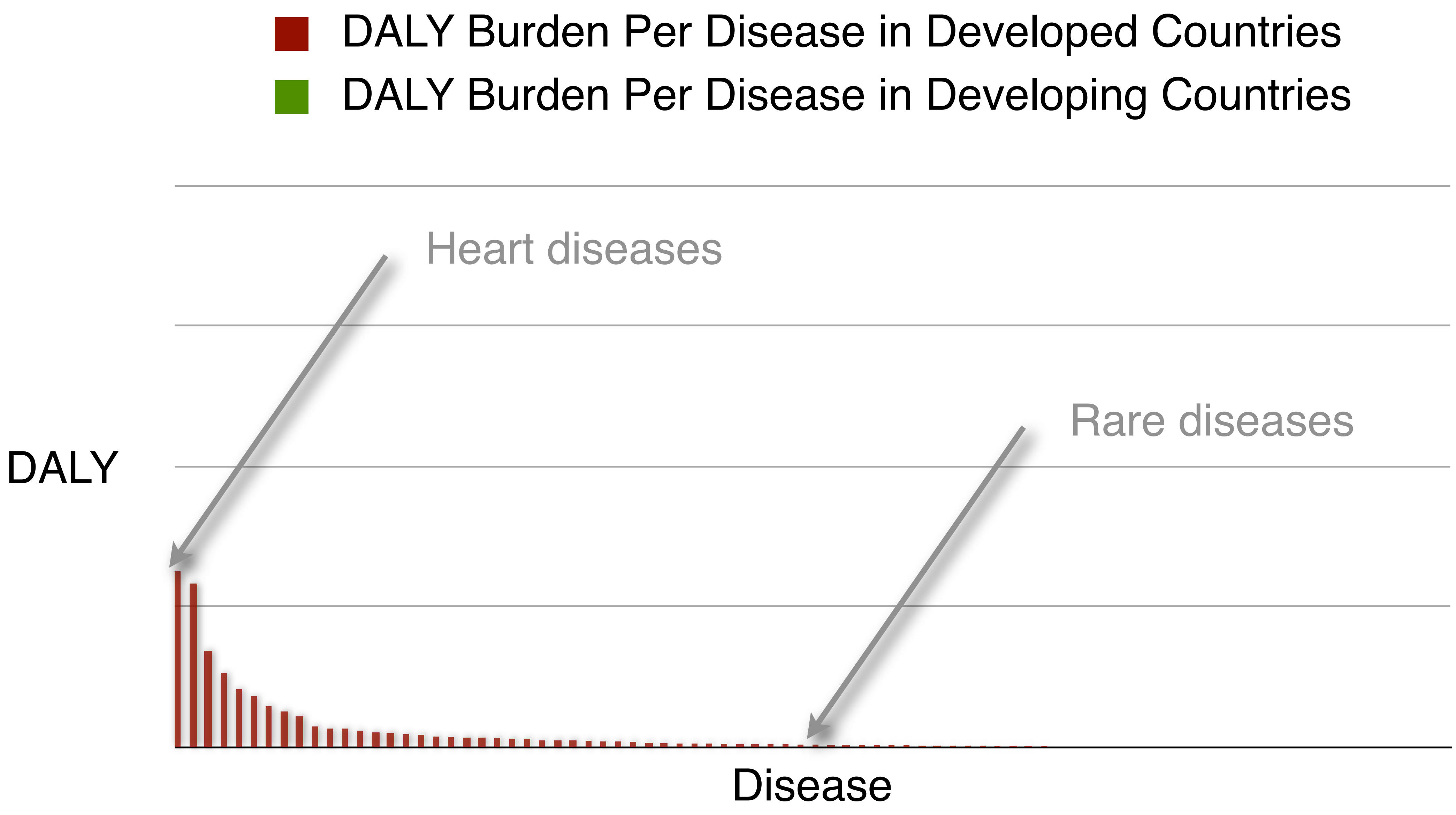
ASMQ (Malaria) Fixed-Dose Artesunate/Mefloquine

ASAQ (Malaria) Fixed-Dose Artesunate/Amodiaquine

HAT: Human African trypanosomiasis
VL: Visceral leishmaniasis
All: HAT, VL, and Chagas

Munos (2006) Nature Reviews. Drug Discovery.

Need is High in the Tail



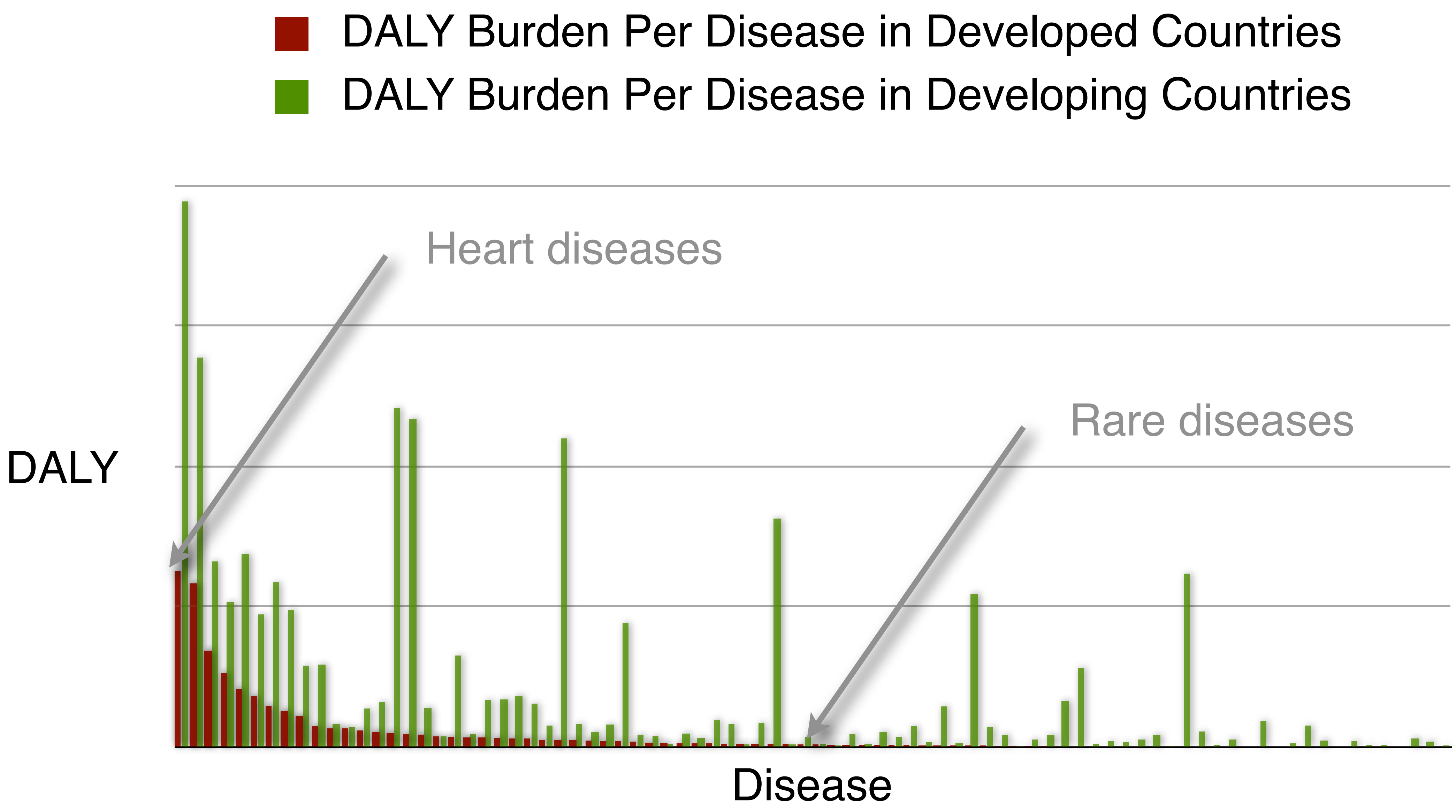
Disease data taken from WHO, *World Health Report 2004*

DALY - Disability adjusted life years

DALY is not a perfect measure of market size, but is certainly a good measure for importance.

DALYs for a disease are the sum of the years of life lost due to premature mortality (YLL) in the population and the years lost due to disability (YLD) for incident cases of the health condition. The DALY is a health gap measure that extends the concept of potential years of life lost due to premature death (PYLL) to include equivalent years of 'healthy' life lost in states of less than full health, broadly termed disability. One DALY represents the loss of one year of equivalent full health.

Need is High in the Tail



Disease data taken from WHO, *World Health Report 2004*

DALY - Disability adjusted life years

DALY is not a perfect measure of market size, but is certainly a good measure for importance.

DALYs for a disease are the sum of the years of life lost due to premature mortality (YLL) in the population and the years lost due to disability (YLD) for incident cases of the health condition. The DALY is a health gap measure that extends the concept of potential years of life lost due to premature death (PYLL) to include equivalent years of 'healthy' life lost in states of less than full health, broadly termed disability. One DALY represents the loss of one year of equivalent full health.

“Unprofitable” Diseases and Global DALY (in 1000’s)

Malaria*	**46,486**
Tetanus	7,074
Lymphatic filariasis*	**5,777**
Syphilis	4,200
Trachoma	2,329
Leishmaniasis*	**2,090**
Ascariasis	1,817
Schistosomiasis*	**1,702**
Trypanosomiasis*	**1,525**

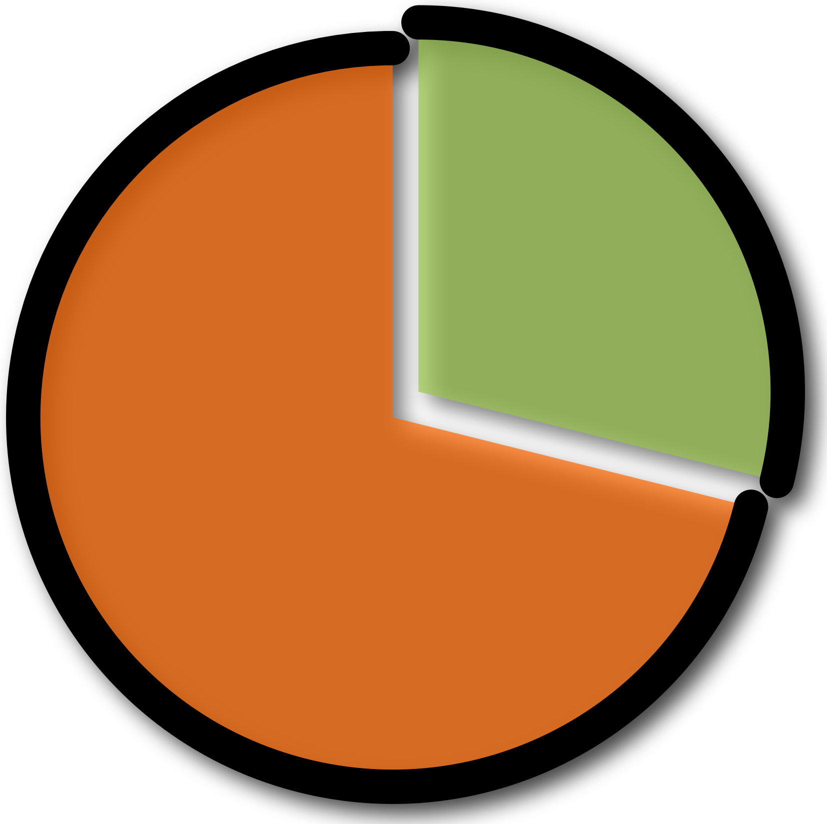
Trichuriasis	1,006
Japanese encephalitis	709
Chagas Disease*	**667**
Dengue*	**616**
Onchocerciasis*	**484**
Leprosy*	**199**
Diphtheria	185
Poliomyelitise	151
Hookworm disease	59

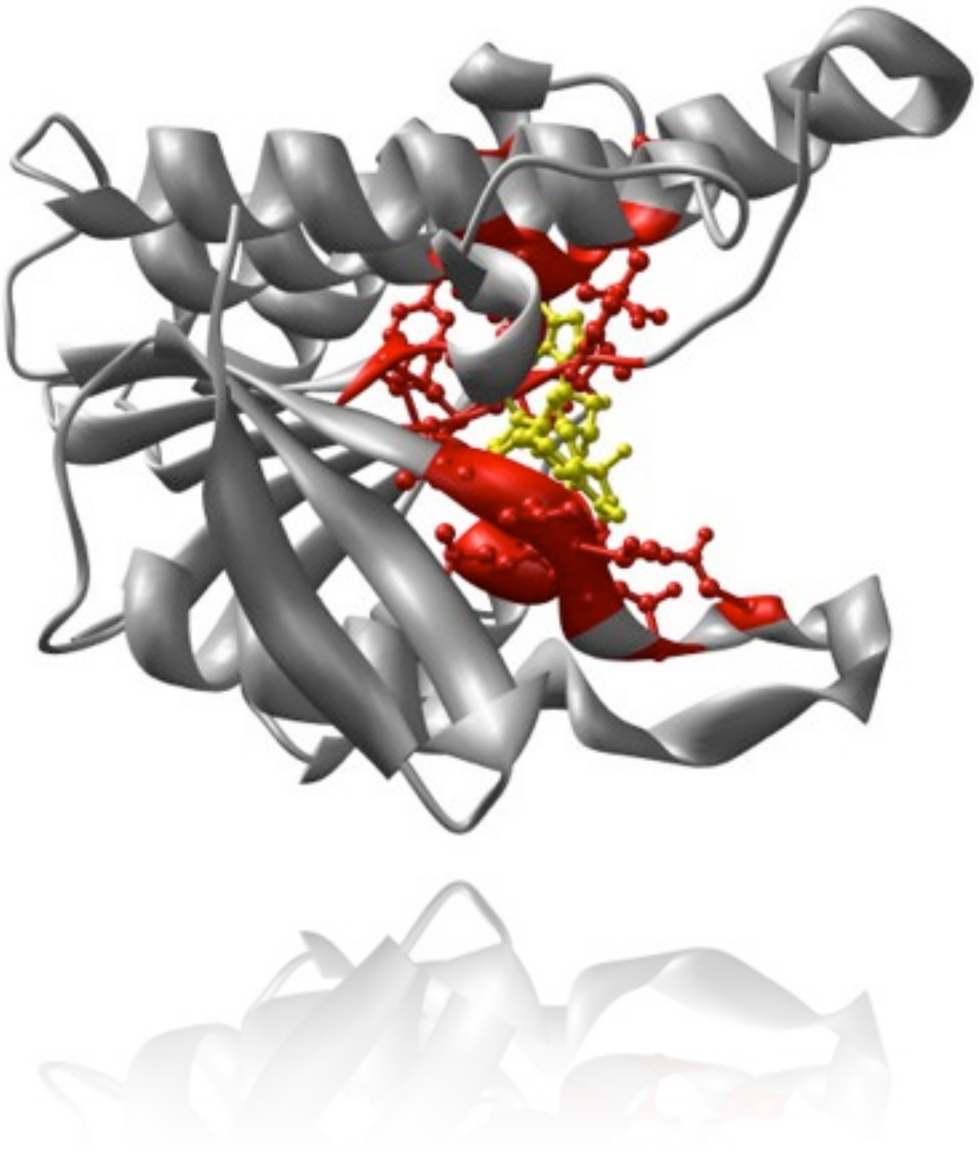
Disease data taken from WHO, *World Health Report 2004*

DALY - Disability adjusted life year in 1000’s.

* Officially listed in the WHO Tropical Disease Research disease portfolio.

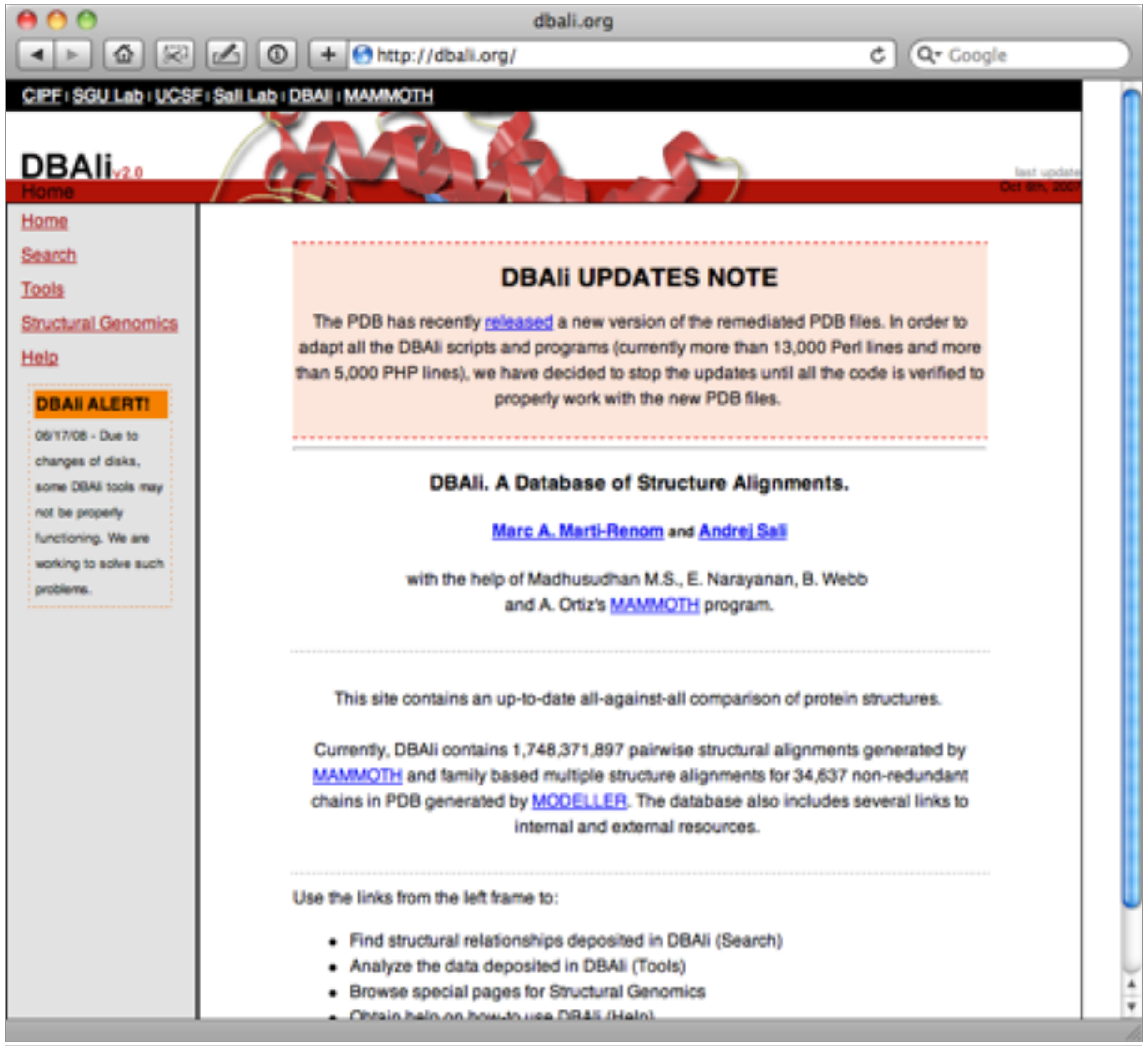
Predicting binding sites in protein structure models.





DBAli$_{v2.0}$ database

http://www.dbali.org



DBAli ALERT!

06/17/08 - Due to changes of disks, some DBAli tools may not be properly functioning. We are working to solve such problems.

DBAli UPDATES NOTE

The PDB has recently released a new version of the remediated PDB files. In order to adapt all the DBAli scripts and programs (currently more than 13,000 Perl lines and more than 5,000 PHP lines), we have decided to stop the updates until all the code is verified to properly work with the new PDB files.

DBAli. A Database of Structure Alignments.

Marc A. Marti-Renom and Andrej Sali

with the help of Madhusudhan M.S., E. Narayanan, B. Webb and A. Ortiz's MAMMOTH program.

This site contains an up-to-date all-against-all comparison of protein structures.

Currently, DBAli contains 1,748,371,897 pairwise structural alignments generated by MAMMOTH and family based multiple structure alignments for 34,637 non-redundant chains in PDB generated by MODELLER. The database also includes several links to internal and external resources.

Use the links from the left frame to:

- Find structural relationships deposited in DBAli (Search)
- Analyze the data deposited in DBAli (Tools)
- Browse special pages for Structural Genomics
- Obtain help on how-to use DBAli (Help)

✓ Fully-automatic
✓ Data is kept up-to-date with PDB releases
✓ Tools for "on the fly" classification of families.
✓ Easy to navigate
✓ Provides tools for structure analysis

Marti-Renom et al. 2001. Bioinformatics. 17, 746

Does not provide a stable classification similar to that of CATH or SCOP

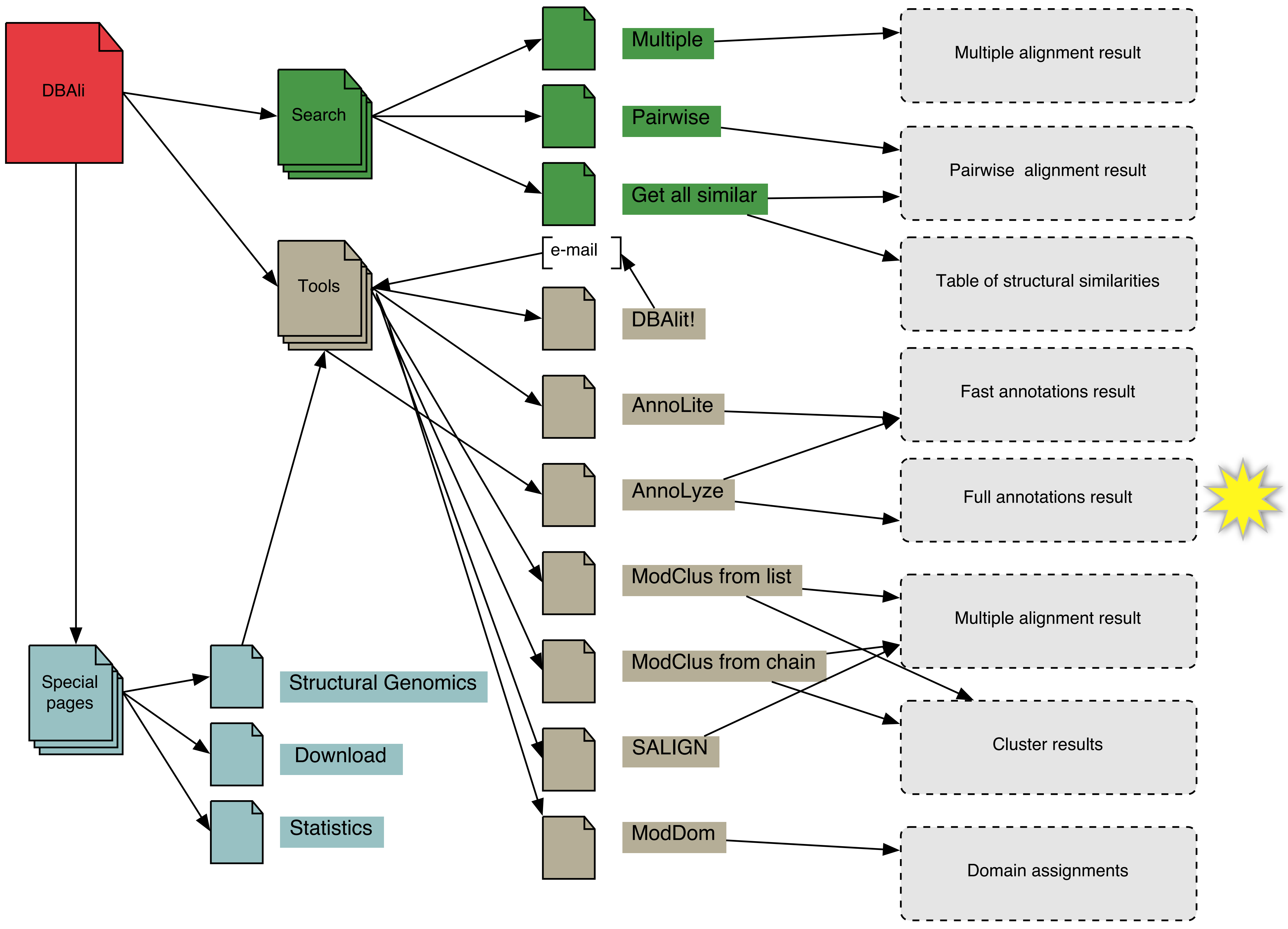
Uses MAMMOTH for similarity detection

✓ VERY FAST!!!
✓ Good scoring system with significance

Ortiz AR, (2002) Protein Sci. 11 pp2606

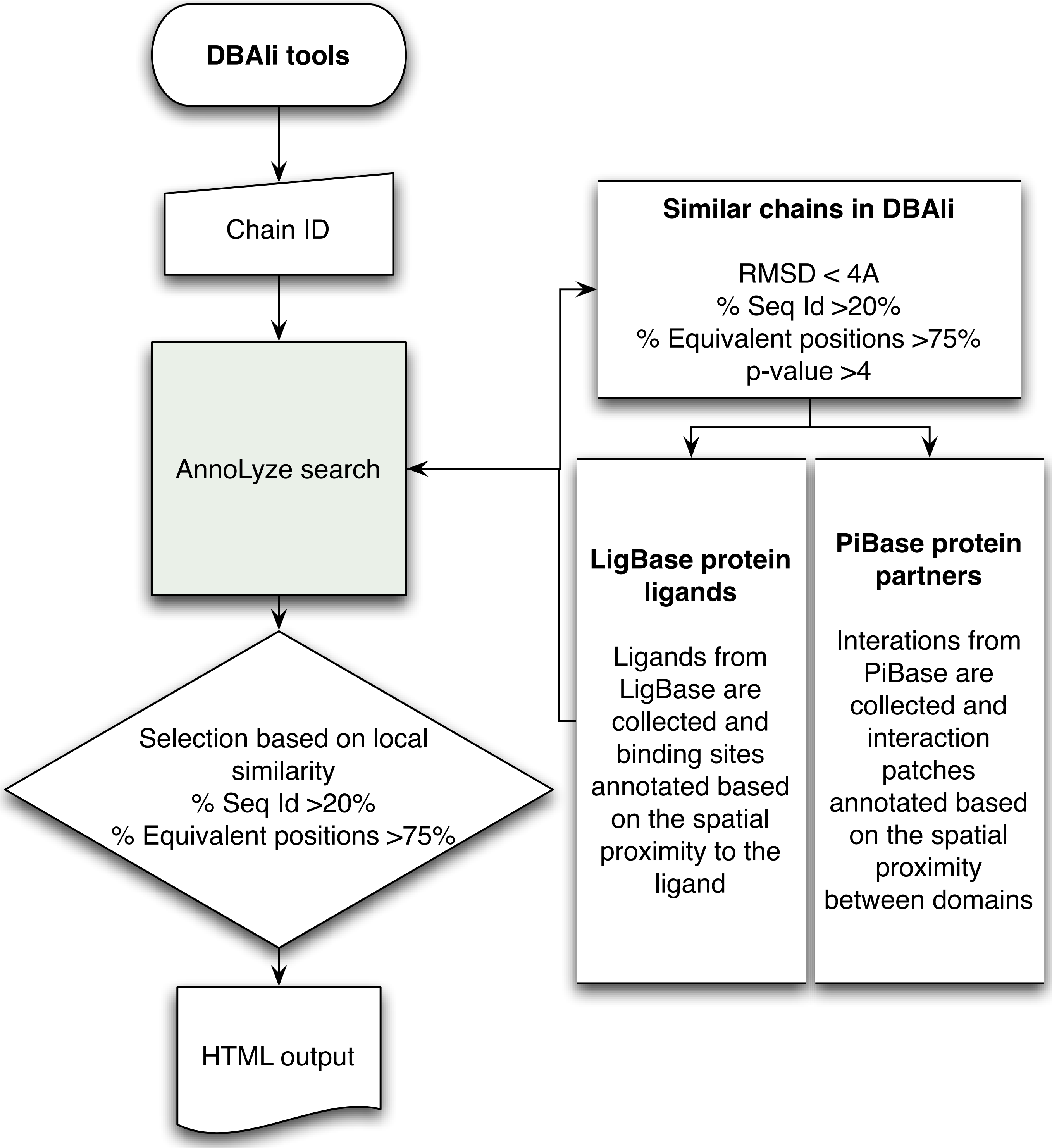
DBAli$_{v2.0}$ database

http://www.dbali.org



Marti-Renom et al. BMC Bioinformatics (2007) Volume 8. Suppl S4

Method



DBAli tools

Chain ID

AnnoLyze search

Selection based on local similarity
% Seq Id >20%
% Equivalent positions >75%

HTML output

Similar chains in DBAli

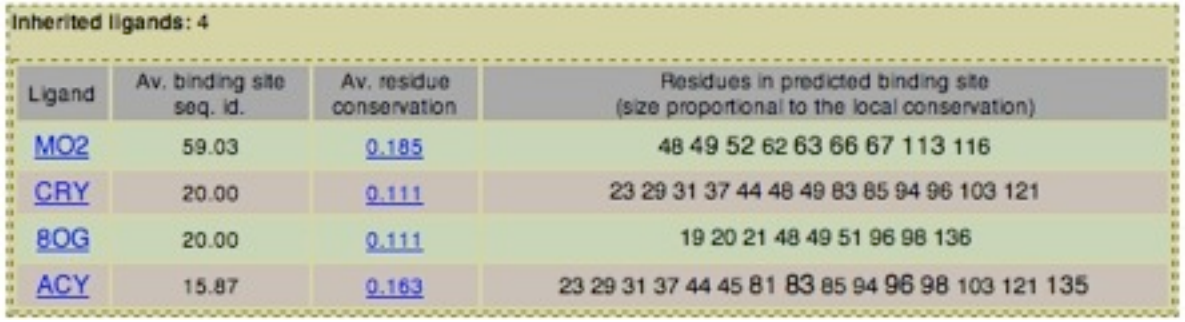
RMSD < 4A
% Seq Id >20%
% Equivalent positions >75%
p-value >4

LigBase protein ligands

Ligands from LigBase are collected and binding sites annotated based on the spatial proximity to the ligand

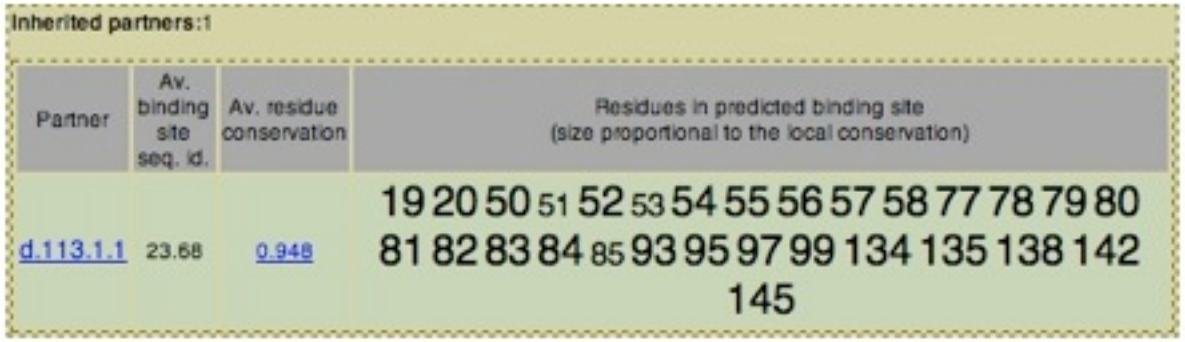
PiBase protein partners

Interations from PiBase are collected and interaction patches annotated based on the spatial proximity between domains



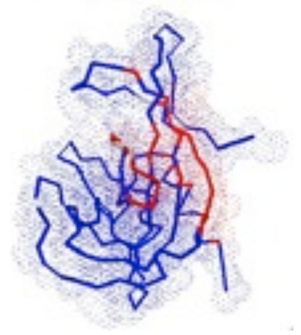
Inherited ligands: 4

Ligand	Av. binding site seq. id.	Av. residue conservation	Residues in predicted binding site (size proportional to the local conservation)
MO2	59.03	0.185	48 49 52 62 63 66 67 113 116
CRY	20.00	0.111	23 29 31 37 44 48 49 83 85 94 96 103 121
8OG	20.00	0.111	19 20 21 48 49 51 96 98 136
ACY	15.87	0.163	23 29 31 37 44 45 81 83 85 94 96 98 103 121 135



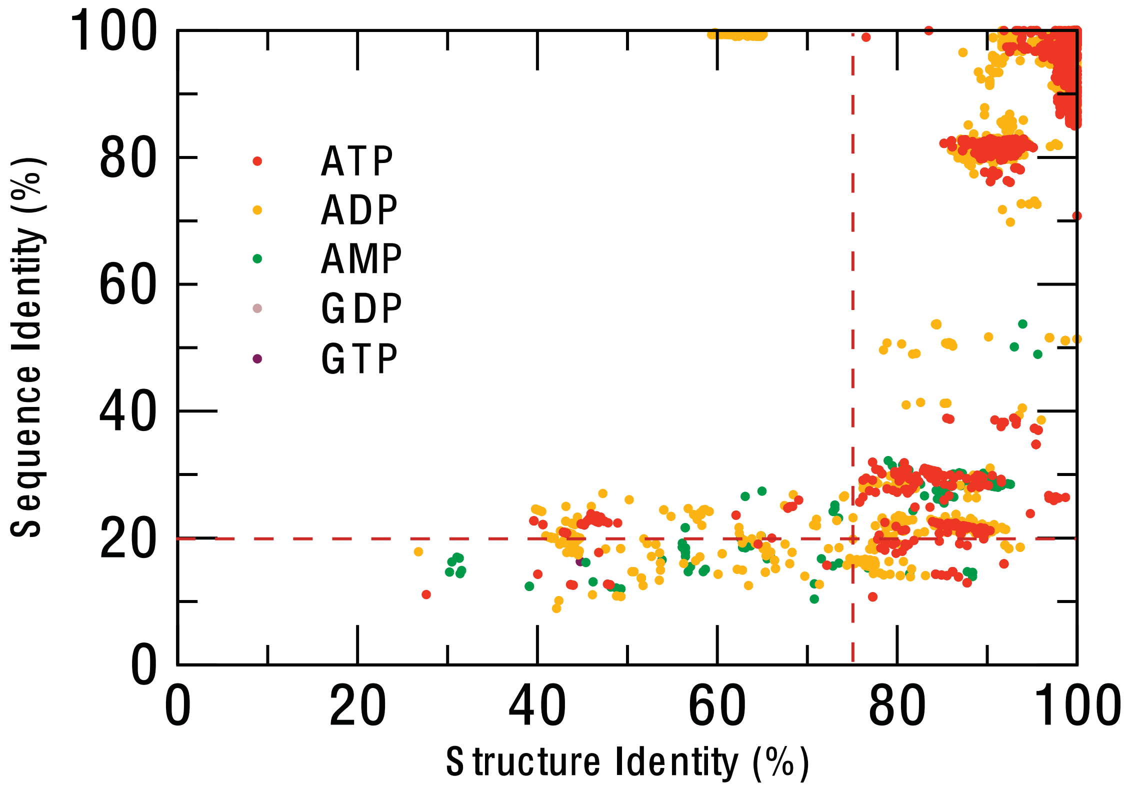
Inherited partners:1

Partner	Av. binding site seq. id.	Av. residue conservation	Residues in predicted binding site (size proportional to the local conservation)
d.113.1.1	23.68	0.948	19 20 50 51 52 53 54 55 56 57 58 77 78 79 80 81 82 83 84 85 93 95 97 99 134 135 138 142 145

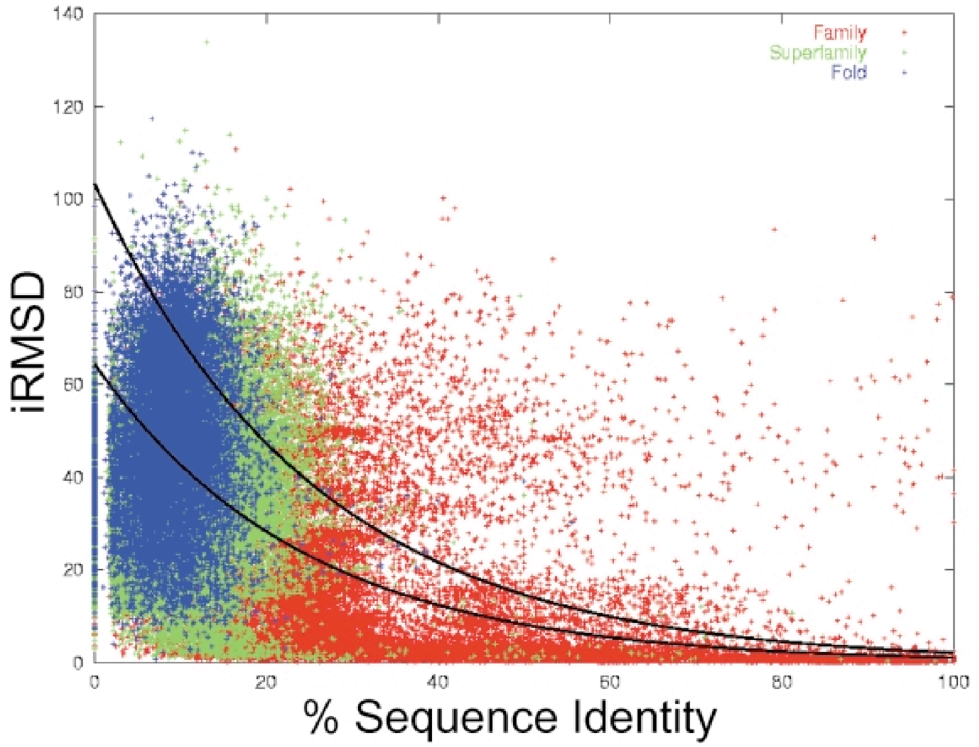


Scoring function

Ligands



Partners



Aloy *et al.* (2003) J.Mol.Biol. 332(5):989-98.

Benchmark

	Number of chains
Initial set*	78,167
LigBase**	30,126
Non-redundant set***	4,948 (8,846 ligands)

**all PDB chains larger than 30 aminoacids in length (8th of August, 2006)*

***annotated with at least one ligand in the LigBase database*

****not two chains can be structurally aligned within 3A, superimposing more than 75% of their Cα atoms, result in a sequence alignment with more than 30% identity, and have a length difference inferior to 50aa*

Sensitivity .vs. Precision

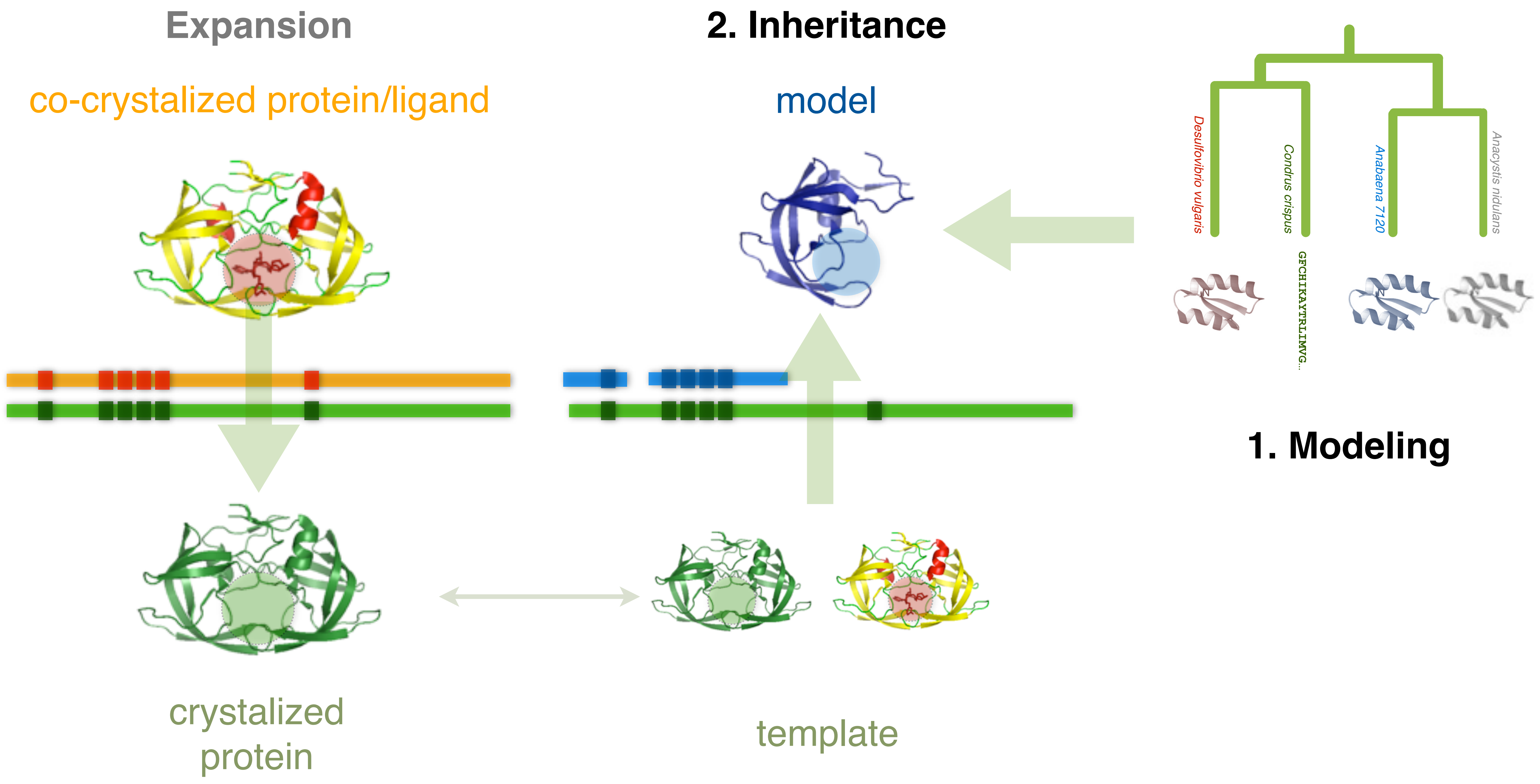
	Optimal cut-off	Sensitivity (%) Recall or TPR	Precision (%)
Ligands	30%	71.9	13.7

$$\text{Sensitivity} = \frac{TP}{TP + FN} \qquad \text{Precision} = \frac{TP}{TP + FP}$$

Marti-Renom et al. BMC Bioinformatics (2007) Volume 8. Suppl S4

~90-95% of residues correctly predicted

Comparative docking



Expansion

co-crystalized protein/ligand

crystalized protein

2. Inheritance

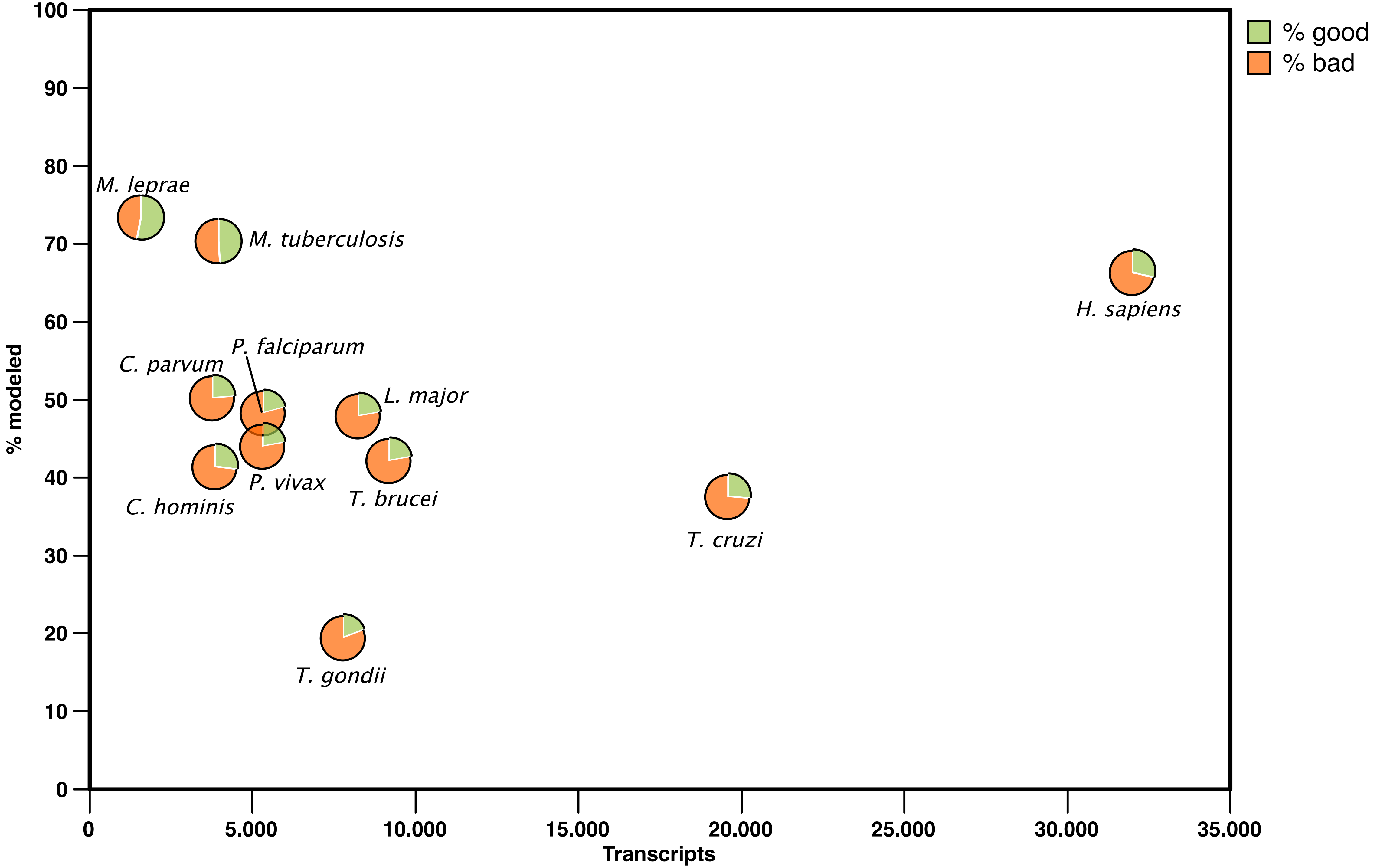
model

template

1. Modeling

Modeling Genomes

data from models generated by ModPipe (Eswar, Pieper & Sali)



A good model has MPQS of 1.0 or higher

Summary table

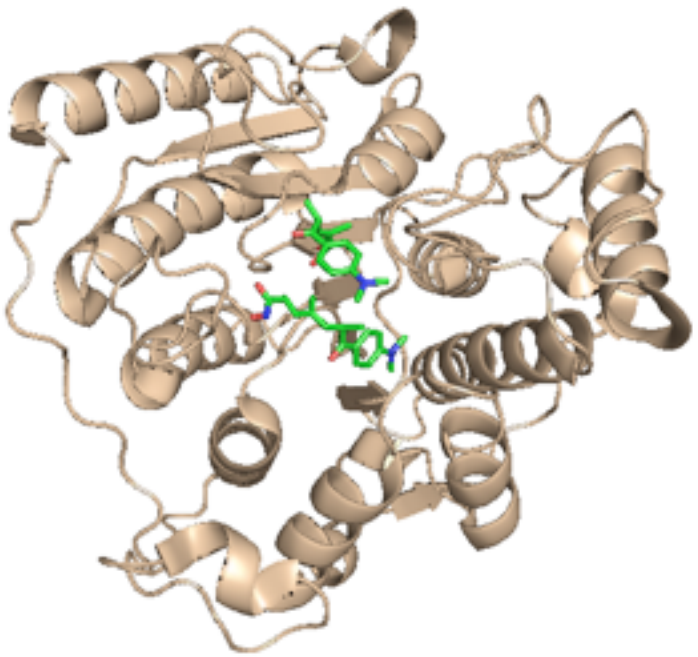
models with inherited ligands

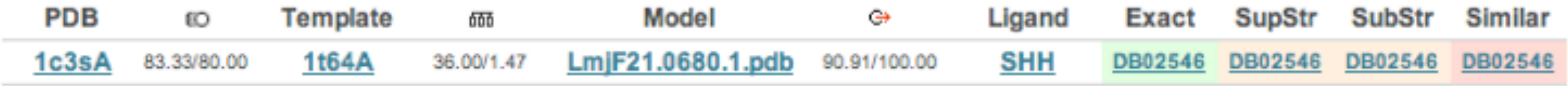
29,271 targets with good models, 297 inherited a ligand/substance similar to a known drug in DrugBank

	Transcripts	Modeled targets	Selected models	Inherited ligands	Similar to a drug	Drugs
C. hominis	3,886	1,614	666	197	20	13
C. parvum	3,806	1,918	742	232	24	13
L. major	8,274	3,975	1,409	478	43	20
M. leprae	1,605	1,178	893	310	25	6
M. tuberculosis	3,991	2,808	1,608	365	30	10
P. falciparum	5,363	2,599	818	284	28	13
P. vivax	5,342	2,359	822	268	24	13
T. brucei	7,793	1,530	300	138	13	6
T. cruzi	19,607	7,390	3,070	769	51	28
T. gondii	9,210	3,900	1,386	458	39	21
TOTAL	**68,877**	**29,271**	**11,714**	**3,499**	**297**	**143**

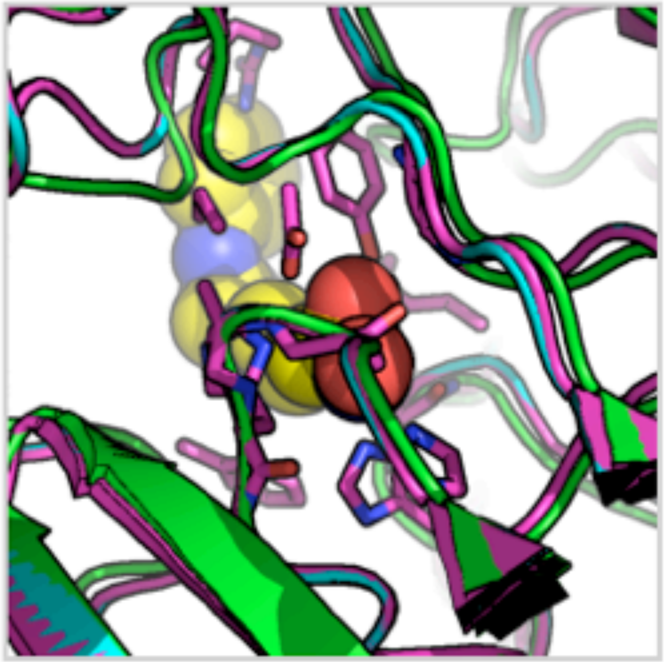
L. major Histone deacetylase 2 + Vorinostat

Template 1t64A a human HDAC8 protein.

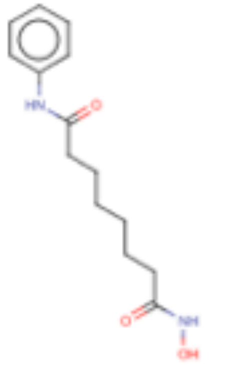




PDB		Template		Model		Ligand	Exact	SupStr	SubStr	Similar
1c3sA	83.33/80.00	1t64A	36.00/1.47	LmjF21.0680.1.pdb	90.91/100.00	SHH	DB02546	DB02546	DB02546	DB02546



DB02546 Vorinostat



Small Molecule; Approved; Investigational

Drug categories:

Anti-Inflammatory Agents, Non-Steroidal

Anticarcinogenic Agents

Antineoplastic Agents

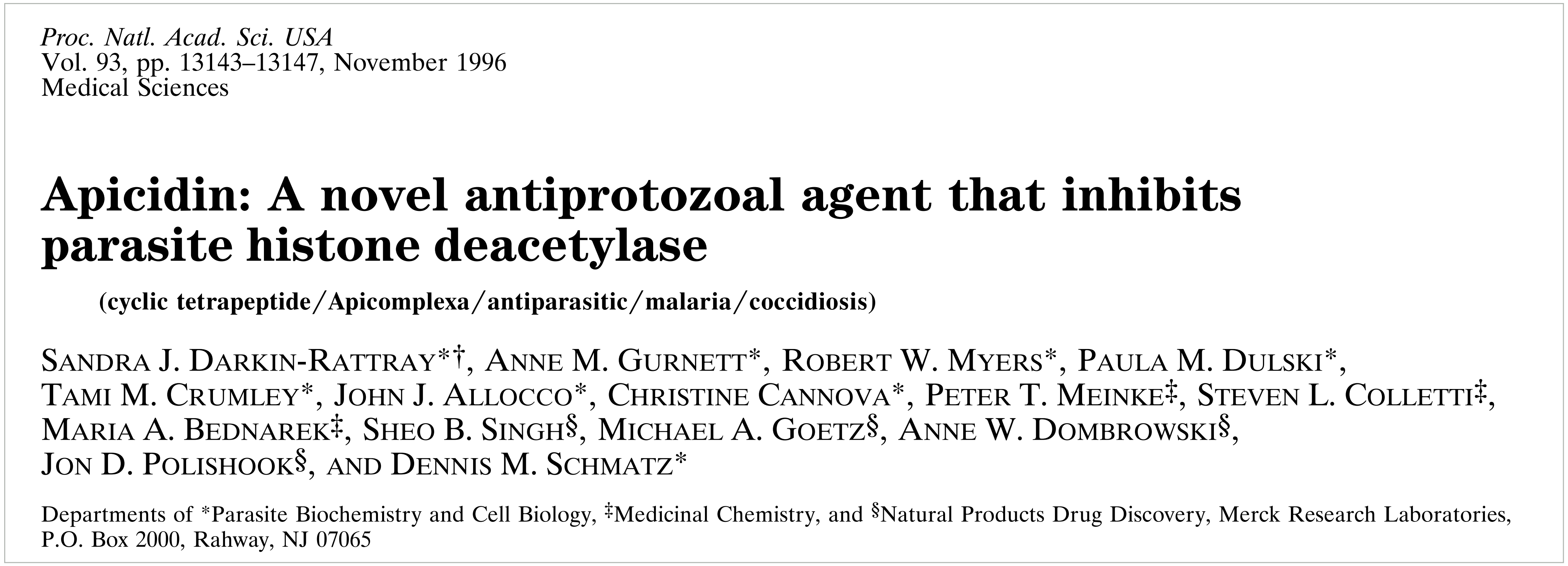
Enzyme Inhibitors

Drug indication:

For the treatment of cutaneous manifestations in patients with cutaneous T-cell lymphoma who have progressive, persistent or recurrent disease on or following two systemic therapies.

L. major Histone deacetylase 2 + Vorinostat

Literature



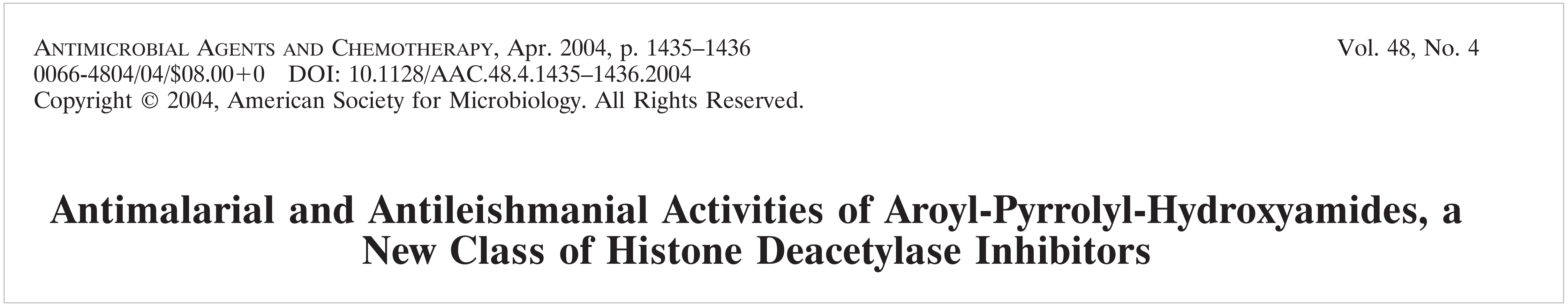
Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 13143–13147, November 1996
Medical Sciences

Apicidin: A novel antiprotozoal agent that inhibits parasite histone deacetylase

(cyclic tetrapeptide/Apicomplexa/antiparasitic/malaria/coccidiosis)

SANDRA J. DARKIN-RATTRAY*†, ANNE M. GURNETT*, ROBERT W. MYERS*, PAULA M. DULSKI*, TAMI M. CRUMLEY*, JOHN J. ALLOCCO*, CHRISTINE CANNOVA*, PETER T. MEINKE‡, STEVEN L. COLLETTI‡, MARIA A. BEDNAREK‡, SHEO B. SINGH§, MICHAEL A. GOETZ§, ANNE W. DOMBROWSKI§, JON D. POLISHOOK§, AND DENNIS M. SCHMATZ*

Departments of *Parasite Biochemistry and Cell Biology, ‡Medicinal Chemistry, and §Natural Products Drug Discovery, Merck Research Laboratories, P.O. Box 2000, Rahway, NJ 07065

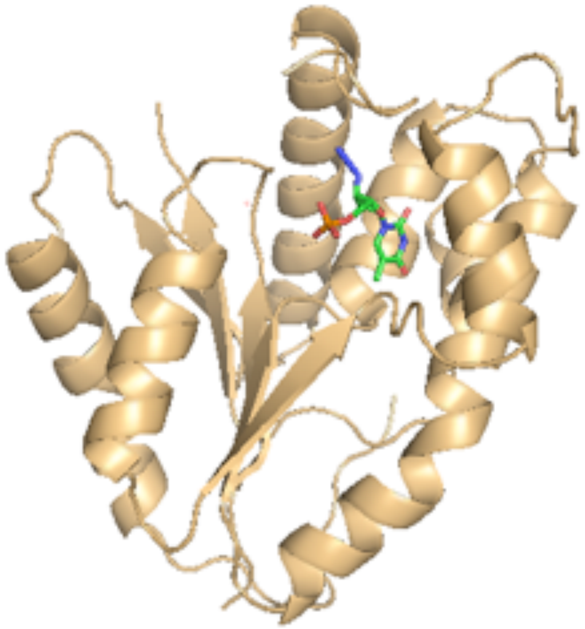


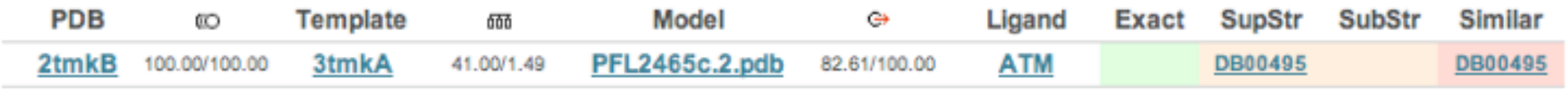
ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Apr. 2004, p. 1435–1436 Vol. 48, No. 4
0066-4804/04/$08.00+0 DOI: 10.1128/AAC.48.4.1435–1436.2004
Copyright © 2004, American Society for Microbiology. All Rights Reserved.

Antimalarial and Antileishmanial Activities of Aroyl-Pyrrolyl-Hydroxyamides, a New Class of Histone Deacetylase Inhibitors

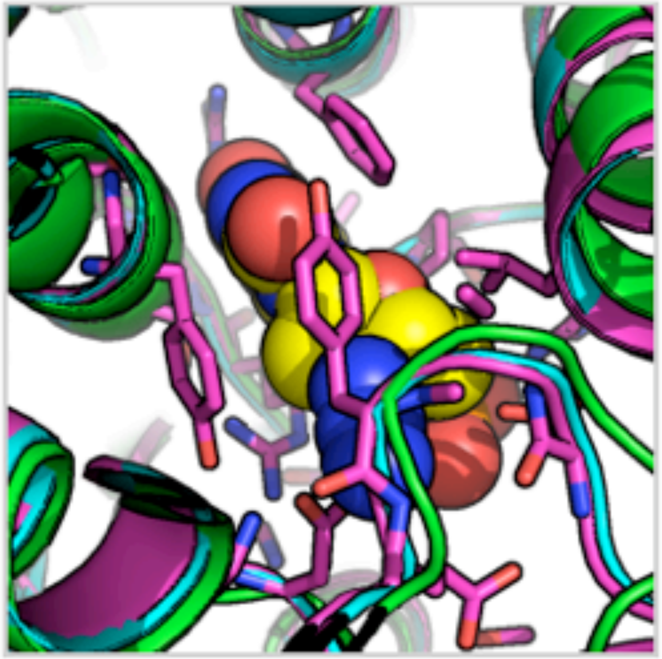
P. falciparum tymidylate kinase + zidovudine

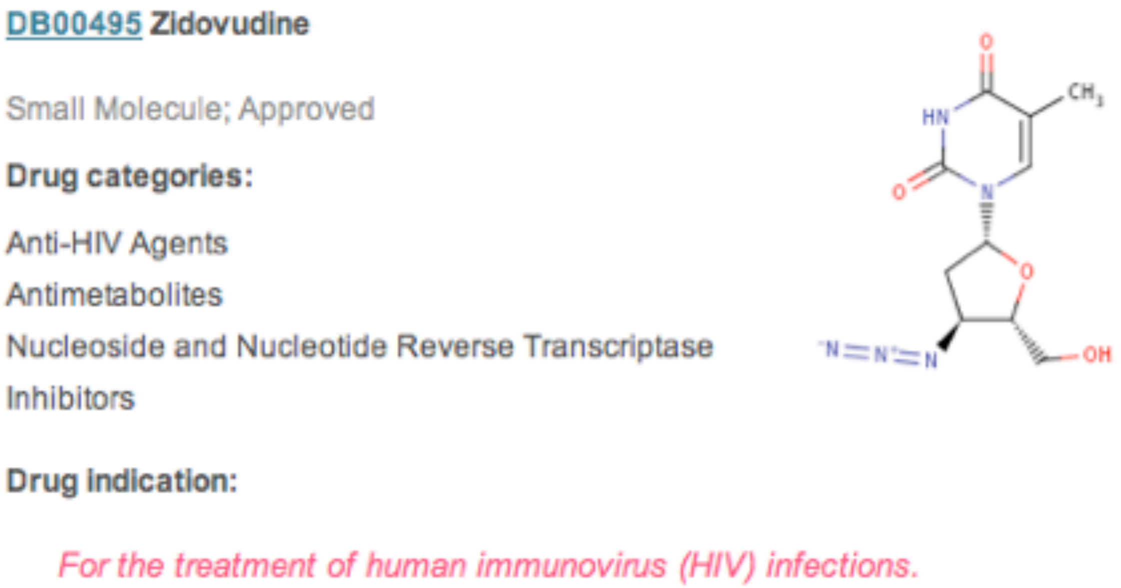
Template 3tmkA a yeast tymidylate kinase.





PDB		Template		Model		Ligand	Exact	SupStr	SubStr	Similar
2tmkB	100.00/100.00	3tmkA	41.00/1.49	PFL2465c.2.pdb	82.61/100.00	ATM		DB00495		DB00495





DB00495 Zidovudine

Small Molecule; Approved

Drug categories:

Anti-HIV Agents

Antimetabolites

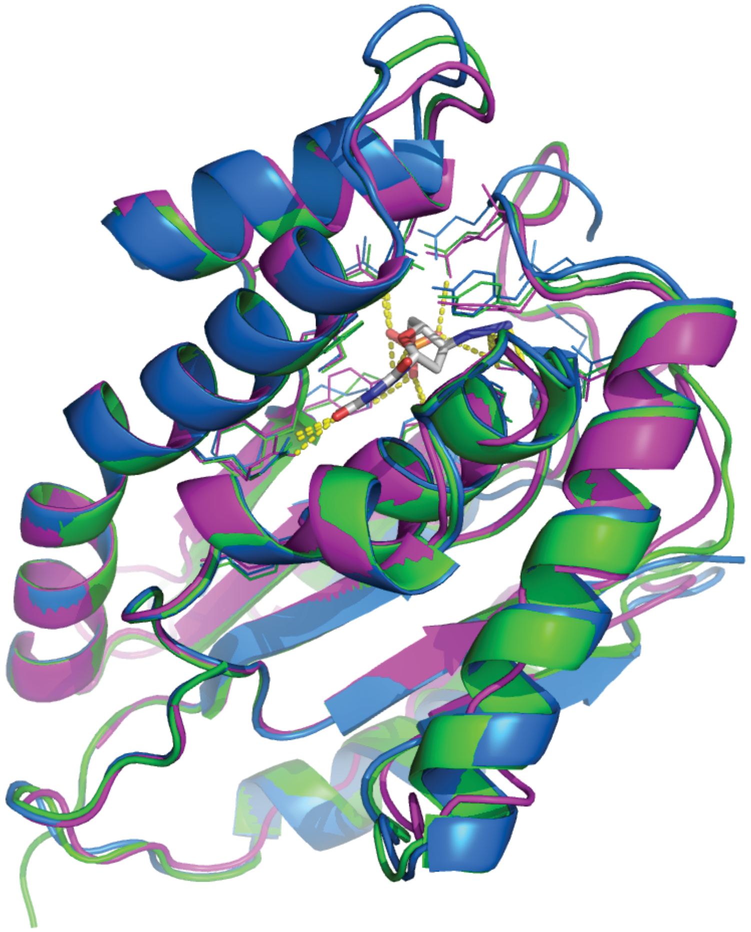
Nucleoside and Nucleotide Reverse Transcriptase Inhibitors

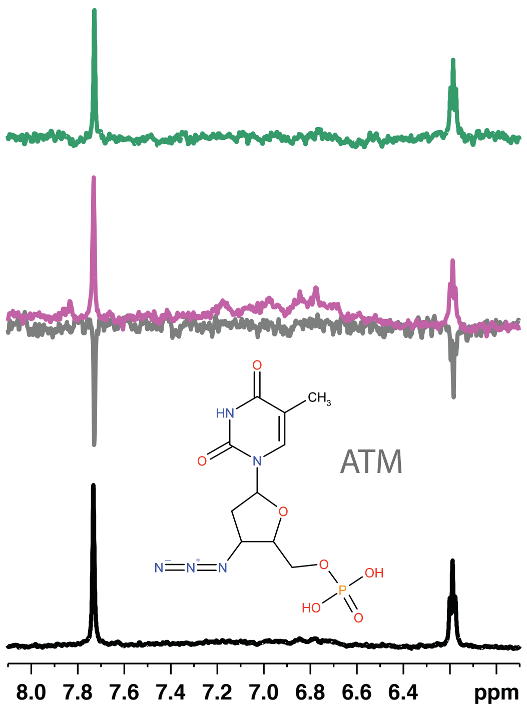
Drug Indication:

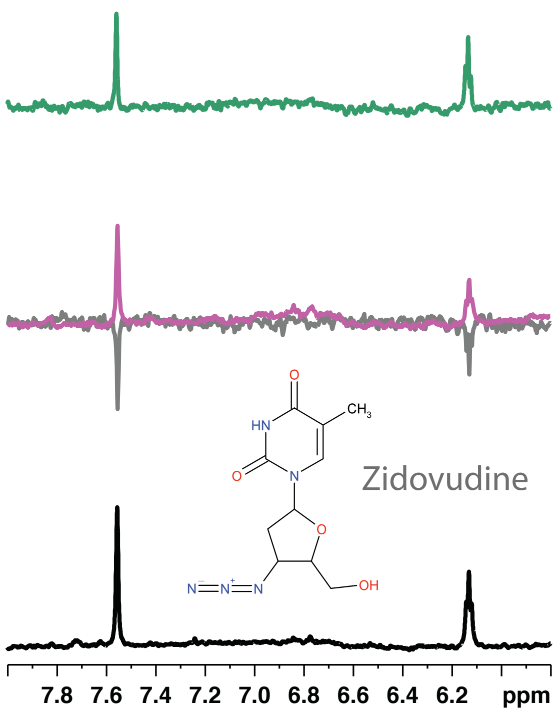
For the treatment of human immunovirus (HIV) infections.

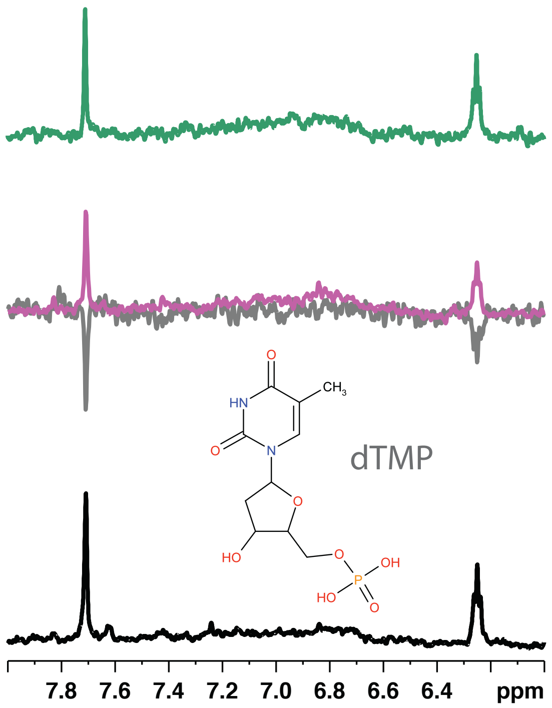
P. falciparum thymidylate kinase + zidovudine

NMR Water-LOGSY and STD experiments





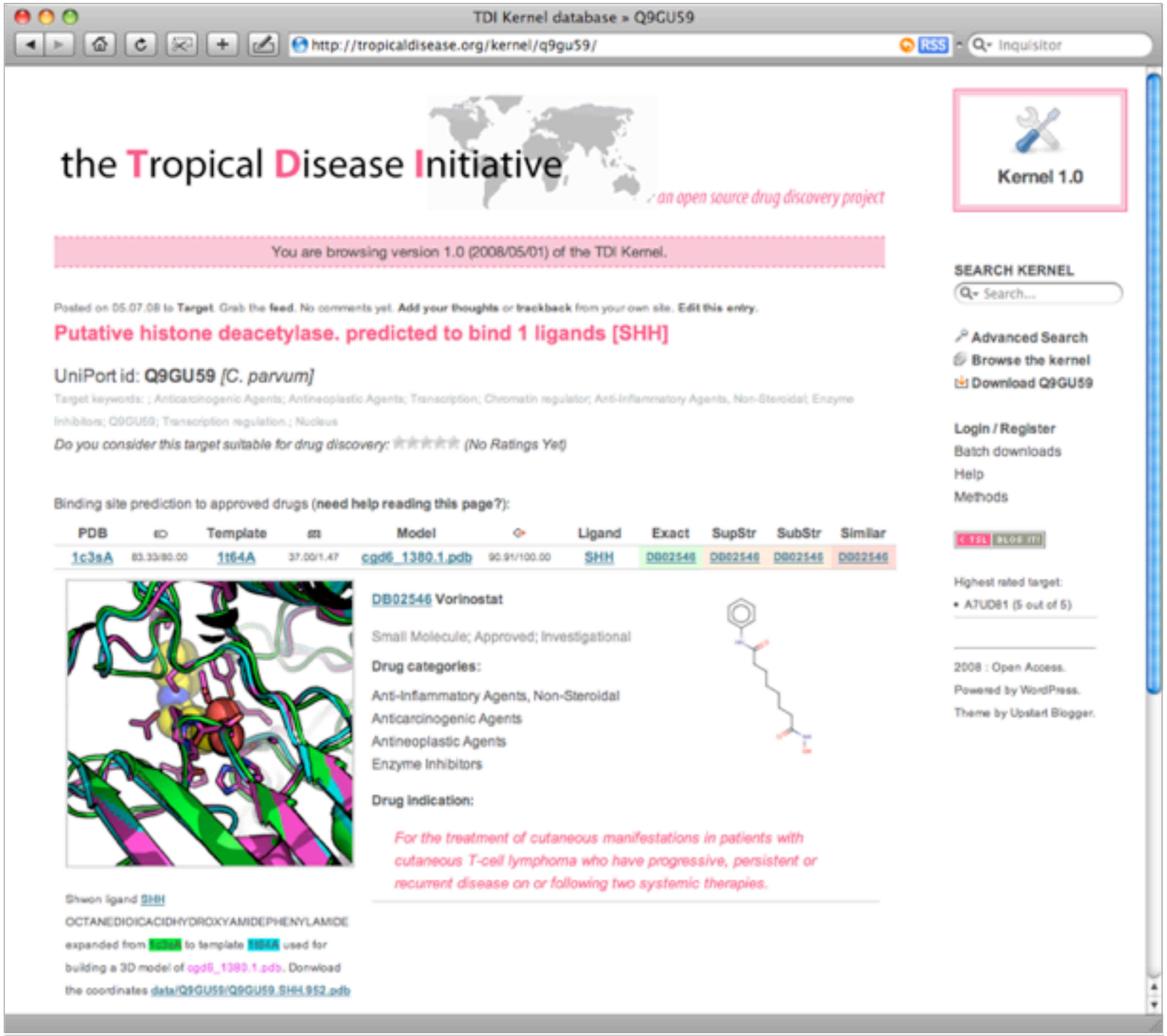




Leticia Ortí, Rodrigo J. Carbajo, and Antonio Pineda-Lucena

TDI’s kernel

http://tropicaldisease.org/kernel



TDI Kernel database » Q9GU59

http://tropicaldisease.org/kernel/q9gu59/

the Tropical Disease Initiative

an open source drug discovery project

Kernel 1.0

You are browsing version 1.0 (2008/05/01) of the TDI Kernel.

Posted on 05.07.08 to **Target**. Grab the **feed**. No comments yet. **Add your thoughts** or **trackback** from your own site. **Edit this entry**.

Putative histone deacetylase. predicted to bind 1 ligands [SHH]

UniPort id: **Q9GU59** *[C. parvum]*

Target keywords: ; Anticarcinogenic Agents; Antineoplastic Agents; Transcription; Chromatin regulator; Anti-Inflammatory Agents, Non-Steroidal; Enzyme Inhibitors; Q9GU59; Transcription regulation.; Nucleus

Do you consider this target suitable for drug discovery: ☆☆☆☆☆ *(No Ratings Yet)*

Binding site prediction to approved drugs (**need help reading this page?**):

PDB		Template		Model		Ligand	Exact	SupStr	SubStr	Similar
1c3sA	83.33/80.00	1t64A	37.00/1.47	cgd6_1380.1.pdb	90.91/100.00	SHH	DB02546	DB02546	DB02546	DB02546

DB02546 Vorinostat

Small Molecule; Approved; Investigational

Drug categories:

Anti-Inflammatory Agents, Non-Steroidal
Anticarcinogenic Agents
Antineoplastic Agents
Enzyme Inhibitors

Drug indication:

For the treatment of cutaneous manifestations in patients with cutaneous T-cell lymphoma who have progressive, persistent or recurrent disease on or following two systemic therapies.

Shwon ligand SHH OCTANEDIOICACIDHYDROXYAMIDEPHENYLAMIDE expanded from 1c3sA to template 1t64A used for building a 3D model of cgd6_1380.1.pdb. Donwload the coordinates data/Q9GU59/Q9GU59.SHH.952.pdb

SEARCH KERNEL

Advanced Search
Browse the kernel
Download Q9GU59

Login / Register
Batch downloads
Help
Methods

Highest rated target:

- A7UD81 (5 out of 5)

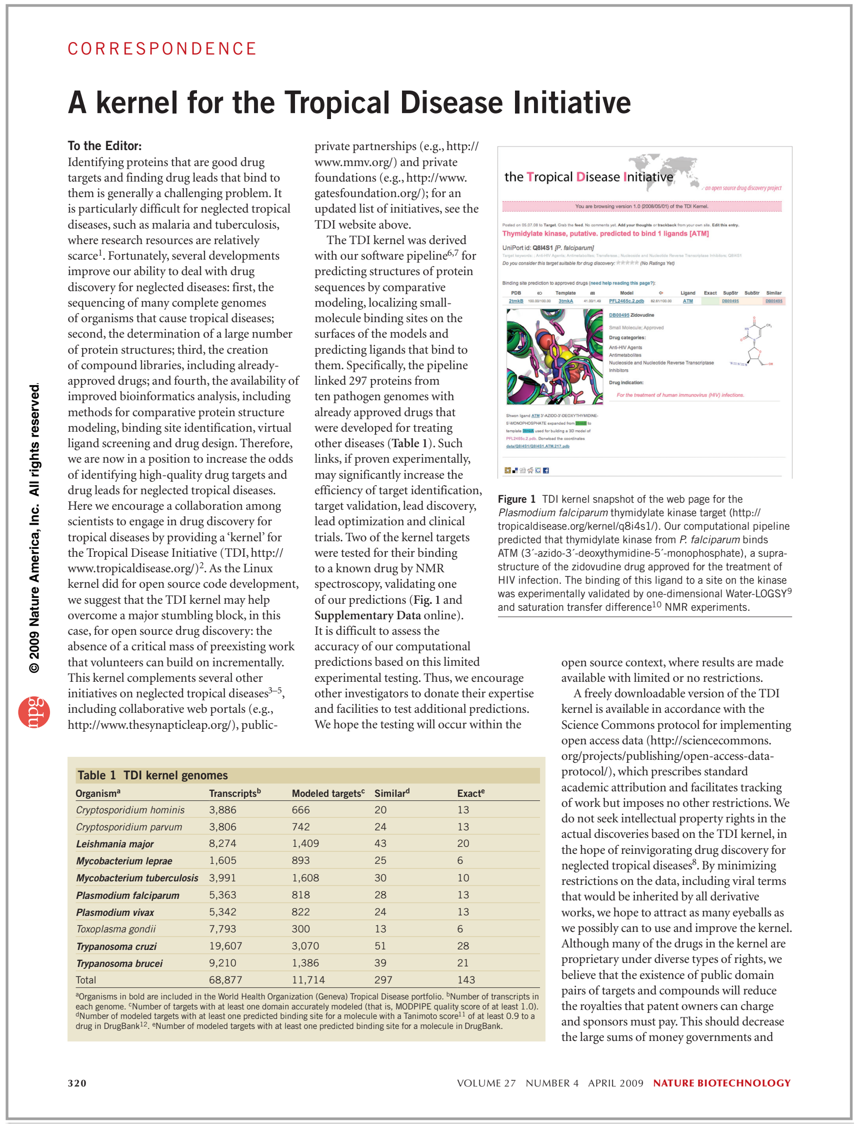
2008 : Open Access.
Powered by WordPress.
Theme by Upstart Blogger.

TDI's kernel

http://tropicaldisease.org/kernel

L. Orti *et al.*, *Nat Biotechnol* **27**, 320 (2009).

L. Orti *et al.*, *PLoS Negl Trop Dis* **3**, e418 (2009).



CORRESPONDENCE

A kernel for the Tropical Disease Initiative

To the Editor:

Identifying proteins that are good drug targets and finding drug leads that bind to them is generally a challenging problem. It is particularly difficult for neglected tropical diseases, such as malaria and tuberculosis, where research resources are relatively scarce[1]. Fortunately, several developments improve our ability to deal with drug discovery for neglected diseases: first, the sequencing of many complete genomes of organisms that cause tropical diseases; second, the determination of a large number of protein structures; third, the creation of compound libraries, including already-approved drugs; and fourth, the availability of improved bioinformatics analysis, including methods for comparative protein structure modeling, binding site identification, virtual ligand screening and drug design. Therefore, we are now in a position to increase the odds of identifying high-quality drug targets and drug leads for neglected tropical diseases. Here we encourage a collaboration among scientists to engage in drug discovery for tropical diseases by providing a 'kernel' for the Tropical Disease Initiative (TDI, http://www.tropicaldisease.org/)[2]. As the Linux kernel did for open source code development, we suggest that the TDI kernel may help overcome a major stumbling block, in this case, for open source drug discovery: the absence of a critical mass of preexisting work that volunteers can build on incrementally. This kernel complements several other initiatives on neglected tropical diseases[3–5], including collaborative web portals (e.g., http://www.thesynapticleap.org/), public-private partnerships (e.g., http://www.mmv.org/) and private foundations (e.g., http://www.gatesfoundation.org/); for an updated list of initiatives, see the TDI website above.

The TDI kernel was derived with our software pipeline[6,7] for predicting structures of protein sequences by comparative modeling, localizing small-molecule binding sites on the surfaces of the models and predicting ligands that bind to them. Specifically, the pipeline linked 297 proteins from ten pathogen genomes with already approved drugs that were developed for treating other diseases (**Table 1**). Such links, if proven experimentally, may significantly increase the efficiency of target identification, target validation, lead discovery, lead optimization and clinical trials. Two of the kernel targets were tested for their binding to a known drug by NMR spectroscopy, validating one of our predictions (**Fig. 1** and **Supplementary Data** online). It is difficult to assess the accuracy of our computational predictions based on this limited experimental testing. Thus, we encourage other investigators to donate their expertise and facilities to test additional predictions. We hope the testing will occur within the open source context, where results are made available with limited or no restrictions.

A freely downloadable version of the TDI kernel is available in accordance with the Science Commons protocol for implementing open access data (http://sciencecommons.org/projects/publishing/open-access-data-protocol/), which prescribes standard academic attribution and facilitates tracking of work but imposes no other restrictions. We do not seek intellectual property rights in the actual discoveries based on the TDI kernel, in the hope of reinvigorating drug discovery for neglected tropical diseases[8]. By minimizing restrictions on the data, including viral terms that would be inherited by all derivative works, we hope to attract as many eyeballs as we possibly can to use and improve the kernel. Although many of the drugs in the kernel are proprietary under diverse types of rights, we believe that the existence of public domain pairs of targets and compounds will reduce the royalties that patent owners can charge and sponsors must pay. This should decrease the large sums of money governments and

the Tropical Disease Initiative

an open source drug discovery project

You are browsing version 1.0 (2008/05/01) of the TDI Kernel.

Posted on 05.07.08 to Target. Grab the feed. No comments yet. Add your thoughts or trackback from your own site. Edit this entry.

Thymidylate kinase, putative. predicted to bind 1 ligands [ATM]

UniPort id: Q8I4S1 [P. falciparum]

Target keywords: ; Anti-HIV Agents; Antimetabolites; Transferase.; Nucleoside and Nucleotide Reverse Transcriptase Inhibitors; Q8I4S1

Do you consider this target suitable for drug discovery: ☆☆☆☆☆ (No Ratings Yet)

Binding site prediction to approved drugs (need help reading this page?):

PDB		Template		Model		Ligand	Exact	SupStr	SubStr	Similar
2tmkB	100.00/100.00	3tmkA	41.00/1.49	PFL2465c.2.pdb	82.61/100.00	ATM		DB00495		DB00495

DB00495 Zidovudine

Small Molecule; Approved

Drug categories:

Anti-HIV Agents
Antimetabolites
Nucleoside and Nucleotide Reverse Transcriptase Inhibitors

Drug indication:

For the treatment of human immunovirus (HIV) infections.

Shwon ligand ATM 3'-AZIDO-3'-DEOXYTHYMIDINE-5'-MONOPHOSPHATE expanded from 2tmkB to template 3tmkA used for building a 3D model of PFL2465c.2.pdb. Donwload the coordinates data/Q8I4S1/Q8I4S1.ATM.217.pdb

Figure 1 TDI kernel snapshot of the web page for the *Plasmodium falciparum* thymidylate kinase target (http://tropicaldisease.org/kernel/q8i4s1/). Our computational pipeline predicted that thymidylate kinase from *P. falciparum* binds ATM (3′-azido-3′-deoxythymidine-5′-monophosphate), a suprastructure of the zidovudine drug approved for the treatment of HIV infection. The binding of this ligand to a site on the kinase was experimentally validated by one-dimensional Water-LOGSY[9] and saturation transfer difference[10] NMR experiments.

Table 1 TDI kernel genomes

Organism[a]	Transcripts[b]	Modeled targets[c]	Similar[d]	Exact[e]
Cryptosporidium hominis	3,886	666	20	13
Cryptosporidium parvum	3,806	742	24	13
Leishmania major	8,274	1,409	43	20
Mycobacterium leprae	1,605	893	25	6
Mycobacterium tuberculosis	3,991	1,608	30	10
Plasmodium falciparum	5,363	818	28	13
Plasmodium vivax	5,342	822	24	13
Toxoplasma gondii	7,793	300	13	6
Trypanosoma cruzi	19,607	3,070	51	28
Trypanosoma brucei	9,210	1,386	39	21
Total	68,877	11,714	297	143

[a]Organisms in bold are included in the World Health Organization (Geneva) Tropical Disease portfolio. [b]Number of transcripts in each genome. [c]Number of targets with at least one domain accurately modeled (that is, MODPIPE quality score of at least 1.0). [d]Number of modeled targets with at least one predicted binding site for a molecule with a Tanimoto score[11] of at least 0.9 to a drug in DrugBank[12]. [e]Number of modeled targets with at least one predicted binding site for a molecule in DrugBank.

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

A Kernel for Open Source Drug Discovery in Tropical Diseases

Leticia Ortí[1,2], Rodrigo J. Carbajo[2], Ursula Pieper[3], Narayanan Eswar[3¤], Stephen M. Maurer[4], Arti K. Rai[5], Ginger Taylor[6], Matthew H. Todd[7], Antonio Pineda-Lucena[2], Andrej Sali[3]*, Marc A. Marti-Renom[1]*

1 Structural Genomics Unit, Bioinformatics and Genomics Department, Centro de Investigación Príncipe Felipe, Valencia, Spain, **2** Structural Biology Laboratory, Medicinal Chemistry Department, Centro de Investigación Príncipe Felipe, Valencia, Spain, **3** Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, and California Institute for Quantitative Biosciences, University of California San Francisco, San Francisco, California, United States of America, **4** Gould School of Law, University of Southern California, Los Angeles, California, United States of America, **5** School of Law, Duke University, Durham, North Carolina, United States of America, **6** The Synaptic Leap, San Ramon, California, United States of America, **7** School of Chemistry, University of Sydney, Sydney, New South Wales, Australia

Abstract

Background: Conventional patent-based drug development incentives work badly for the developing world, where commercial markets are usually small to non-existent. For this reason, the past decade has seen extensive experimentation with alternative R&D institutions ranging from private–public partnerships to development prizes. Despite extensive discussion, however, one of the most promising avenues—open source drug discovery—has remained elusive. We argue that the stumbling block has been the absence of a critical mass of preexisting work that volunteers can improve through a series of granular contributions. Historically, open source software collaborations have almost never succeeded without such "kernels".

Methodology/Principal Findings: Here, we use a computational pipeline for: (i) comparative structure modeling of target proteins, (ii) predicting the localization of ligand binding sites on their surfaces, and (iii) assessing the similarity of the predicted ligands to known drugs. Our kernel currently contains 143 and 297 protein targets from ten pathogen genomes that are predicted to bind a known drug or a molecule similar to a known drug, respectively. The kernel provides a source of potential drug targets and drug candidates around which an online open source community can nucleate. Using NMR spectroscopy, we have experimentally tested our predictions for two of these targets, confirming one and invalidating the other.

Conclusions/Significance: The TDI kernel, which is being offered under the Creative Commons attribution share-alike license for free and unrestricted use, can be accessed on the World Wide Web at http://www.tropicaldisease.org. We hope that the kernel will facilitate collaborative efforts towards the discovery of new drugs against parasites that cause tropical diseases.

Citation: Ortí L, Carbajo RJ, Pieper U, Eswar N, Maurer SM, et al. (2009) A Kernel for Open Source Drug Discovery in Tropical Diseases. PLoS Negl Trop Dis 3(4): e418. doi:10.1371/journal.pntd.0000418

Editor: Timothy G. Geary, McGill University, Canada

Received December 29, 2008; **Accepted** March 23, 2009; **Published** April 21, 2009

Funding: MAM-R acknowledges the support from a Spanish Ministerio de Educación y Ciencia grant (BIO2007/66670). AS acknowledges the support from the Sandler Family Supporting Foundation and the National Institutes of Health (R01 GM54762, U54 GM074945, P01 AI035707, and P01 GM71790). AP-L acknowledges the support from a Spanish Ministerio de Ciencia e Innovación grant (SAF2008-01845). RJC acknowledges the support from the Ramon y Cajal Program of the Spanish Ministerio de Educación y Ciencia. We are also grateful for computer hardware gifts to AS from Ron Conway, Mike Homer, Intel, IBM, Hewlett-Packard, and NetApp. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sali@salilab.org (AS); mmarti@cipf.es (MAM-R)

¤ Current address: DuPont Knowledge Center, Hyderabad, India

Introduction

There is a lack of high-quality protein drug targets and drug leads for neglected diseases [1,2]. Fortunately, many genomes of organisms that cause tropical diseases have already been sequenced and published. Therefore, we are now in a position to leverage this information by identifying potential protein targets for drug discovery. Atomic-resolution structures can facilitate this task. In the absence of an experimentally determined structure, comparative modeling can provide useful models for sequences that are detectably related to known protein structures [3,4]. Approximately half of known protein sequences contain domains that can be currently predicted by comparative modeling [5,6]. This coverage will increase as the number of experimentally determined structures grows and modeling software improves. A protein model can facilitate at least four important tasks in the early stages of drug discovery [7]: prioritizing protein targets for drug discovery [8], identifying binding sites for small molecules [9,10], suggesting drug leads [11,12], and optimizing these leads [13–15].

Here, we address the first three tasks by assembling our computer programs into a software pipeline that automatically and on large-scale predicts protein structures, their ligand binding sites, and known drugs that interact with them. As a proof of principle, we applied the pipeline to the genomes of ten organisms that cause tropical diseases ("target genomes"). We also experimentally tested two predicted drug-target interactions using Nuclear Magnetic

Acknowledgments

http://sgu.bioinfo.cipf.es
http://tropicaldisease.org - http://thesynapticleap.org

COMPARATIVE MODELING
Andrej Sali
Narayanan Eswar
Ursula Pieper
M. S. Madhusudhan
Min-Yi Shen
Ben Webb

Tropical Disease Initiative
Stephen Maurer
Arti Rai
Andrej Sali
Ginger Taylor
Matthew Todd

Bissan Al-Lazikani
James McKerrow
Brian Shoichet
David S. Roos

NMR
Antonio Pineda-Lucena
Leticia Ortí
Rodrigo J. Carbajo

FUNCTIONAL ANNOTATION
Stefania Bosi
Anna Tramontano

FUNDING
Ministerio de Educación y Ciencia
STREP UE Grant
Marie Curie Reintegration Grant

كاوست
KAUST