

CIPF / Nuevos Fármacos / Bioinformática

Descubrimiento de dianas terapéuticas para enfermedades tropicales



Marc A. Marti-Renom

<http://sgu.bioinfo.cipf.es>

Structural Genomics Unit
Bioinformatics Department

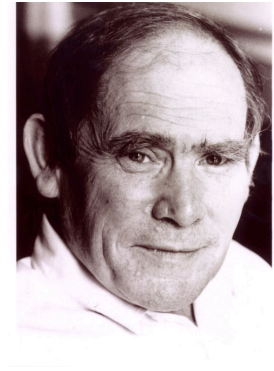
Centro de Investigación Príncipe Felipe (CIPF), Valencia, Spain



Data in the post-genomic era

Progress in science depends on new techniques, new discoveries and new ideas, **probably in that order.**

Sydney Brenner, 1980



The introduction and popularization of high-throughput techniques has drastically changed the way in which biological problems **can** be addressed and hypotheses can be tested.

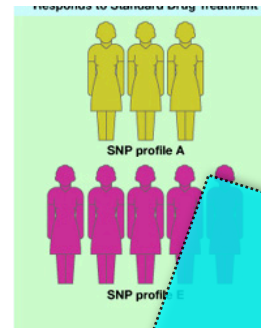
But not necessarily the way in which we really address or test them...

Genes in the DNA...



...code for proteins...

>protein kinase
acctgtgatggcgacaggactgtatgctg
atctatgctgatgcatgcatgctgactactgat

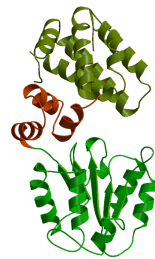


...produces the final phenotype

From genotype to phenotype.

...whose structure accounts

...plus the environment...



...which can be different because of the variability. 10 million SNPs



...whose final effect configures the phenotype...

Genes in the DNA...

Now: 22240 (NCBI build 35 12/04)
50-70% display alternative splicing
25%-60% unknown

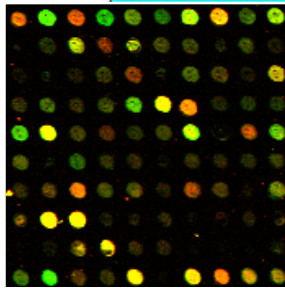
>protein kinase
acctgttgatggcgacagggactgtatgctgatctat
gctgatgcatgcatgctgactactgatggtgggcta
ttgactgatgtctatc....



Transfrags

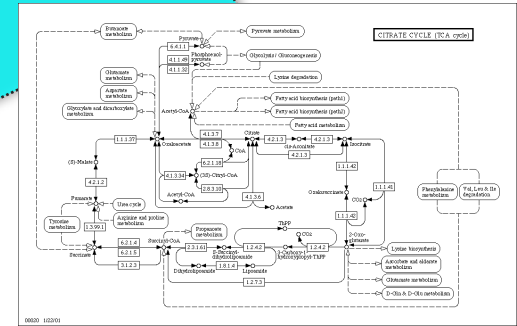
...when expressed in the proper moment and place...

A typical tissue is expressing among 5000 and 10000 genes



From genotype to phenotype.

(post-genomics scenario)



...code for proteins...

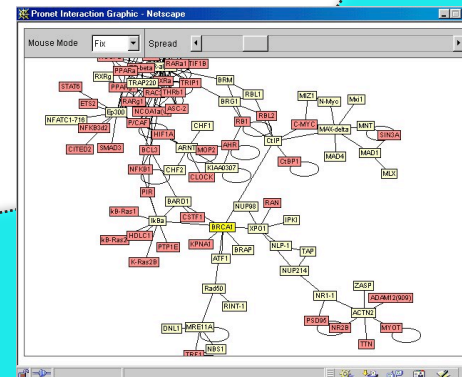
...conforming complex interaction networks...

That undergo post-translational modifications, somatic recombination...

100K-500K proteins



...whose structures account for function...

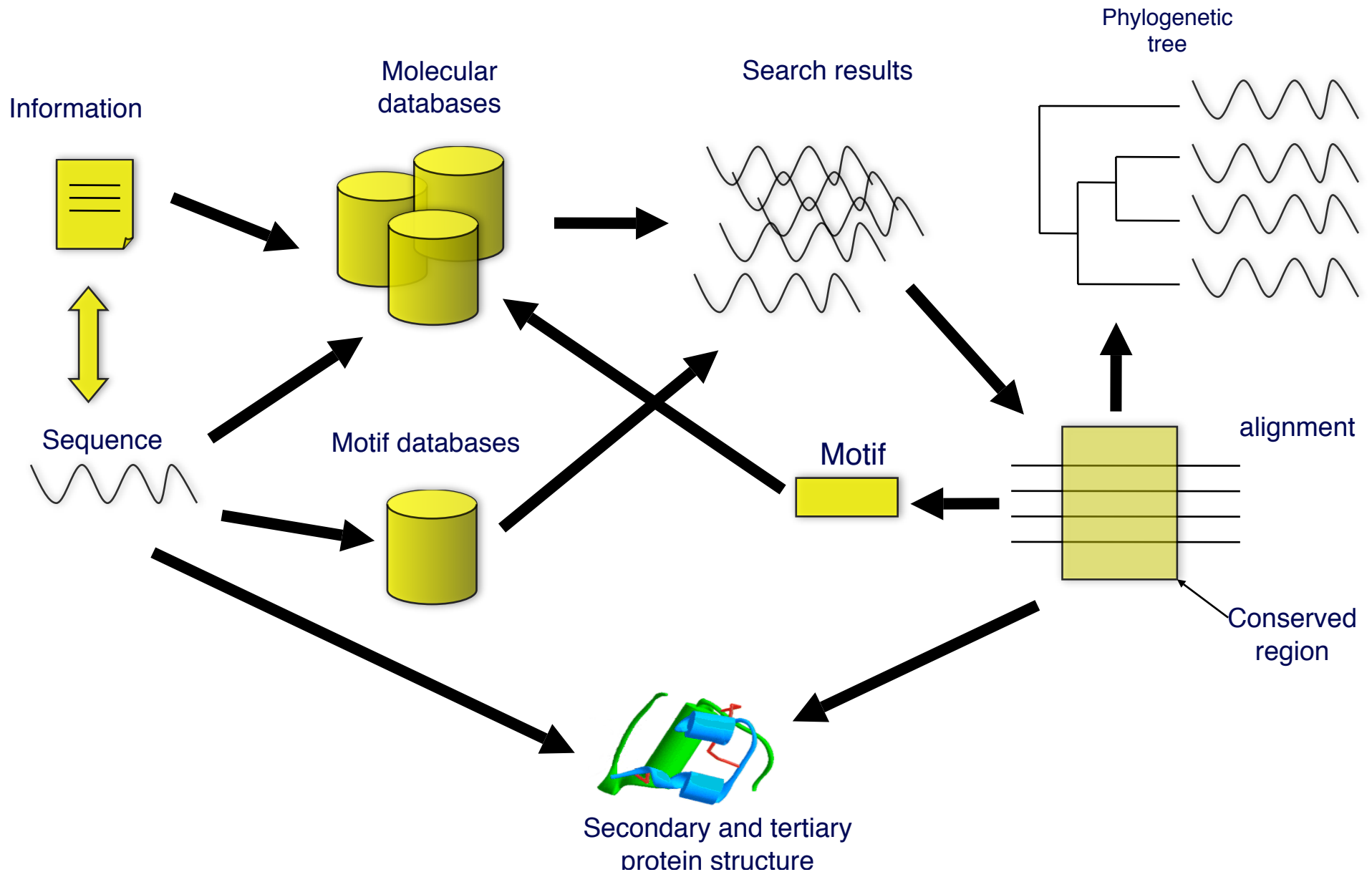


...in cooperation with other proteins...

Each protein has an average of 8 interactions

"The Aim"

Extracting as much information as possible from/for one single data



Bioinformatics Department <http://bioinfo.cipf.es>

Bioinformatics at CIPF

<http://bioinfo.cipf.es/> Google

Bioinformatics at CIPF

The Department

People
Functional Genomics Unit
Pharmacogenomics & Comparative Genomics Unit
Structural Genomics Unit

Tools

DNA array data analysis
SNPs data analysis
Functional profiling
Downloads

Documents & Publications

Papers
Communications
Supplementary material

Meetings & Courses

Meetings & workshops
Courses
On line courses
Accommodation

Coming events...

 **CAMDA 2007, Dec 13-14**
Seventh international conference for the Critical Assessment of Microarray Data Analysis

 **PRINCIPE FELIPE**
CENTRO DE INVESTIGACION

 **INB** INSTITUTO NACIONAL DE BIOINFORMATICA

 **CEGEN**
Centro Nacional de Genotipado

 **ciberer**
Centro de Investigación Biomédica En Red de Enfermedades Raras

 **GECOBIO**
VALENCIA

 Instituto Nacional de Bioinformática

 **INDIGO**
Integrated Highly Sensitive Fluorescence-based Biosensor for Diagnostic applications

Spotlight Tools

 **Access**
Tool usage around the World.

 **Prophet**
A tool for building a class predictor

 **PupaSuite**
Interactive selection of optimal sets of SNPs for large-scale genotyping.

 **MARMITE**
Functional profiling with PubMed words.

 **FatiGO+**
Functional interpretation of large-scale experiments using GO, KEGG, Interpro, Transfac, CisRed...

 **CAAT**
Draw, browse, analyze and validate interactively your hierarchical clustering results.

 **FatiScan**
Detect blocks of functionally related genes (GO, KEGG) with significant coordinate (although modest) over- or under-expression.

 **SIDE**
Interactive design of Small interfering RNA.

 **Blast2GO**
A universal Gene Ontology annotation, visualization and analysis suite for functional genomics research.

News

May 2007 - Positions available: one biostatistician and one bioinformatician

April 2007 - **Blast2GO**, a universal Gene Ontology annotation, visualization and analysis suite for functional genomics research.

January 2007 - **Phylemon**, a suite of tools for molecular evolution, phylogenetics and phylogenomics.

September 2006 - **Prophet**, a tool for building a class predictor.

March 2006 - Computing journal award to the best R&D project.

February 2006 - New releases: **GEPAS v3.0** and **Babelomics v2.0**. Also **PupaSuite**, interactive selection of optimal sets of SNPs for large-scale genotyping.

December 2005 - CIPF receives autelsi award: "The First European Cluster for Scientific Computing with Free software on GNU/Linux."

More news...

Google Search

Search WWW @ Search bioinfo.cipf.es

Postal address

Bioinformatics Department (CIPF)
Av. Autopista del Saler, 16
(Camino de las moreras)
46013 Valencia, Spain
Tel.: + 34 96 328 96 80
Fax: + 34 96 328 97 01
(see map)

Send comments to the webmaster. Last updated: January 15, 2007

Packages

 **GEPAS**
Gene Expression Pattern Analysis Suite (v3.0).

 **BABELOMICS**
Suite of tools for Functional Profiling (v2.0).

 **Phylemon**
Suite of tools for molecular evolution & phylogenetics (v1.0).

 **Data Analysis and Visualization in Genomics and Proteomics.**
Francisco Azuaje, Joaquín Dopazo (Editors)
ISBN: 0-470-09439-7

Bioinformatics Department <http://bioinfo.cipf.es>



Functional Genomics
Dr. Joaquín Dopazo

GEPAS



BABELOMICS

<http://gepas.bioinfo.cipf.es>

Gene Expression Pattern Analysis Suite

<http://pupasuite.bioinfo.cipf.es>

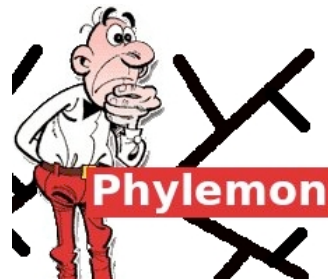
SNP Analysis Suite

<http://babelomics.bioinfo.cipf.es>

Functional Profiling Analysis Suite



Comparative Genomics
Pharmacogenomics
Dr. Hernán Dopazo



<http://phylemon.bioinfo.cipf.es>

Molecular Evolution Analysis Suite



Structural Genomics
Dr. Marc A. Marti-Renom

DBAli_{v2.0}
Home

<http://www.dbali.org>

Structural Biology Analysis Suite

Bioinformatics Department <http://bioinfo.cipf.es>



Tropical Disease Initiative (TDI)

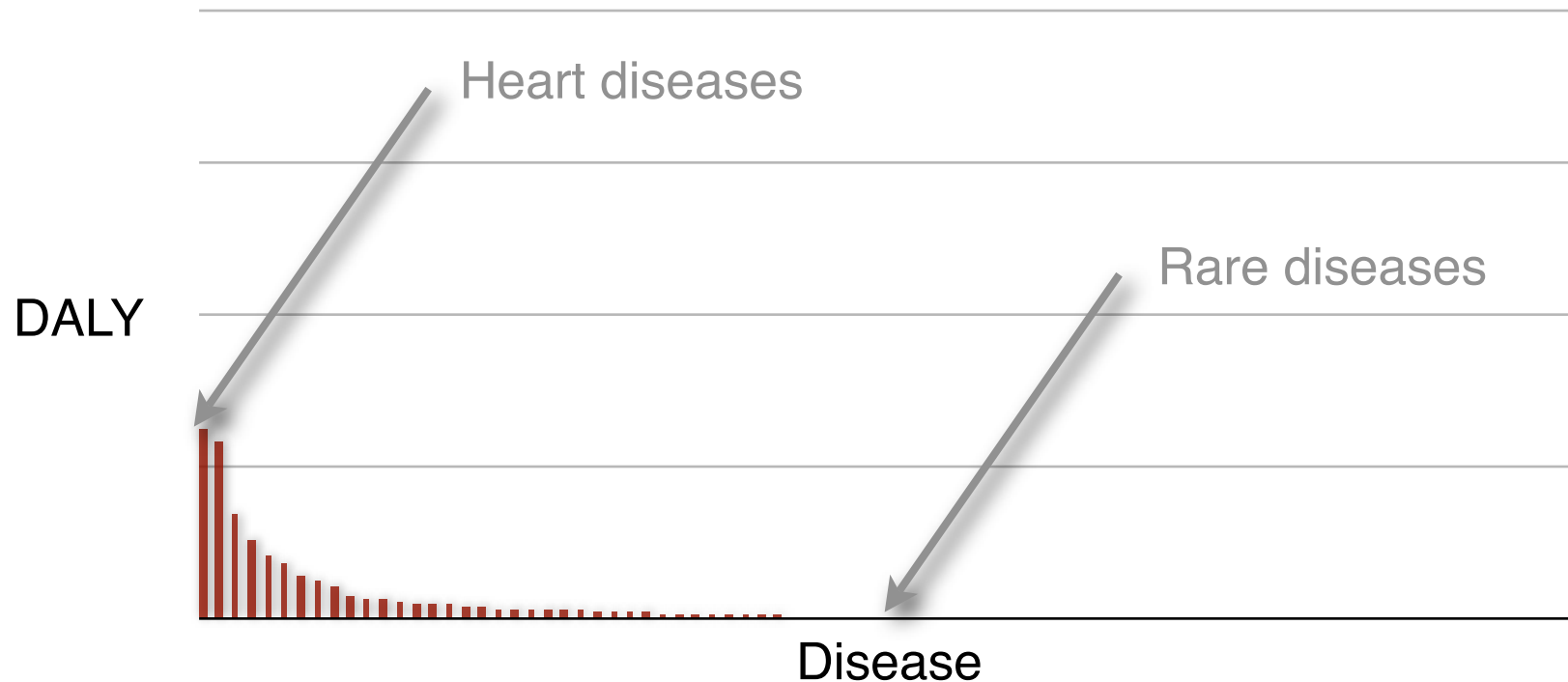
Predicting binding sites in protein structure models.



<http://www.tropicaldisease.org>

Need is High in the Tail

- DALY Burden Per Disease in Developed Countries
- DALY Burden Per Disease in Developing Countries



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

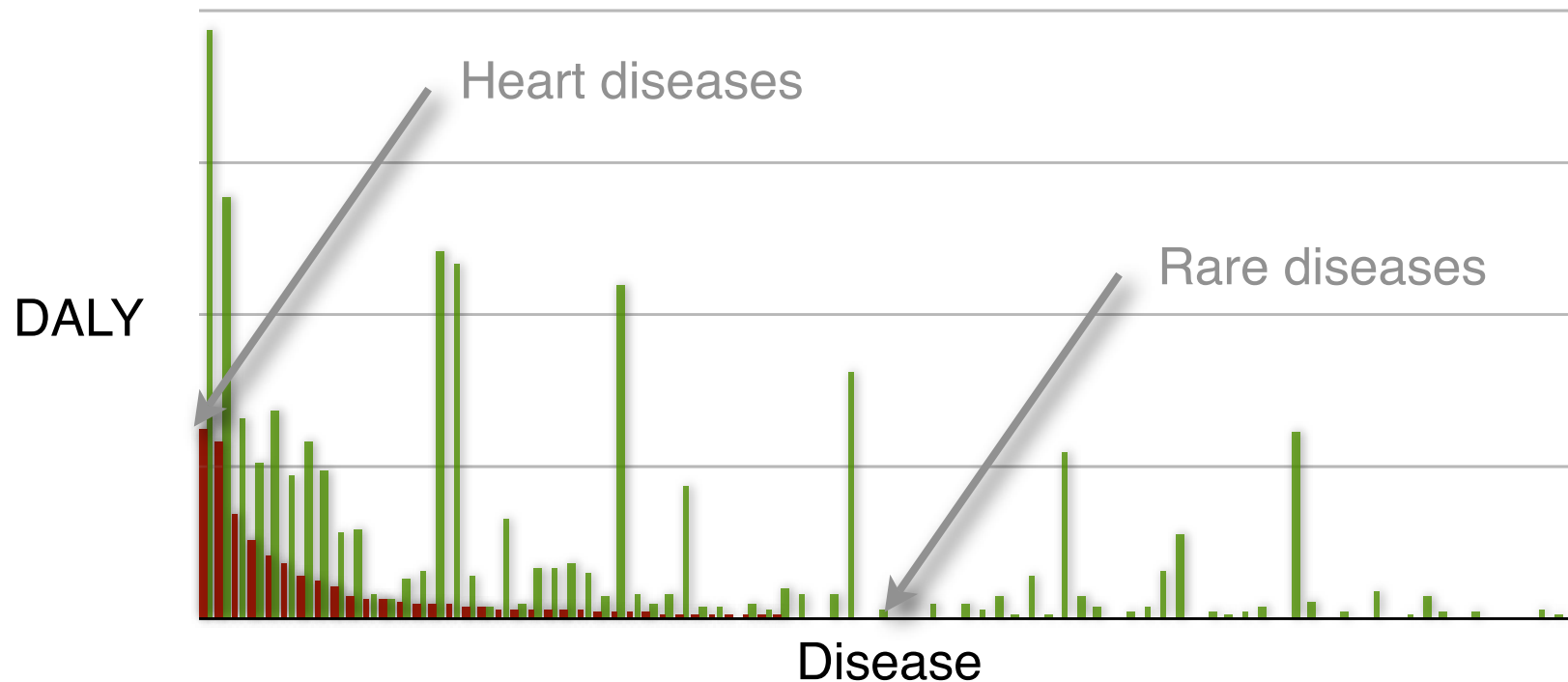
DALY - Disability adjusted life years

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

- DALY Burden Per Disease in Developed Countries
- DALY Burden Per Disease in Developing Countries



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

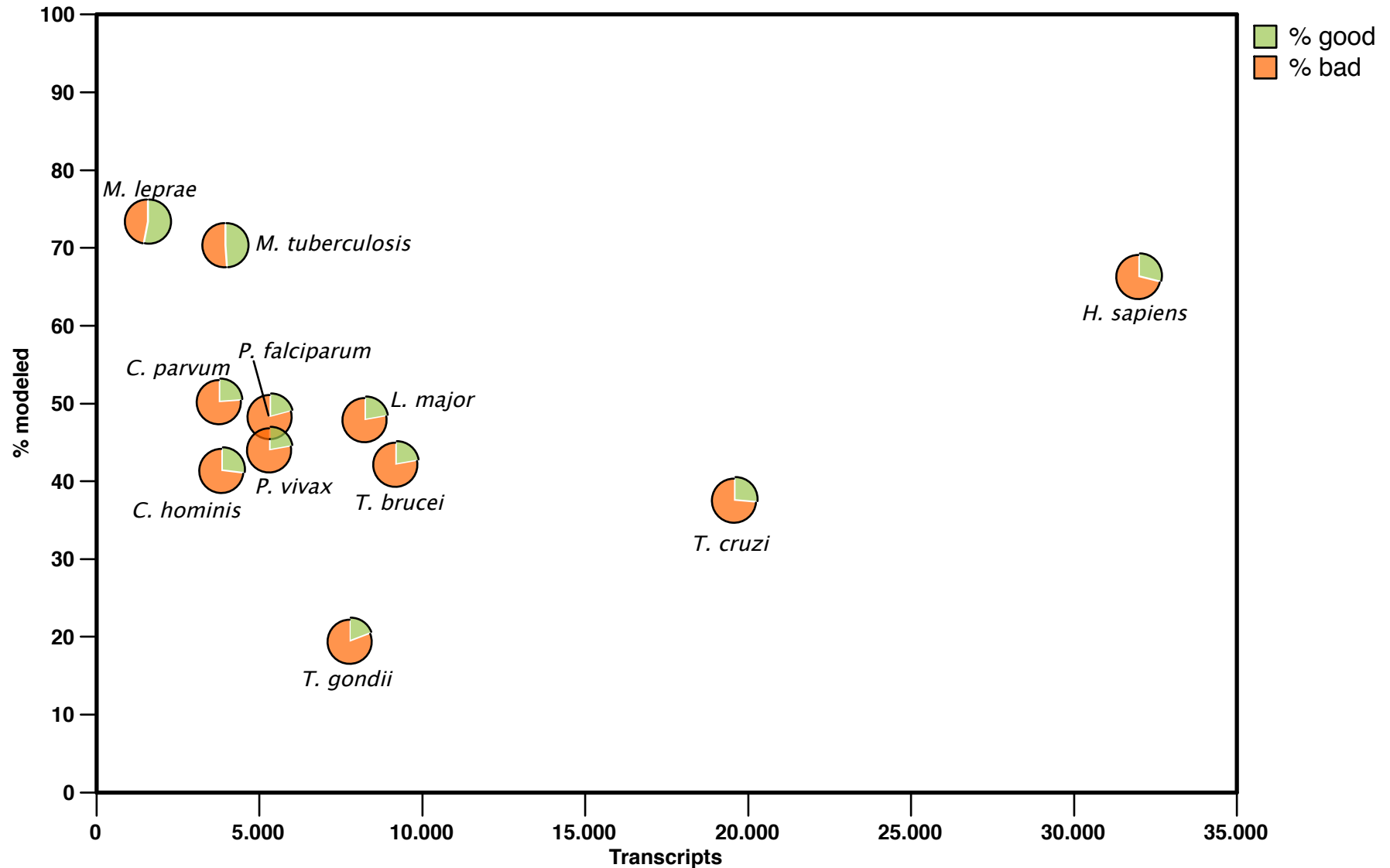
DALY - Disability adjusted life years

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.

Modeling Genomes

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

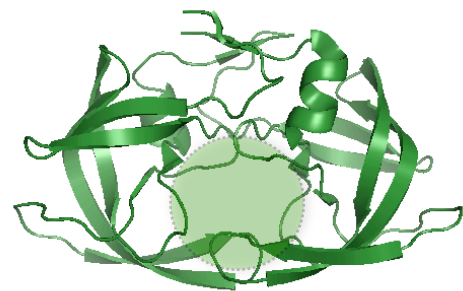
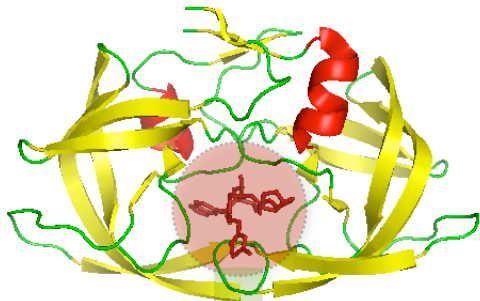


A good model has MPQS of 1.1 or higher

Comparative docking

1. Expansion

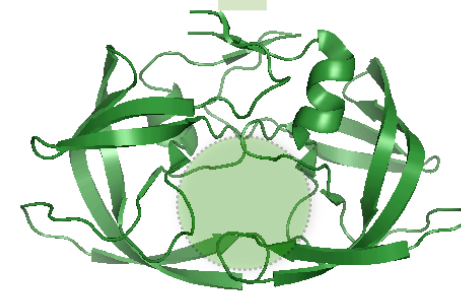
co-crystallized protein/ligand



crystallized protein

2. Inheritance

model



template

Summary table

models with inherited ligands

from 16,284 good models, 295 inherited a ligand/substance with at least one compound already approved by FDA and ready to be used from ZINC

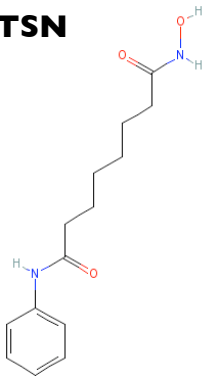
	Transcripts	Good	Ligands	Lipinski	Lipinski+ZINC	FDA+ZINC
<i>C. hominis</i>	3,886	886	183	131	28	12 (10)
<i>C. parvum</i>	3,806	949	219	145	30	12 (10)
<i>L. major</i>	8,274	1,845	488	334	84	44 (34)
<i>M. leprae</i>	1,605	1,321	286	189	39	29 (25)
<i>M. tuberculosis</i>	3,991	2,887	404	285	71	44 (37)
<i>P. falciparum</i>	5,363	1,057	271	191	48	20 (16)
<i>P. vivax</i>	5,342	1,042	267	177	37	18 (15)
<i>T. brucei</i>	921	1,795	440	309	94	46 (36)
<i>T. cruzi</i>	19,607	3,915	730	493	127	62 (52)
<i>T. gondii</i>	7,793	587	174	124	28	8 (7)
TOTAL	60,588	16,284	3,462	2,378	586	295 (242)

Example of inheritance (inheritance)

LmjF21.0680 from *L. major* "Histone deacetylase 2" (model 1)

	Formula	Name	Cov.	Seq. Id. (%)	Residues
TSN	C ₁₇ H ₂₂ N ₂ O ₃	Trichostatin A	100.00	90.9	90 131 132 140 141 167
SHH	C ₁₄ H ₂₀ N ₂ O ₃	Octadenioic acid hydroxyamide phenylamide	100.00	90.9	169 256 263 293 295

TSN



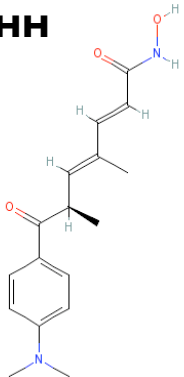
suberoylanilide hydroxamic acid

Pharmacological Action:

[Anti-Inflammatory Agents, Non-Steroidal](#)
[Antineoplastic Agents](#)
[Enzyme Inhibitors](#)
[Anticarcinogenic Agents](#)

Inhibits histone deacetylase 1 and 3

SHH



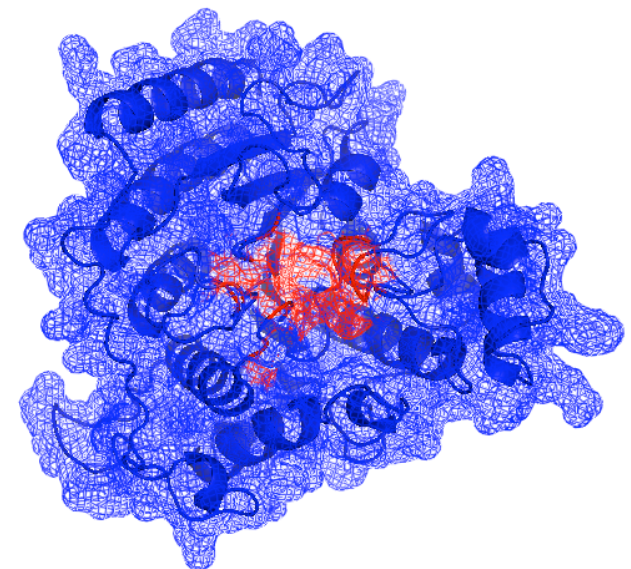
trichostatin A

Pharmacological Action:

[Antibiotics, Antifungal](#)
[Enzyme Inhibitors](#)
[Protein Synthesis Inhibitors](#)

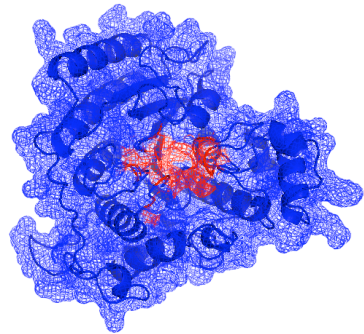
chelates zinc ion in the active site of histone deacetylases, resulting in preventing histone unpacking so DNA is less available for transcription

	LmjF21.0680.1.pdb
Template	1t64A
Seq. Id (%)	38.00
MPQS	1.47



Example of inheritance (CDD-Roos-literature)

LmjF21.0680 from L. major "Histone deacetylase 2" (model 1)



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
0066-4804/04/\$08.00+0 DOI: 10.1128/AAC.48.4.1435–1436.2004
Copyright © 2004, American Society for Microbiology. All Rights Reserved.

Vol. 48, No. 4

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

Acknowledgments

Structural Genomics Unit (CIPF)

Marc A. Marti-Renom

Emidio Capriotti

Peio Ziarsolo Areitioaurtena

Comparative Genomics Unit (CIPF)

Hernán Dopazo

Leo Arbiza

Francisco García

Functional Genomics Unit (CIPF)

Joaquín Dopazo

Fátima Al-Shahrour

José Carbonell

Ignacio Medina

David Montaner

Joaquín Tárraga

Ana Conesa

Toni Gabaldón

Eva Alloza

Lucía Conde

Stefan Goetz

Jaime Huerta Cepas

Marina Marcet

Pablo Minguez

Jordi Burguet Castell

FUNDING

Prince Felipe Research Center

Marie Curie Reintegration Grant

STREP EU Grant

Generalitat Valenciana

Tropical Disease Initiative

Stephen Maurer (UC Berkeley)

Arti Rai (Duke U)

Andrej Sali (UCSF)

Ginger Taylor (TSL)

Barri Bunin (CDD)

STRUCTURAL GENOMICS

Stephen Burley (SGX)

John Kuriyan (UCB)

NY-SGXRC

MAMMOTH

Angel R. Ortiz

BIOLOGY

Jeff Friedman (RU)

James Hudsped (RU)

Partho Ghosh (UCSD)

Alvaro Monteiro (Cornell U)

Stephen Krilis (St. George H)

FUNCTIONAL ANNOTATION

Fatima Al-Shahrour

Joaquín Dopazo

COMPARATIVE MODELING

Andrej Sali

M. S. Madhusudhan

Narayanan Eswar

Min-Yi Shen

Ursula Pieper

Bino John

Maya Topf

FUNCTIONAL ANNOTATION

Andrea Rossi

Fred Davis



<http://bioinfo.cipf.es>
<http://sgu.bioinfo.cipf.es>