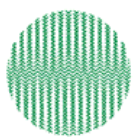
in silico strategies for target discovery

Eric Maréchal

Comparative chemogenomics team,
Institut de Recherches en Sciences et Technologies pour le Vivant
CEA Grenoble, France

UNIVERSITE JOSEPH FOURIER
SCIENCES. TECHNOLOGIE. MEDECINE. GRENOBLE - ALPES



INRA
Institut National de la Recherche Agronomique

Cerep

GENE·IT

A **target** is something you aim at.

In the context of a pathology, a **target** can be

- a ***biological entity*** ← ***Ontology***
- and/or a ***biological phenomenon*** ← ***Functional representations***

Intervention on the target depends on

- the **level of knowledge of the disease**,
- the **cultural background of the physician**
  - **traditional medicine** will focus more on the symptoms in order to **end the symptoms.**
  - **allopathic** (western, orthodox) **medicine** will focus on the causes (*i.e.* the **etiology**) in order to **oppose the causes**.
  - **homeopathic medicine** (in the sense of non-allopathic medicine) will focus on the etiology in order to **treat likes with likes**. In the widest sense of the word, vaccination is a category of homeopathic treatment.
- the **technics and technologies for the intervention**
  - **without exogenous agents**
    - musculoskeletal manipulations, surgery, diets
  - **with exogenous biochemicals**
    - **natural substances** based on **traditional pharmacopoeia**
    - **drugs** derived of traditional pharmacopoeia or from synthetic chemistry
    - **vaccines**
  - **with exogenous genes (gene therapy)**

Malaria ?

***In silico* strategies for target discovery, should** :

1) comply with knowledge of the disease in ETIOLOGIC terms

- **objects**, attributes and **data** that are **consistent with**
  - **biological entities** (as defined or understood by bench biologists in reductionist terms); i.e. genes, RNA, proteins, solutes, cellular structures, tissues, organs, etc.
    Ex.: a protein "E"
  - and their **measures**, scales, i.e. concentration
- **processes** and functions that represent biological operations
  -alterations over time (growth, transformations, movements) :
    Ex.: a reaction "A → B"
- **relations** between objects and processes
    Ex.: "C" controls "E" that catalyzes "A → B"
      If "B" is related to the pathology, "C" might be a target

2) benefit from biological knowledge in widest terms

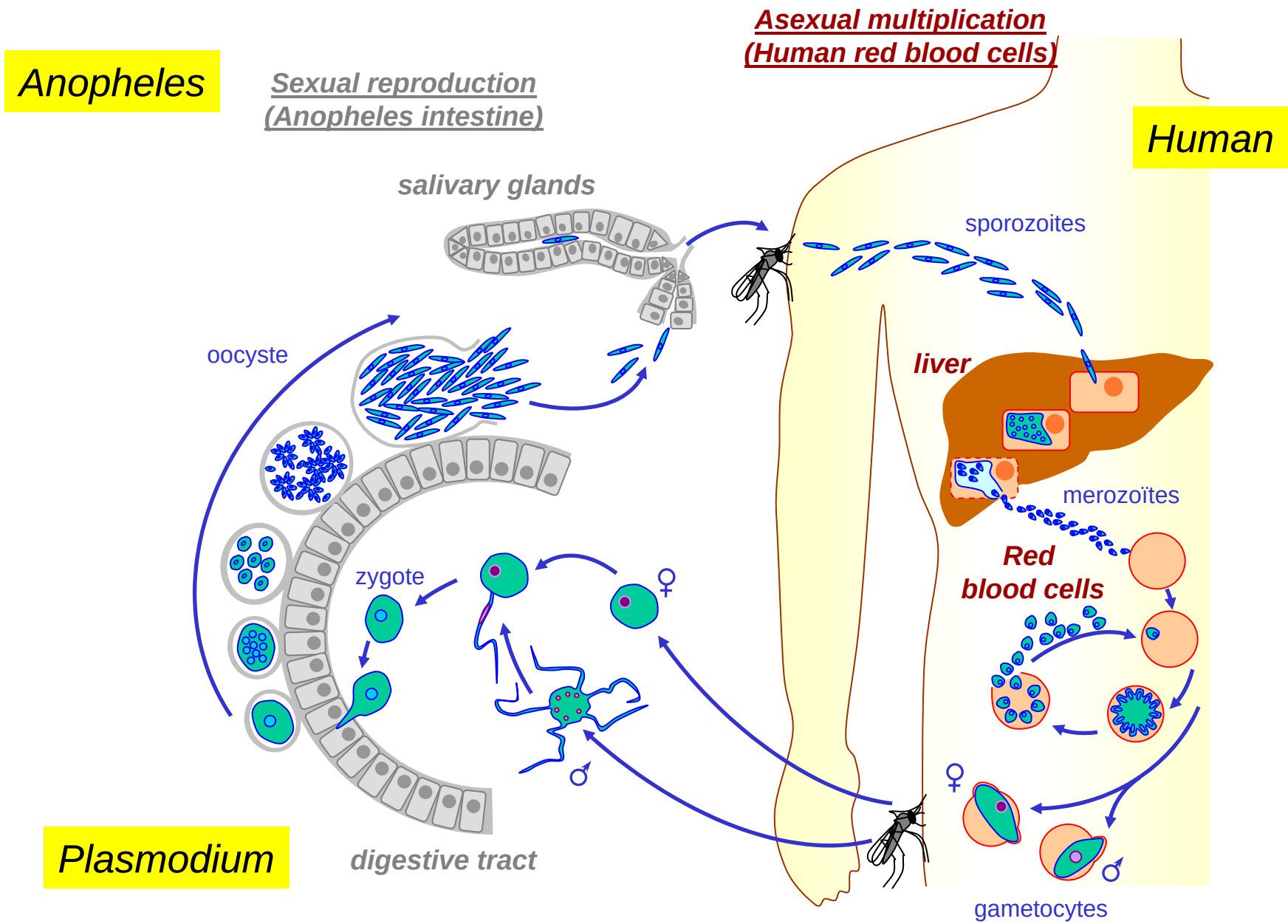
- **access** to all data, experiments, literature
- **mining**, **comparisons**, **knowledge transfers**

3) include, enable, enhance, maximize creative approaches

4) comply with feasibility of a treatment

- **genetic variabilitiy of target? drug? vaccine? gene therapy?**

Malaria


Asexual multiplication
(Human red blood cells)
Anopheles
Sexual reproduction
(Anopheles intestine)
Human
salivary glands
sporozoites
oocyste
liver
merozoïtes
Red
blood cells
zygote
♀
♂
Plasmodium
digestive tract
gametocytes

Malaria

FlyBase

Anopheles

Fly

Integr8
UniProt

Mouse

Human

Rat

ensembl!

Ensembl
VectorBase

Bacillus

coliBASE

E. coli

Chicken

CyanoBase

Synechocystis

TAIR

Arabidopsis

Leishmania

PlasmoDB
ApiDB

Euglena

Trypanosoma

Toxoplasma

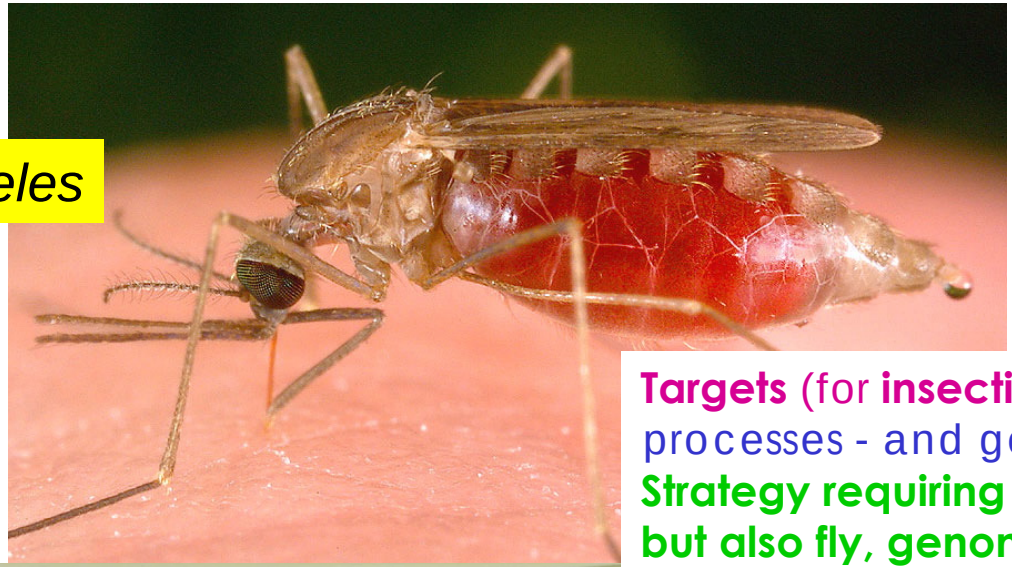
Cryptosporidium

Babesia

Plasmodium

Desp. waiting for Google Bio-earth…

Malaria



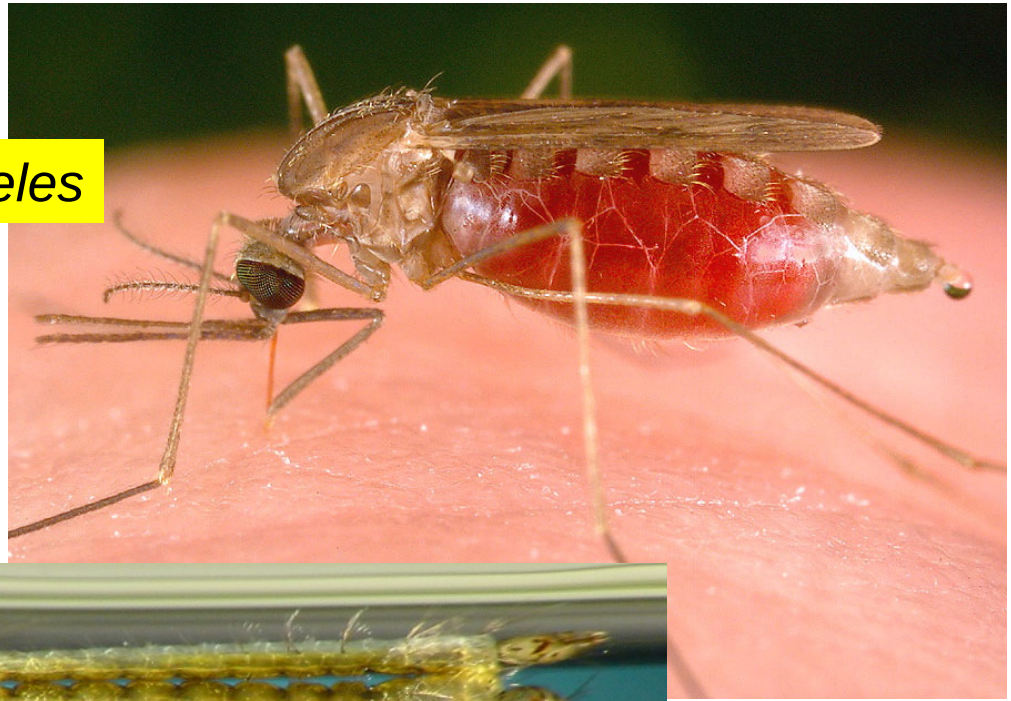
Anopheles

Human

Targets (for **insecticides**): mostly vital processes - and genes –
Strategy requiring an exploration of mosquito, but also fly, genomic/postgenomic molecular, and functional data

Plasmodium

Malaria



Anopheles

Human

Targets (for **vector resistance and population control**): Anopheles **genes introducing resistance to *Plasmodium*** (with possible genetic manipulations to introduce this resistance) **+ fitness genes** => promote the introduction of competing populations that resist to *Plasmodium*.

Strategy requiring an exploration of mosquito genomic/postgenomic molecular, and functional data + *Plasmodium* genomic/postgenomic data focusing on parasitic stages within mosquitoes.

Plasmodium

Malaria

Anopheles

Targets ?

Human

Plasmodium

Malaria

Anopheles

Targets (for **resistance**):
Human **genes introducing resistance to *Plasmodium*** (with possible genetic manipulations to introduce this resistance) => promote the human processes involved in resistance (with immune response enhancer, but also including gene therapy approaches on a long, long, long term…)
Strategy requiring an exploration of human genomic/postgenomic molecular, and functional data + mouse/rat models requiring comparative tools (homologies, syntenies) + epidemiological information.
Chromosome 5q31-q33

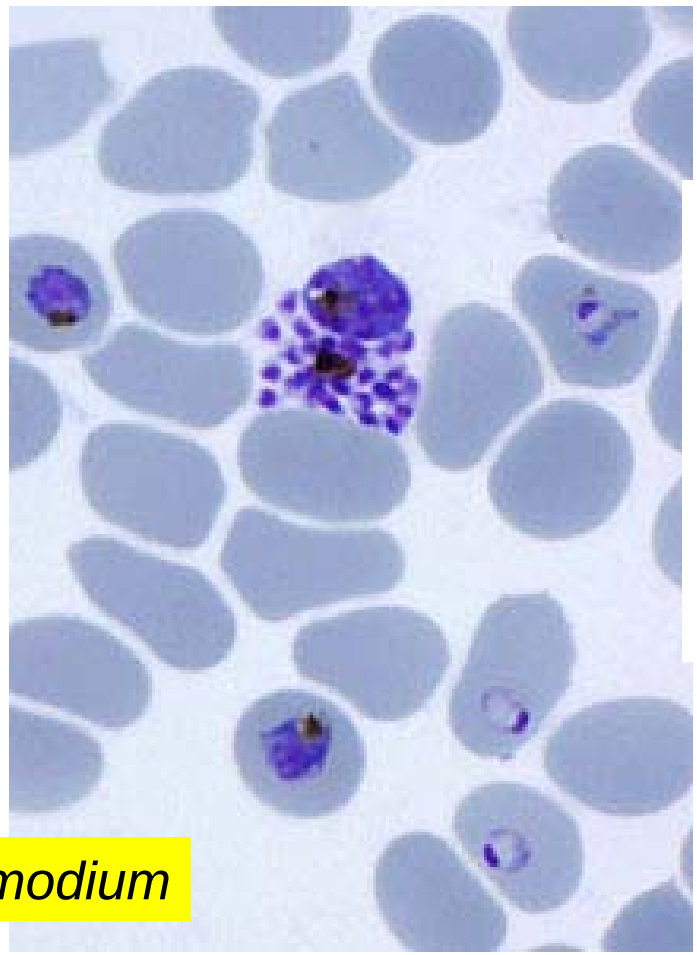
Plasmodium

Human

Malaria

Anopheles

Human



Plasmodium

Photo Jason Socrates Bardi

Targets (for **drugs/vaccines**): mostly vital processes - and genes -
Strategy requiring an exploration of *Plasmodium (P. falciparum* and other malaria species), genomic/postgenomic molecular, and functional data + a study of genetic variations of the target candidate + a sufficient difference with human features to avoid toxicity.

Malaria

Anopheles

Targets (for **insecticides**): mostly vital processes - and genes –
Mosquito + fly, genomic/postgenomic molecular, and functional data

Human

Targets (for **vector resistance and population control**): Anopheles **genes introducing resistance to *Plasmodium* + fitness genes**
Mosquito genomic/postgenomic molecular, and functional data + *Plasmodium* genomic/postgenomic data focusing on parasitic stages within mosquitoes + population genetic data

Targets (for **resistance**):
Human **genes introducing resistance to *Plasmodium***
Human genomic/postgenomic molecular, and functional data + mouse/rat models requiring comparative tools (homolgogies, syntenies) + epidemiological information.

+ Experiments for Target validation (can be knock out but not only)
+ Experiment to develop treatment (drug discovery + development; vaccine development)

Plasmodium

Targets (for **drugs/vaccines**): mostly vital processes - and genes -
***Plasmodium (P. falciparum* and other malaria species), genomic/postgenomic molecular, and functional data +human genomic/postgenomic data + a study of genetic variations of the target candidate**

Malaria

Anopheles

Targets (for **insecticides**): mostly vital processes - and genes –
Mosquito + fly, genomic/postgenomic molecular, and functional data

Targets (for **vector resistance an**
population control): Anopheles **g**
introducing resistance to *Plasmo*
fitness genes
Mosquito genomic/postgenomic
molecular, and functional data +
***Plasmodium* genomic/postgeno**
focusing on parasitic stages with
mosquitoes + population genetic

+ Experiments for Target valid
knock out but not only)
+ Experiment to develop treat
discovery + development; va
development)

Invaluable collaborative work to generate/collect/control data and put some order in this avalanche of information. Endorsed by all individual scientists + dedicated teams having sufficient critical masses. No cost evaluation (gigantic); access considered as a « universal human right »; only limit is the investment in the maintenance, improvement and development of wise data storage models and systems, navigation maps and mining tools (DATABASES, KNOWLEDGE BASES, KNOWLEDGE SPACES) and accesses (internet portals and grid distribution)

Human

e):
ucing resistance

stgenomic
tional data +
equiring
omolgogies,
iological

Targets
genes -
esses - and

Plasmodium

***Plasmodium (P. falciparum* and other malaria species), genomic/postgenomic molecular, and functional data +human genomic/postgenomic data + a study of genetic variations of the target candidate**

Malaria

Anopheles

Targets (for **insecticides**): mostly vital processes - and genes –
Mosquito + fly, genomic/postgenomic molecular, and functional data

Human

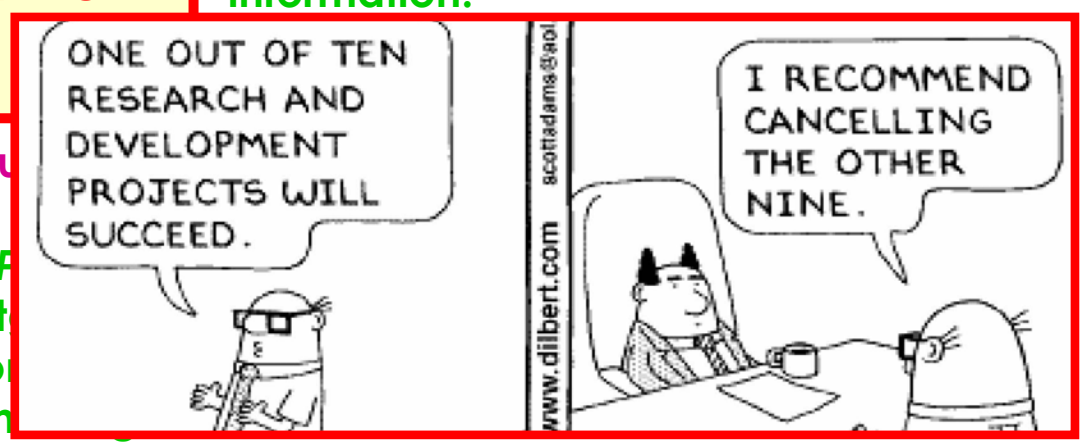
Targets (for **vector resistance and population control**): Anopheles **genes introducing resistance to *Plasmodium* + fitness genes**
Mosquito genomic/postgenomic molecular, and functional data + *Plasmodium* genomic/postgenomic data focusing on parasitic stages within mosquitoes + population genetic data

Linear workflows of experiments that can be stopped as soon as failure indicators appear. Cost evaluation includes the cost of failures (>70 % of drug discovery/development projects fail), *i.e.* 800 M€.
Endorsed by small teams.
Time scale is ~10 yr per project.
This is clearly the LIMITING STEP.
In silico information are non limiting by comparison.

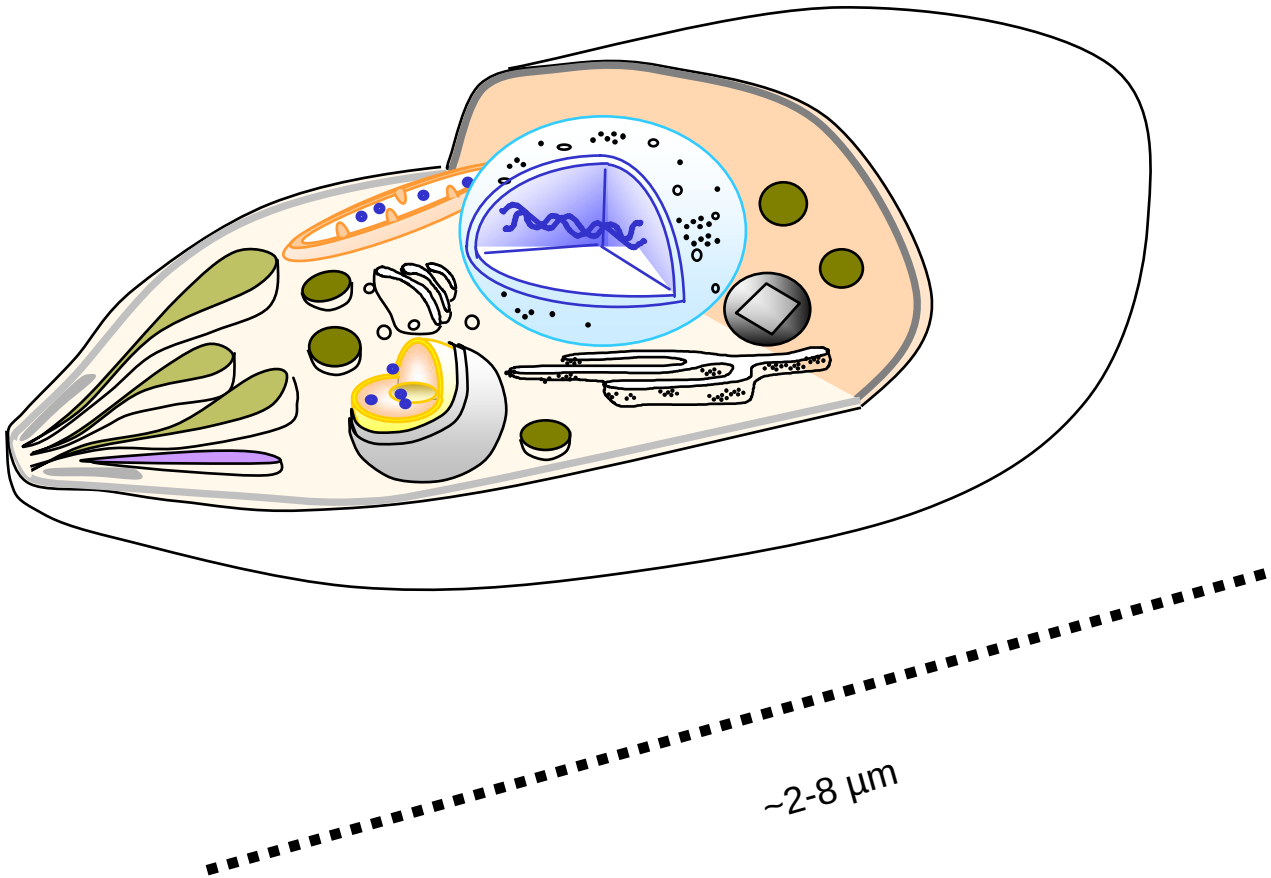
+ Experiments for Target validation (can be knock out but not only)
+ Experiment to develop treatment (drug discovery + development; vaccine development)

Plasmodium

Targets (for **dru**
genes -
***Plasmodium* (P**
genomic/post
+human genor
variations of th



Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.



Plasmodium falciparum an apicomplexan

Malaria targets for new drugs ?
An example of a drug discovery project,
combining bioinformatics and experiments.

1991

J. Protozool. 1991 38:243-5.

The putative mitochondrial genome of *Plasmodium falciparum.*

Feagin JE, Gardner MJ, Williamson DH, Wilson RJ.

Intraerythrocytic stages of mammalian malarial parasites employ glycolysis for energy production but some aspects of mitochondrial function appear crucial to their survival since inhibitors of mitochondrial protein synthesis and electron transport have antimalarial effects. Investigations of the putative mitochondrial genome of Plasmodium falciparum have detected organellar rRNAs and tRNAs encoded by a 35 kb circular DNA. Some features of the organization and sequence of the rRNA genes are reminiscent of chloroplast DNAs. The 35 kb DNA also encodes open reading frames for proteins normally found in chloroplast but not mitochondrial genomes. An apparently unrelated 6 kb tandemly repeated element which encodes two mitochondrial protein coding genes and fragments of rRNA genes is also found in malarial parasites. The malarial mitochondrial genome thus appears quite unusual. Further investigations are expected to provide insights into the possible functional relationships between these molecules and perhaps their evolutionary history.

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.

1996

J. Mol. Biol. (1996) **261**, 155–172

JMB

Complete Gene Map of the Plastid-like DNA of the Malaria Parasite *Plasmodium falciparum*

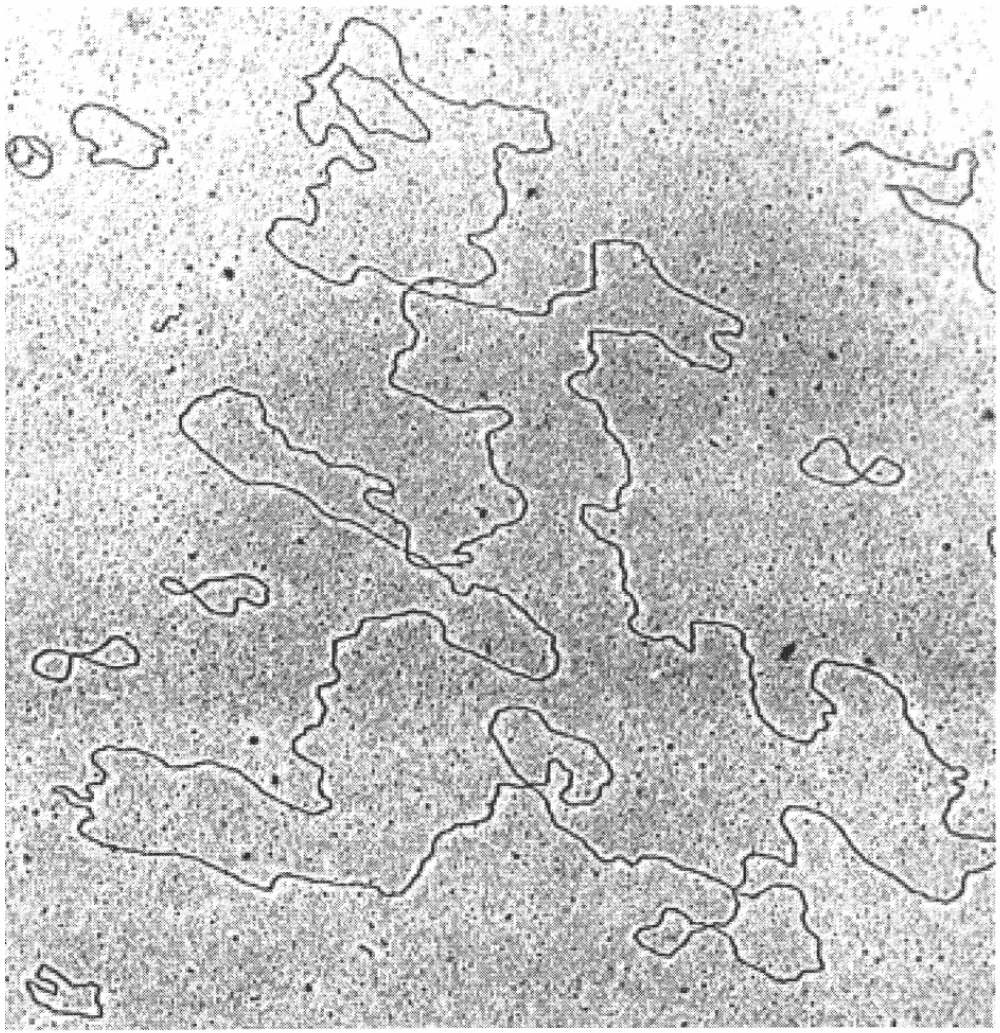
R. J. M. (Iain) Wilson*, Paul W. Denny, Peter R. Preiser Kaveri Rangachari, Kate Roberts, Anjana Roy, Andrea Whyte Malcolm Strath, Daphne J. Moore, Peter W. Moore and Donald H. Williamson

National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

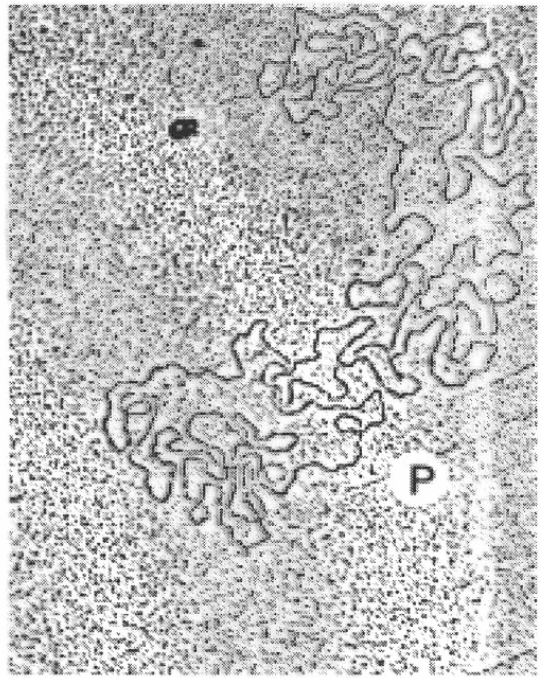
Malaria parasites, and other parasitic protists of the Phylum Apicomplexa, carry a plastid-like genome with greatly reduced sequence complexity. This 35 kb DNA circle resembles the plastid DNA of non-photosynthetic plants, encoding almost exclusively components involved in gene expression. The complete gene map described here includes genes for duplicated large and small subunit rRNAs, 25 species of tRNA, three subunits of a eubacterial RNA polymerase, 17 ribosomal proteins, and a translation elongation factor. In addition, it codes for an unusual member of the Clp family of chaperones, as well as an open reading frame of unknown function found in red algal plastids. Transcription is polycistronic. This plastid-like DNA molecule is conserved in several genera of apicomplexans and is conjectured to have been acquired by an early progenitor of the Phylum by secondary endosymbiosis. The function of the organelle (plastid) carrying this DNA remains obscure, but appears to be specified by genes transferred to the nucleus.

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.

1997



Lettuce circular CpDNA ~140 kb
(Kolodner and Tewari, 1975)



Plasmodium berghei ~35 kb
(Yap et al., 1997)

Synechocystis	3500 kb
Lettuce	140 kb
Plasmodium	35 kb

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.

1997

A Plastid of Probable Green Algal Origin in Apicomplexan Parasites

Sabine Köhler,*† Charles F. Delwiche,*‡ Paul W. Denny, Lewis G. Tilney, Paul Webster, R. J. M. Wilson, Jeffrey D. Palmer, David S. Roos§

Protozoan parasites of the phylum Apicomplexa contain three genetic elements: the nuclear and mitochondrial genomes characteristic of virtually all eukaryotic cells and a 35-kilobase circular extrachromosomal DNA. In situ hybridization techniques were used to localize the 35-kilobase DNA of *Toxoplasma gondii* to a discrete organelle surrounded by four membranes. Phylogenetic analysis of the *tufA* gene encoded by the 35-kilobase genomes of coccidians *T. gondii* and *Eimeria tenella* and the malaria parasite *Plasmodium falciparum* grouped this organellar genome with cyanobacteria and plastids, showing consistent clustering with green algal plastids. Taken together, these observations indicate that the Apicomplexa acquired a plastid by secondary endosymbiosis, probably from a green alga.

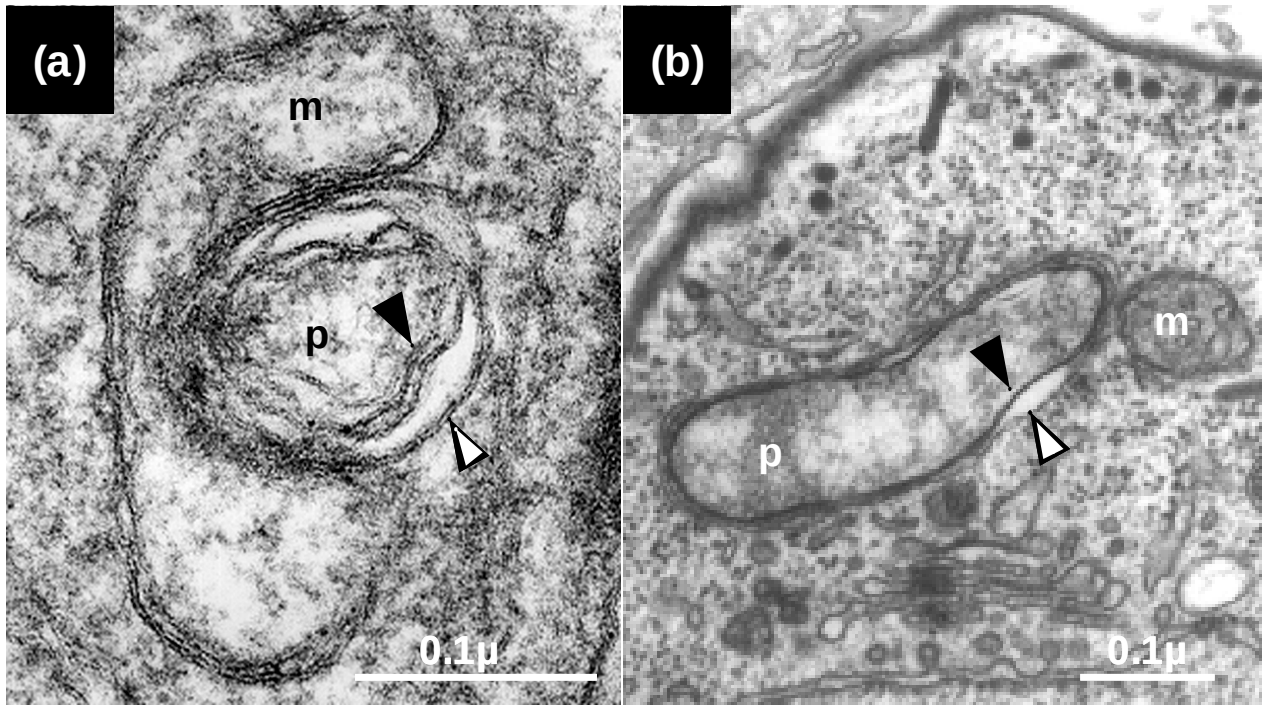
Apicomplexan parasites contain two maternally inherited extrachromosomal DNA elements (*1*). The mitochondrial genome is a multicopy element of ~6 to 7 kb encoding three proteins of the respiratory chain and extensively fragmented ribosomal RNAs (*2*). In addition, these parasites contain a 35-kb circular DNA molecule with no significant similarity to known mitochondrial genomes. The 35-kb element is similar to chloroplast genomes, containing an inverted repeat of ribosomal RNA genes and genes typically found in chloroplasts but not mitochondria (*rpoB/C*, *tufA*, and *clpC*) (*3*). The 35-kb DNA is also predicted to encode a complete set of tRNAs, numerous ribosomal proteins, and several unidentified open reading frames (*3*).

We used in situ hybridization to determine whether the 35-kb DNA is found within the parasite nucleus, mitochondrion, or cytoplasm or, alternatively, whether this

S. Köhler, L. G. Tilney, D. S. Roos, Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.
C. F. Delwiche and J. D. Palmer, Department of Biology, Indiana University, Bloomington, IN 47405, USA.
P. W. Denny and R. J. M. Wilson, National Institute for Medical Research, Mill Hill, London NW7 1AA UK.
P. Webster, Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520, USA.

*These authors contributed equally to this work.
†Present address: School of Pharmacy, University of California, San Francisco, CA 94143, USA.
‡Present address: Department of Plant Biology, University of Maryland, College Park, MD 20742, USA.
§To whom correspondence should be addressed. E-mail: droos@sas.upenn.edu

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.



Plasmodium *Toxoplasma*

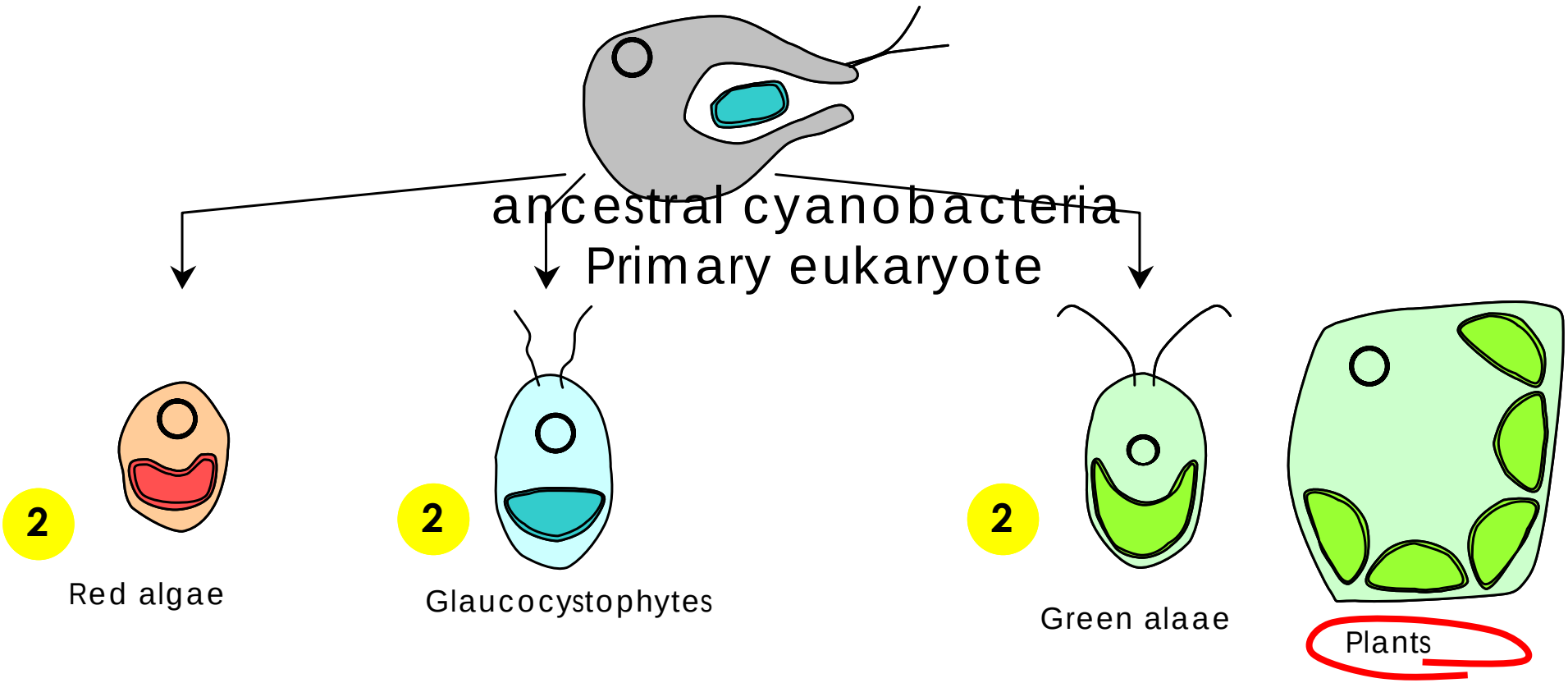
(Apicomplexans)

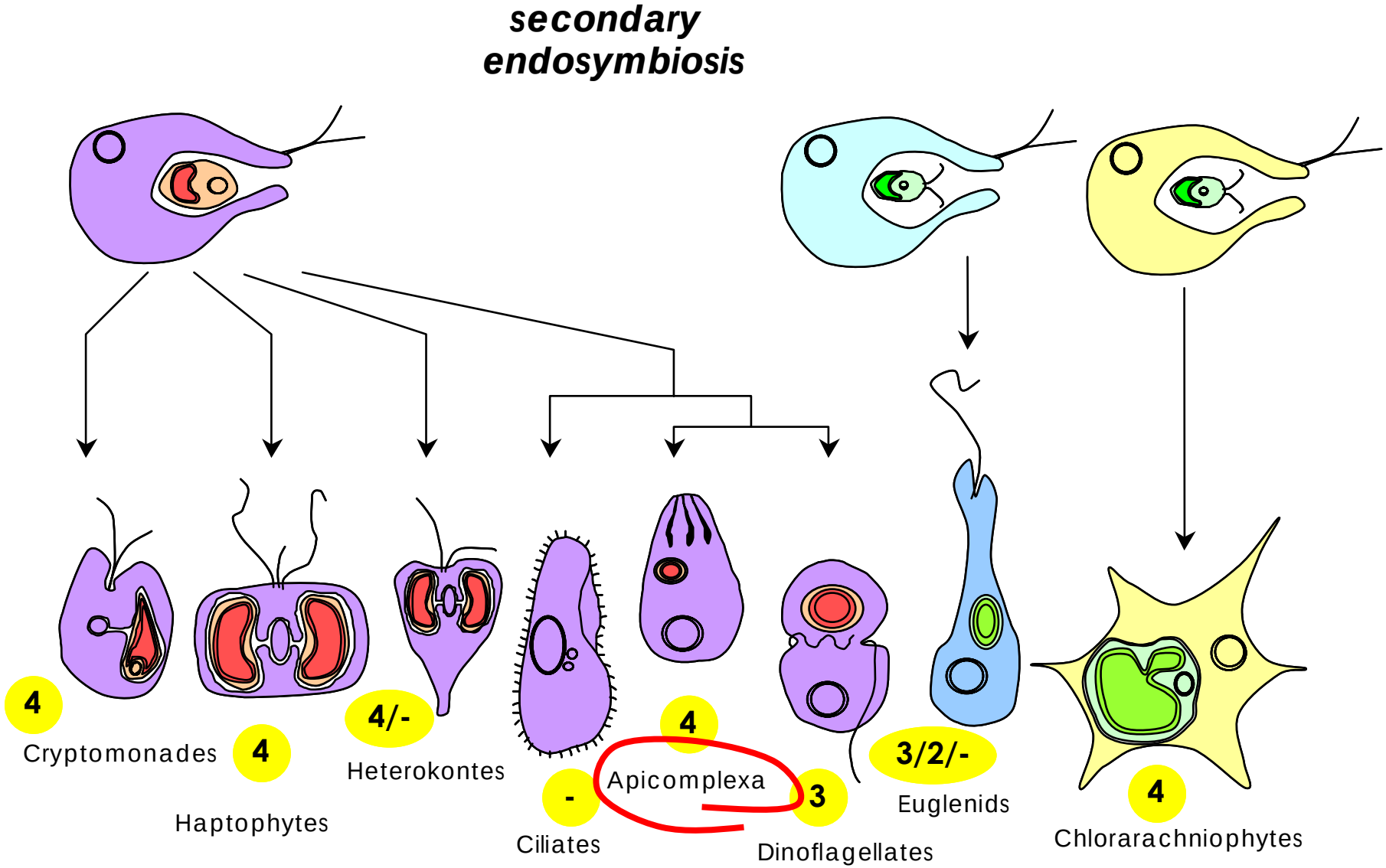
The apicoplast, a plastid limited by 3-4 membranes

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.

What is the apicoplast origin ?

primary endosymbiosis




secondary
endosymbiosis
4
Cryptomonades
4
Haptophytes
4/-
Heterokontes
-
Ciliates
4
Apicomplexa
3
Dinoflagellates
3/2/-
Euglenids
4
Chlorarachniophytes

The two contentious groups include parasites that cause diseases such as malaria and sleeping sickness; Adl hopes his classification will aid drug development. "Placing these organisms in the wrong group is in part responsible for the fact that we do not have specific drugs for these diseases, because of wrong assumptions about their biochemistry," he says.

fairly troubled. But the evidence was not clear enough for a consensus on a new regime, and fierce disagreements became common.

"I've seen people throw things at each other," says Sina Adl, a soil-organism specialist at Dalhousie University in Halifax, Canada, who coordinated the group of 28 protist experts that produced the new classification. It was commissioned by the International Society of

Trypanosomes (purple), which cause sleeping sickness, were one of the hardest protists to classify.

Protistologists and is published in *The Journal of Eukaryotic Microbiology* (S. M. Adl *et al. J. Eukaryot. Microbiol.* **52**, 399–451; 2005).

The experts have given protists a mighty four kingdoms out of six. Animals do not even get their own group — fungi and animals have

The two contentious groups include parasites that cause diseases such as malaria and sleeping sickness; Adl hopes his classification will aid drug development. "Placing these organisms in the wrong group is in part responsible for the fact that we do not have specific drugs for these diseases, because of wrong assumptions about their biochemistry," he says.

fairly troubled. But the evidence was not clear enough for a consensus on a new regime, and fierce disagreements became common.

"I've seen people throw things at each other," says Sina Adl, a soil-organism specialist at Dalhousie University in Halifax, Canada, who coordinated the group of 28 protist experts that produced the new classification. It was commissioned by the International Society of

Trypanosomes (purple), which cause sleeping sickness, were one of the hardest protists to classify.

Protistologists and is published in *The Journal of Eukaryotic Microbiology* (S. M. Adl *et al. J. Eukaryot. Microbiol.* **52**, 399–451; 2005).

The experts have given protists a mighty four kingdoms out of six. Animals do not even get their own group — fungi and animals have

Malaria targets for new drugs ?
An example of a drug discovery project,
combining bioinformatics and experiments.

1999

Is the apicoplast a target ?

- **Fosmidomycin**, a **herbicide** targeting a plant chloroplast process (*i.e.* the non-mevalonate isoprenoid pathway), kills the parasite (Jooma et al., 1999, Science).
- **Genetic impairment of apicoplast division** leads to cell with no apicoplast… that die (He et al., 2001, *EMBO J*)

Malaria targets for new drugs ?
An example of a drug discovery project, combining bioinformatics and experiments.

1997 1998 2000 2001 2002 2003 2004 2005 2008

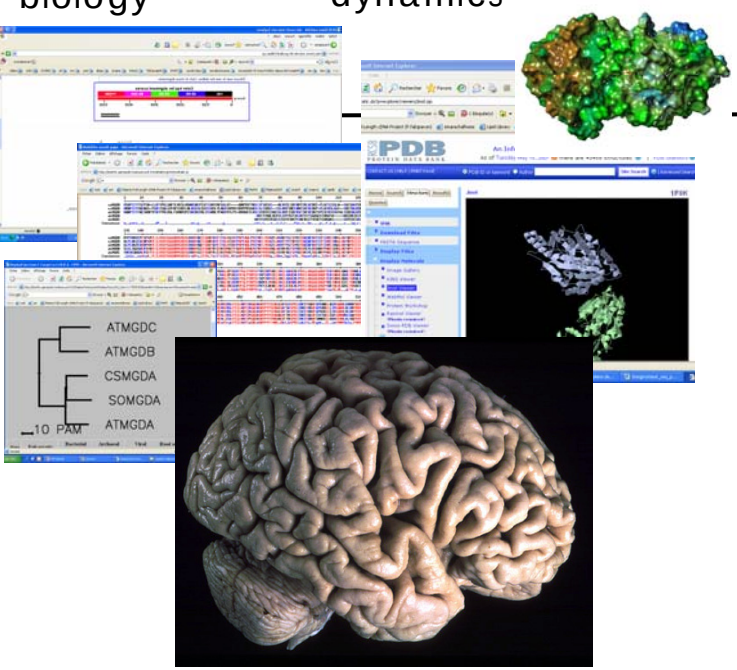
Target identification → **Structural and functional characterization** → **Pharmacological screening** → **Bioactivity assays Plasmodium, Toxoplasma, ...Plants** → **Human cell cytotoxicity** → **Molecule optimization**

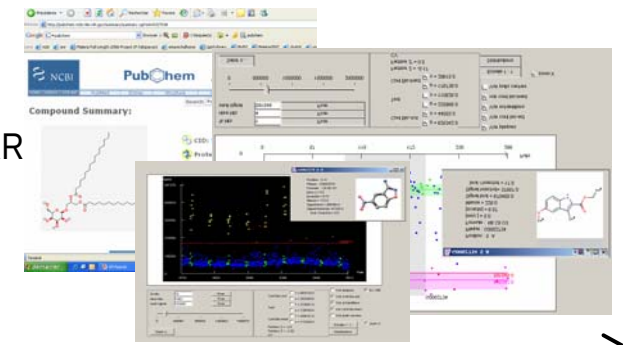
original idea: comparative genomics & biology

Structure prediction, molecular dynamics

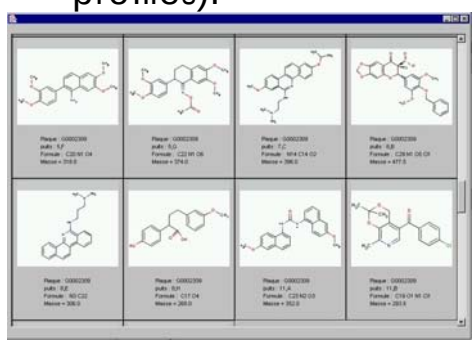
First QSAR analysis

second QSAR analysis (optimization of chemical descriptors constrained by complexe bioactivity profiles).





Molecular phylogeny

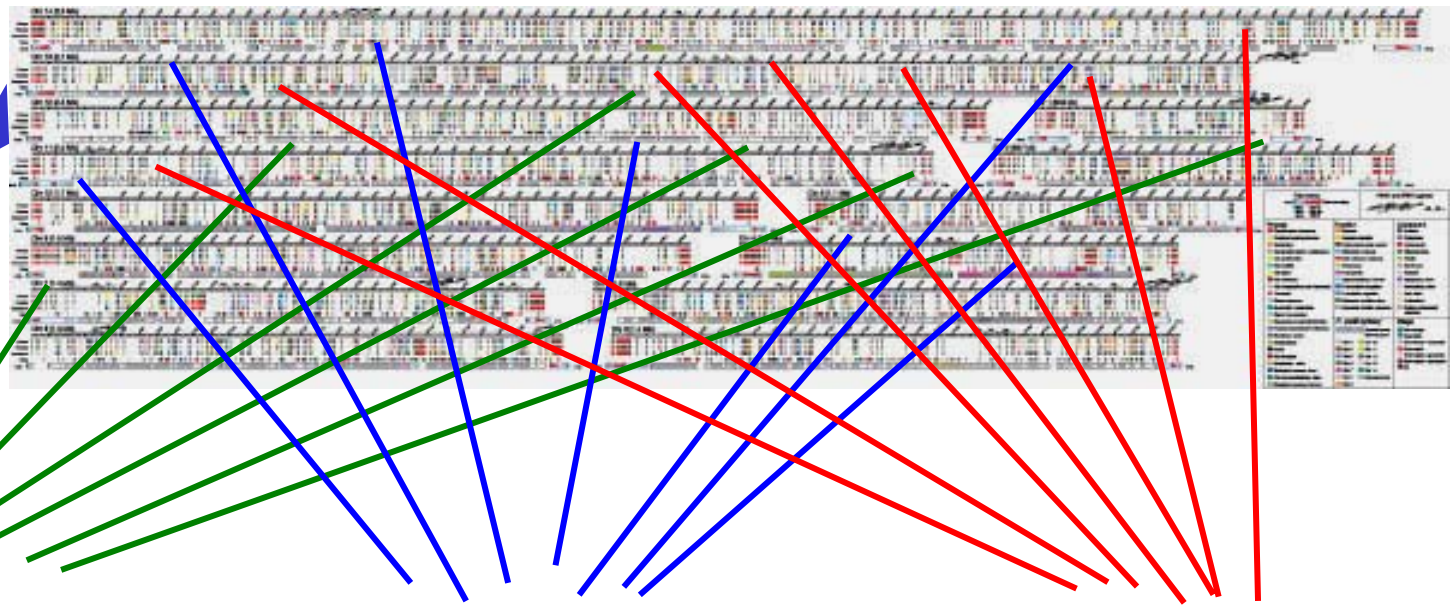


11 years
Marginal cost: 950 K€

Is the apicoplast a target ?
(in more general term any plant-like process in Plasmodium)

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons



Genes inherited from ancestral cyanobacteria

Genes inherited from primary eukaryote

Genes inherited from secondary endosymbionte

red algal sub-genome

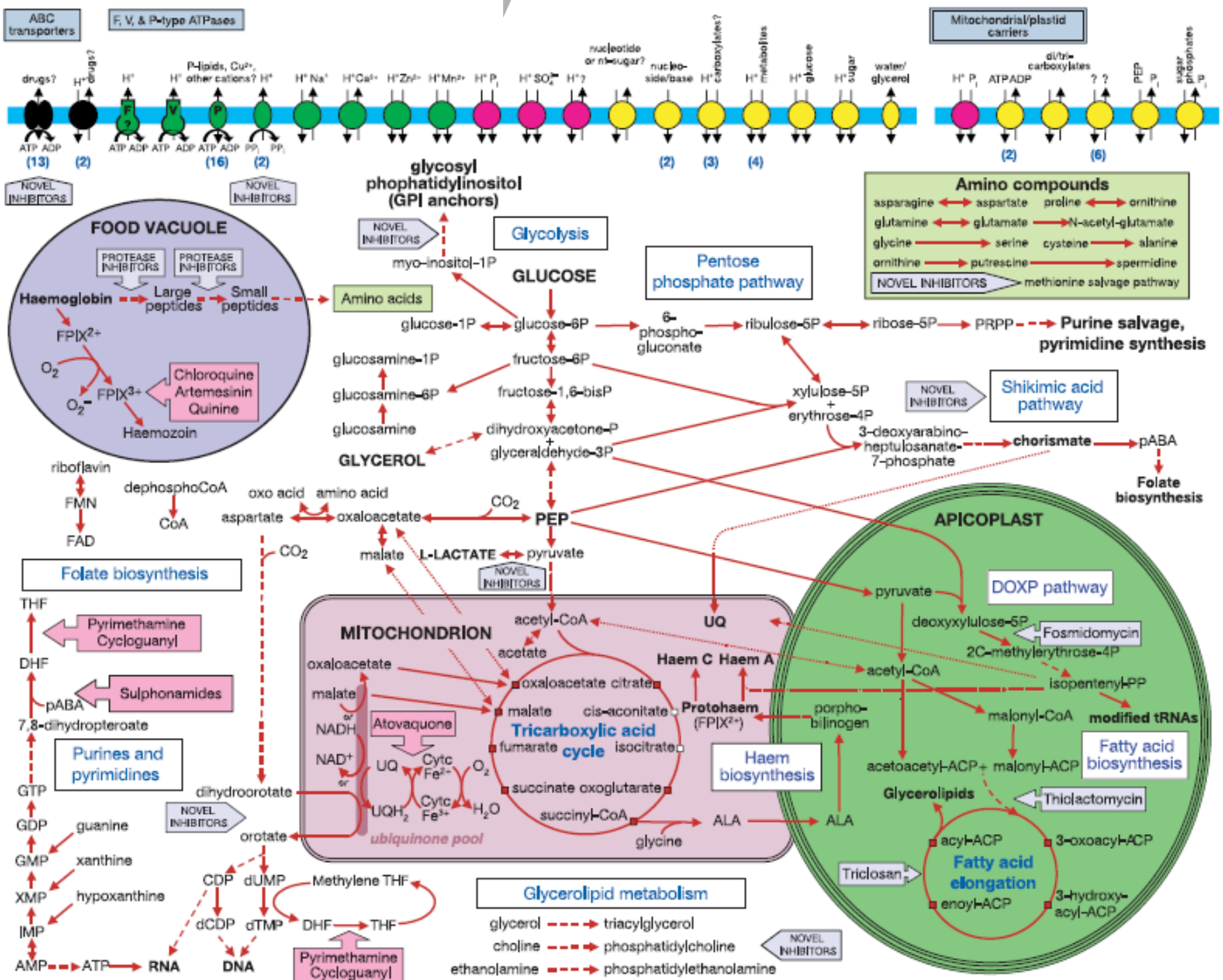
Plant side ?

Is the apicoplast a target ?
(in more general term any plant-like process in Plasmodium)

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes



Is the apicoplast a target ?
(in more general term any plant-like process in Plasmodium)

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes

Are there more we might have missed ?

CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes

Blast (+ motifs) searches

	Number of genes predicted to code for proteins (in original paper)	hit with genes with known function	hit with genes with no known function	no hit
Human (Craig et al., 2001)	~30,000			
Arabidopsis (A.G.I., 2000)	~25,000			
Plasmodium (Gardner et al., 2002)	~5,000		**~3200 sequences**	

CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes

Blast (+ motifs) searches

```
>87088036|chr11_s259
            Length = 17,424

  Plus Strand HSPs:

 Score = 79 (32.9 bits), Expect = 4.1, Sum P(2) = 0.98
 Identities = 23/76 (30%), Positives = 42/76 (55%), Frame = +2

Query:   441 IAGQEAGNVPYVIENGIG--KYLKSPKEIAKTVSQWFGPKANELQIMSQNALKHARPDAV 498
             I G+  G++  ++ N +   KY KSP  + KTVS  FG ++ E  ++     +H + +
Sbjct:  2423 INGKNMGDLQSMLNNLLNSEKYKKSPYTLDKTVSSGFGSRSKESMLLEY---QHNKSEQ- 2590

Query:   499 FKIVHDLDELV-RQKI 513
              +I+ +L EL+ + KI
Sbjct:  2591 -EILRELQELINKNKI 2635

 Score = 54 (24.1 bits), Expect = 4.1, Sum P(2) = 0.98
 Identities = 21/95 (22%), Positives = 45/95 (47%), Frame = +2

Query:    27 NSSLHGNNSNGYSSFSSNSVHFGGLATQNRYKFVNSLSFSKEGSNLKRILSDFNRVIRL- 85
             N++ + NN+N  SS +++S +    ++ N     N+ + +K  + ++++ + F+ +
Sbjct:  1745 NNNNNNNNNNNNSSNNNSSSNNNNSSSNNNNSSNNNNNNAKAANPMEQLTNLFSHINNND 1924

Query:    86 HCDRI----PLGFSSIGLNSGESNGVSDNGHGVLE 116
             H D      PL         G +NG +  G+  LE
Sbjct:  1925 HHDEKGDGHPLEHIMNLFGQGHNNGDAKGGNNPLE 2029
```

~3200 sequences

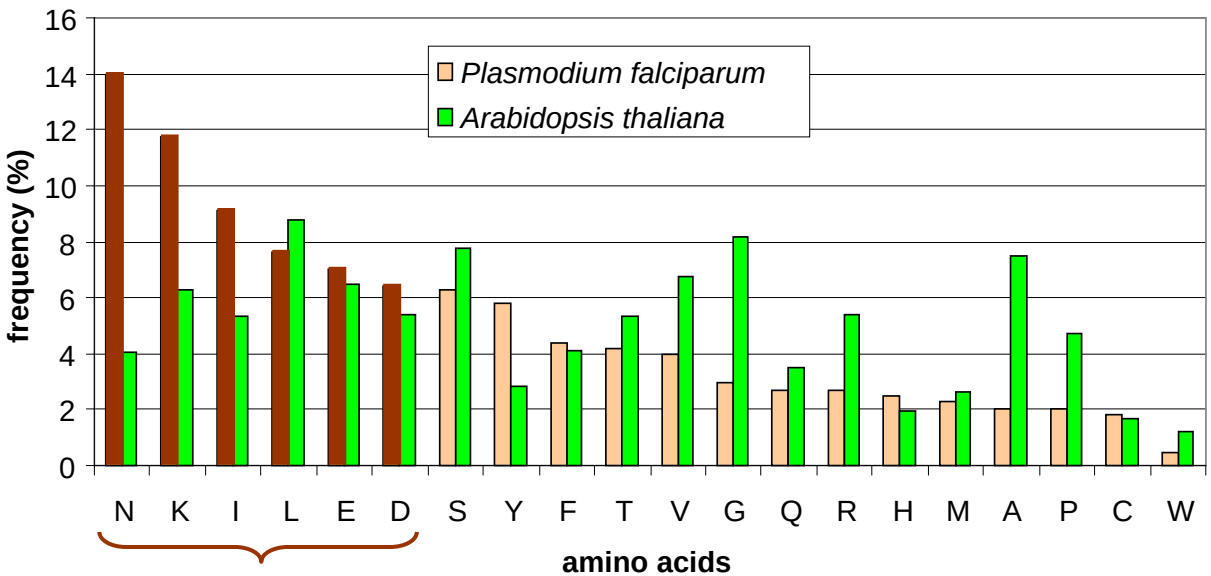
CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes

Blast (+ motifs) searches



> 50% protein content

~3200 sequences

CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular

sequence comparisons

analyses of biological processes

	A	R	N	D	C
A	5	-3	-2	-3	0
R	-1	6	-4	-1	-2
N	-1	0	4	1	-1
D	-2	-3	0	5	-4
C	1	-3	-2	-4	8

abidopsis thaliana

Blast (+ motifs) searches

	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
	-3	0	-2	-2	-2	-4	-2	-3	-1	0	0	-1	-3	0
	-2	-3	0	-3	-5	1	0	-7	-1	-1	0	-3	-5	-3
	0	0	0	-3	-2	0	-2	-3	0	0	0	-2	-2	-2
	1	-1	0	-3	-5	0	-1	-5	0	0	-1	-6	-5	-3
	-5	0	-2	-1	0	-5	-1	-2	-2	0	0	0	0	0
	0	-2	0	-5	-1	0	0	-6	0	0	-2	0	0	-3
	4	-1	-1	-4	-3	0	-2	-5	0	0	0	0	-4	-2
	-4	7	-1	-3	-5	-4	-4	-5	-3	-2	-2	-5	-4	-5
	0	-2	7	-4	-2	-3	-2	-1	-1	0	-1	0	1	-2
	-2	-4	-3	4	1	-2	0	0	-1	-2	0	0	-2	2
	-3	-4	-2	1	4	-3	2	0	-5	-2	-1	0	-1	0
	0	-2	0	-3	-2	4	0	-5	0	0	0	-5	-2	-2
	-4	-3	0	0	2	-2	6	0	-1	-1	-1	-4	-3	0
	-3	-3	-1	0	0	-4	0	6	-3	-3	-4	1	2	0
	-1	-5	-1	-7	-3	-1	-4	-1	7	-1	-1	-4	-2	-3
	0	0	-1	-3	-5	0	-2	-2	0	4	0	-3	-3	-3
T	0	0	-1	-1	-1	0	0	-1	-1	-1	-1	-1	0	-2	-1	1	5	-4	-5	0
W	-6	-1	-5	-3	0	-4	-4	-3	-1	-2	0	-3	0	3	-5	-3	-3	11	2	-4
Y	-1	0	-2	-2	-1	-1	-1	-2	1	-2	0	-2	0	3	-2	-1	-1	3	6	-1
V	0	-3	-5	-7	0	-3	-2	-2	-4	2	0	-2	0	0	-2	-3	0	-2	-2	4

Pla

~3200 sequences

CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes

Blast (+ motifs) searches

⇒ **Sequence alignement methods ?**

- ✓ **Substitution matrices ?**
  - **Non-symetric matrices ⇒ in progress**
- ✓ **Score statistics?**
  - **E-value vs Z-score** ⇒ **Z-score**
    ⇒ **Use at massive scale in progress**

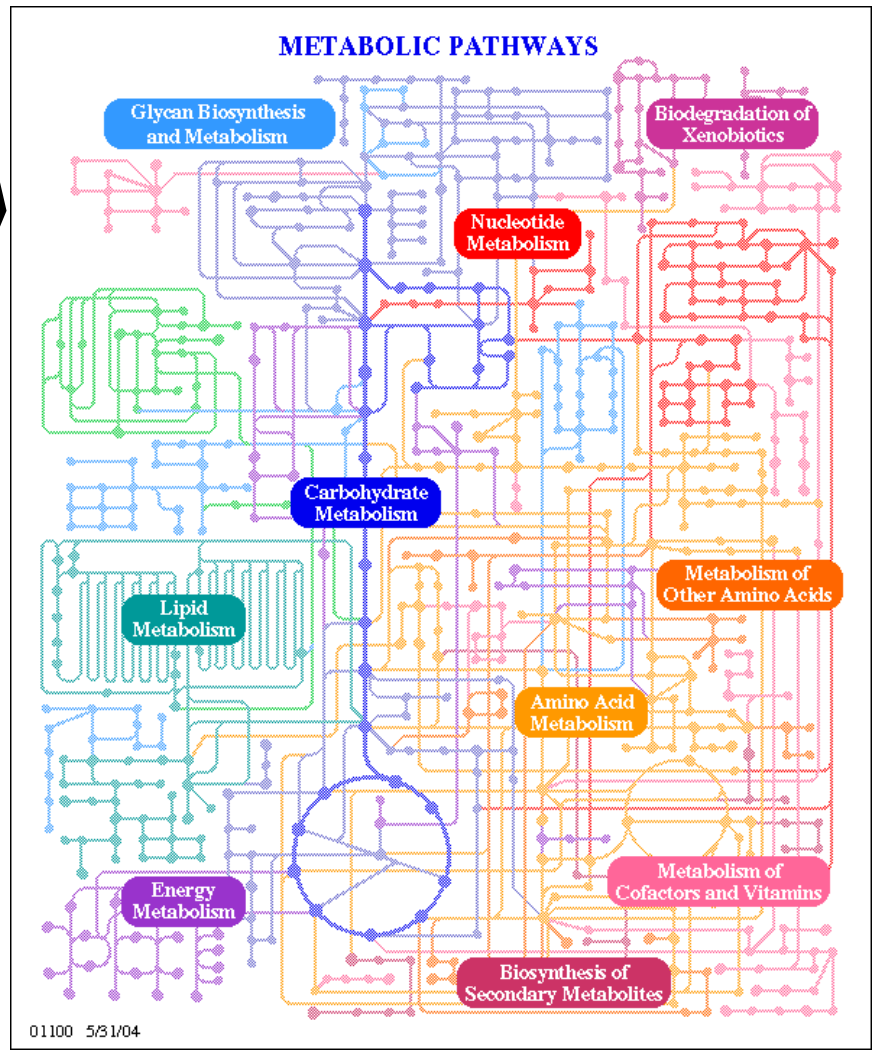
~3200 sequences

CURRENT LIMITS

Available molecular data PlasmoDB, etc. + comparison with other organisms (in particular plants)

sequence comparisons

analyses of biological processes



FUTURE CHALLENGES

Potential of *in silico* approaches for target discovery ?

⇒ more integration of genomic/post-genomic molecular and functional data
⇒ from user-friendly, but rigid, web portals to flexible & creativity-oriented knowledge accesses (with user-designed workflows, implying a strong interoperability)
⇒ improved and new analyzing and mining tools

⇒ integration of biological and chemical information (chemogenomic knowledge) and models for biological and chemical space intersects

- what biological space ? what chemical space ?
- bridging biological targets → ligand
  - bridge #1: recorded effect of drugs/small molecules on known targets
  - bridge # 2: protein structure –small molecule structure (docking)
- bridging small molecule → biological responses
  - OMIC functional responses
- implementing new predictions (*i.e.* druggability)

Comparative chemogenomics **CEA Grenoble**
C Botté, O Bastien, N Saïdani,
D. Grando, A. Zoppé, J Jouhet, H Valadié
MA. Block, E Maréchal

AI for chemogenomics+cheminformatics
CEA Grenoble
S. Wieczorek, S. Aci, S Roy
IMAG Grenoble
G. Bisson, M. Gordon
Université de Lille
D Horvath

Functional chemogenomics **CEA Grenoble**
L Lafanechère

Automated screening **Cerep, France**
E Nicolaï, F Revah

Virtual screening **IN2P3 Clermont-Ferrand**
V Kasam, V Breton

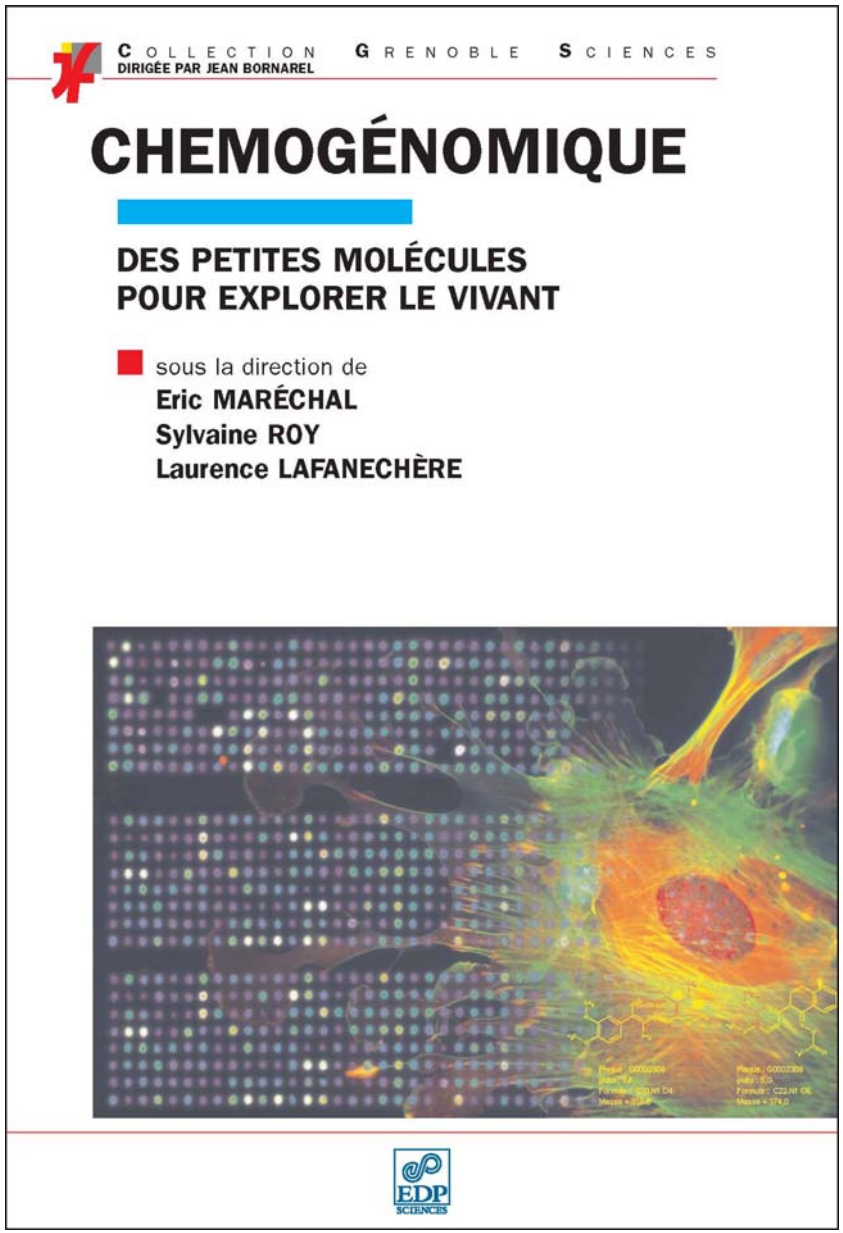
Diversity oriented chemistry **CEA Saclay**
M Deligny, AL Bonneau, R Lasselin,
B Rousseau, R Lopez
Faculté de Pharmacie Meylan
YS Wong

Structural biology **Cermav, Grenoble**
C Jouanneau, M Audry
A Imberty, C Breton

Parasitology **Grenoble**
C. Mercier, C. Bisanz,
MF. Cesbron-Delauw
Montpellier
JF Dubremetz, H Vial
Mexico
R Mondragon
Seattle
J Feagin,
Marburg
N. Azzouz, R Schwarz,

Ontological status of plants
Univ. J Fourier Grenoble
N Aumonier

Bioinformatics + Post genomics analyses
LIRMM Montpellier
O. Gascuel, L. Bréhélin & team
Gene-IT, France
K Metayer, JJ Codani
Pretoria, South Africa
LM Birkholtz, F Joubert, B. Louw
CEA Cadarache
P Ortet



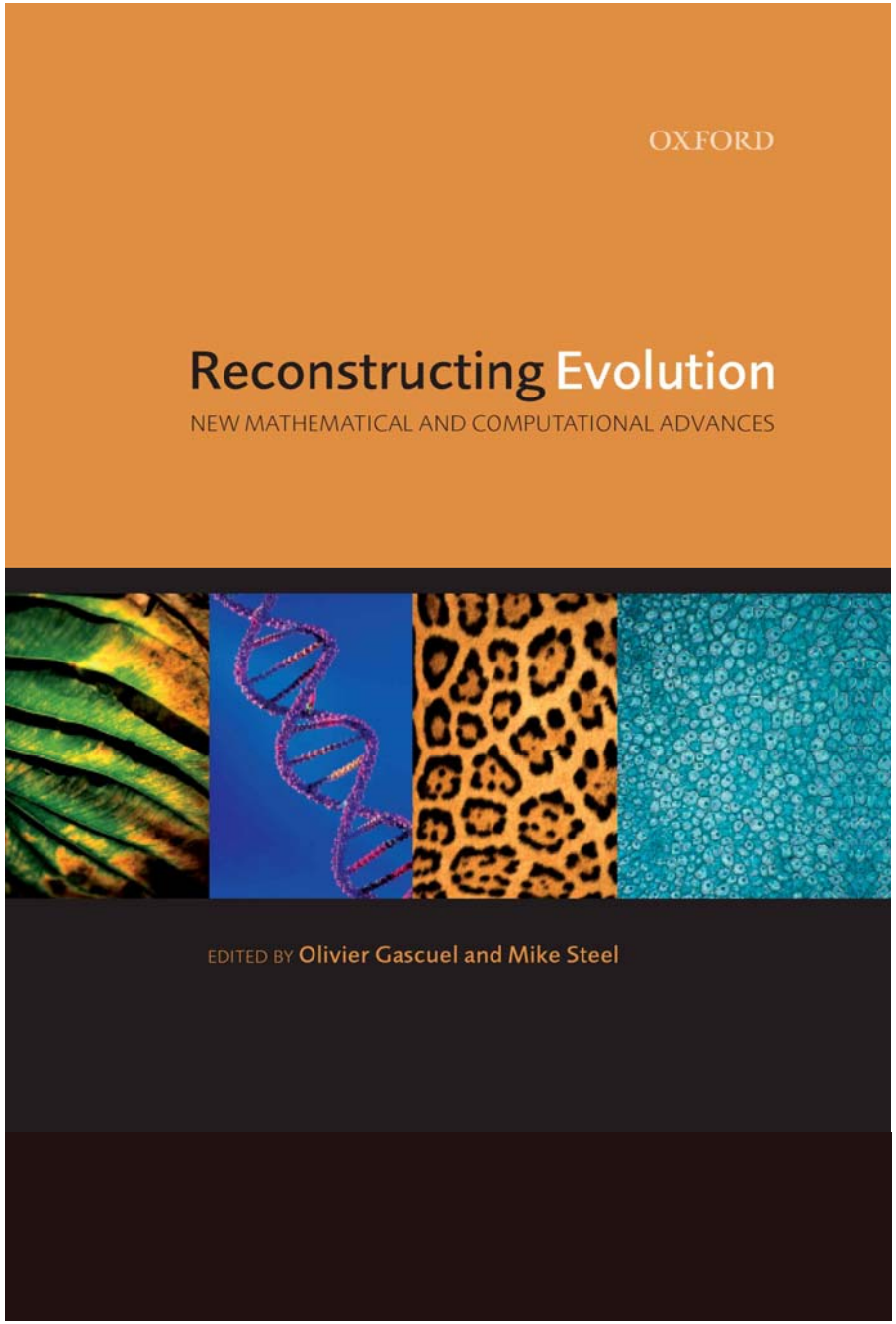
Chemogénomique,
des petites molécules pour explorer le vivant
(Eds. Maréchal, Roy, Lafanechère)
EDP Sciences, Collection Grenoble Science

An *introductory manual* for biologists, chemists and computer scientits
(Master Level ⇒ engineers, researchers)

- **Part I:** Automated pharmacological screening
- **Part II:** High content screening and chemical genetic strategies
- **Part III:** toward an ***in silico*** exploration of the chemical and biological spaces
- **Glossary to help biologists, chemists and computer scientists sharing a vocabulary**

16 authors (biologists, chemists, chemo- and bioinformaticians).

Chemogenomics,
Exploring the living world with small molecules
In progress



Reconstructing evolution
(Eds. Gascuel, Steel)
Oxford University Press