Bioinformatics and other advanced technologies for Africa and all developing regions

- Some general remarks about use and misuse of bioinformatics

(2) Initiatives that may be relevant for spreading advanced technologies in Africa and other southern areas

Some general remarks about use and misuse of bioinformatics

- Preliminary warnings
- A recommendable strategy for the (many) labs that don’t have big technologies
- Tasks of bioinformatics
- Traps
- The relevance of recombination in pathogens
- Relevant megaprogrammes in human genetics
- The gap between genotype and phenotype
- (Another) call for an integrated approach

Preliminary warnings

Relevance of field data

Never forget environment and socioeconomical factors

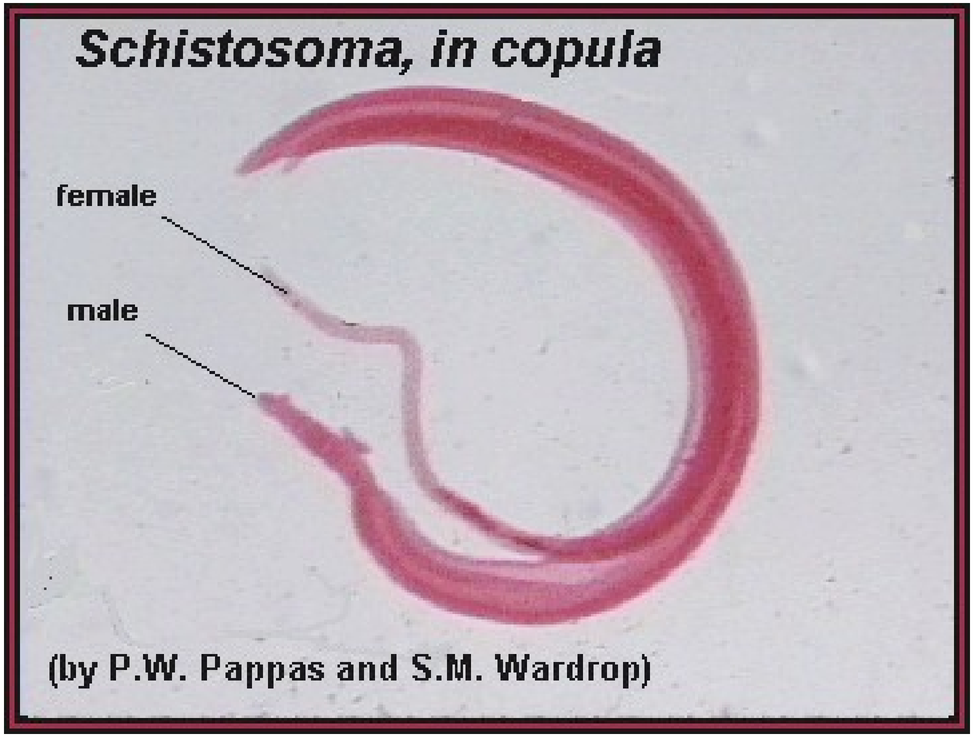
Avoid to make it a big fishing expedition

Deductive research (design working hypotheses and collect data to falsify them)

Vs

Inductive research (collect data and see what’s going on)

Molecular Idiots



Schistosoma, in copula

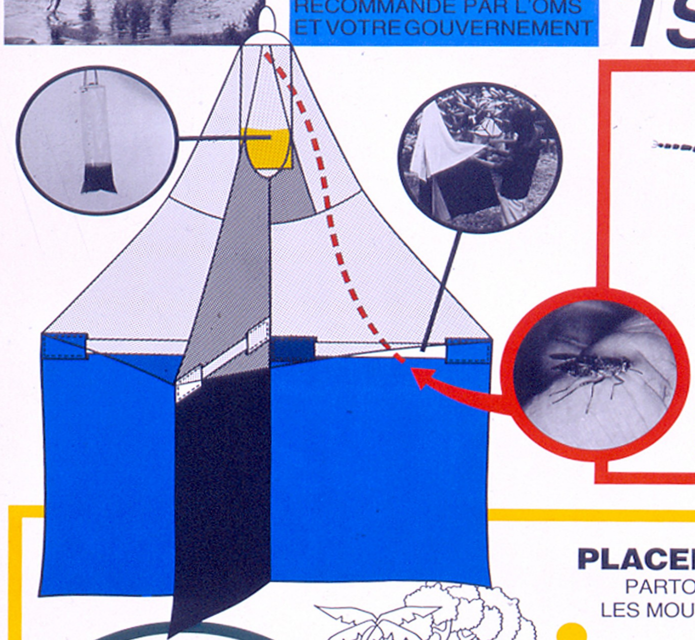
(by P.W. Pappas and S.M. Wardrop)


A
B
C
D
E

Don’t forget the dear old things

POUR LUTTER DANS VOTRE VILLAGE
CONTRE LA
MALADIE DU SOMMEIL
INSTALLEZ DES
PIEGES A MOUCHES
TSÉ-TSÉ

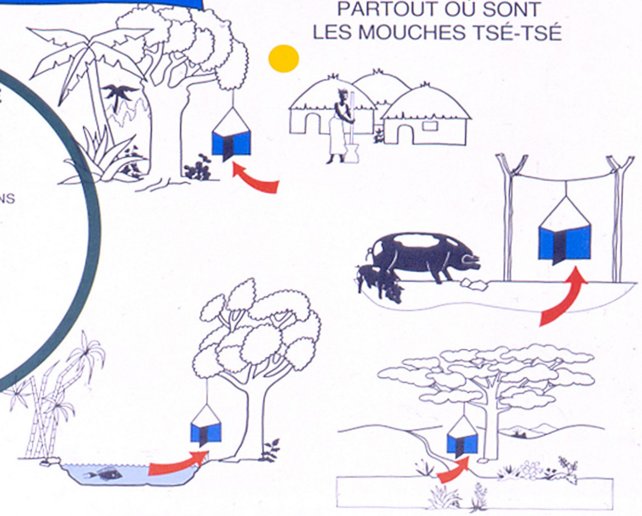
RECOMMANDÉ PAR L'OMS
ET VOTRE GOUVERNEMENT



SA PIQURE PEUT TRANSMETTRE
LA MALADIE DU SOMMEIL
CETTE MALADIE EST MORTELLE
SI ELLE N'EST PAS SOIGNÉE

PLACER LE PIÈGE

PARTOUT OÙ SONT
LES MOUCHES TSÉ-TSÉ



TOUT LE VILLAGE DOIT PARTICIPER

EN NETTOYANT AUTOUR
DU PIEGE ET
EN COMPTANT LES MOUCHES
CAPTUREES CHAQUE MOIS DANS
LE SAC PLASTIQUE

ORSTOM
INSTITUT FRANCAIS DE RECHERCHE SCIENTIFIQUE
POUR LE DEVELOPPEMENT EN COOPERATION

J. LANCIEN

LA COOPERATIVE DE CREATION

Do not forget

Bioinformatics is not only for molecular data:

- Computer-assisted taxonomy

- Morphometrics

A recommendable strategy for the (many) labs that don’t have big technologies

The remora’s strategy

What may have « weak » teams to sell to big sharks?

Biodiversity

Data

Strain banks

Access to the field

Field and clinical expertise

Other complementary expertise:
population genetics, etc

Little remoras, don’t sell yourself for nothing:

Data mining = Gold mining

Tasks of bioinformatics

Either: look for individual relevant genes (susceptibility to disease, drug resistance)

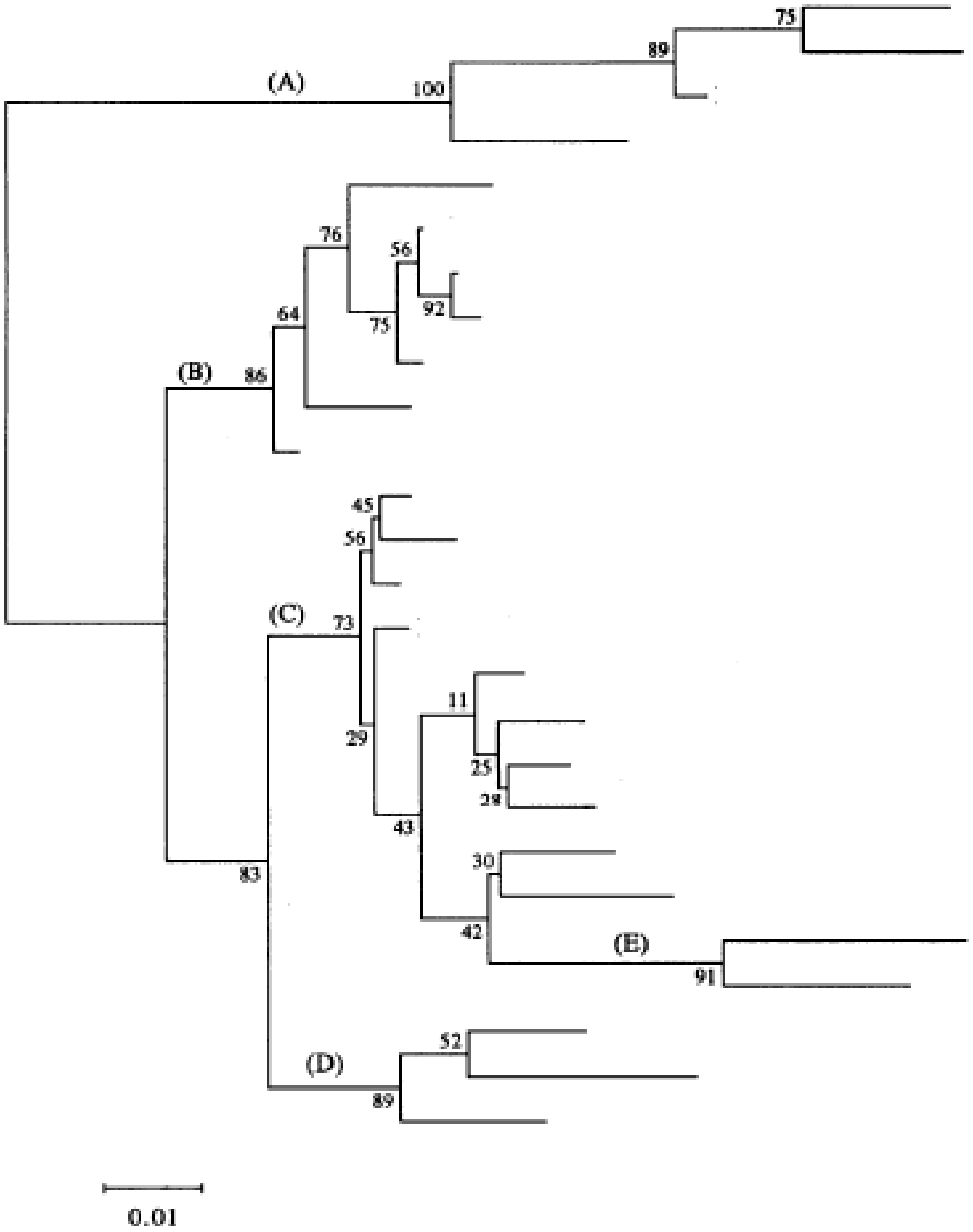
Or (better): draw a comprehensive survey of the whole genetic variability of the species under study

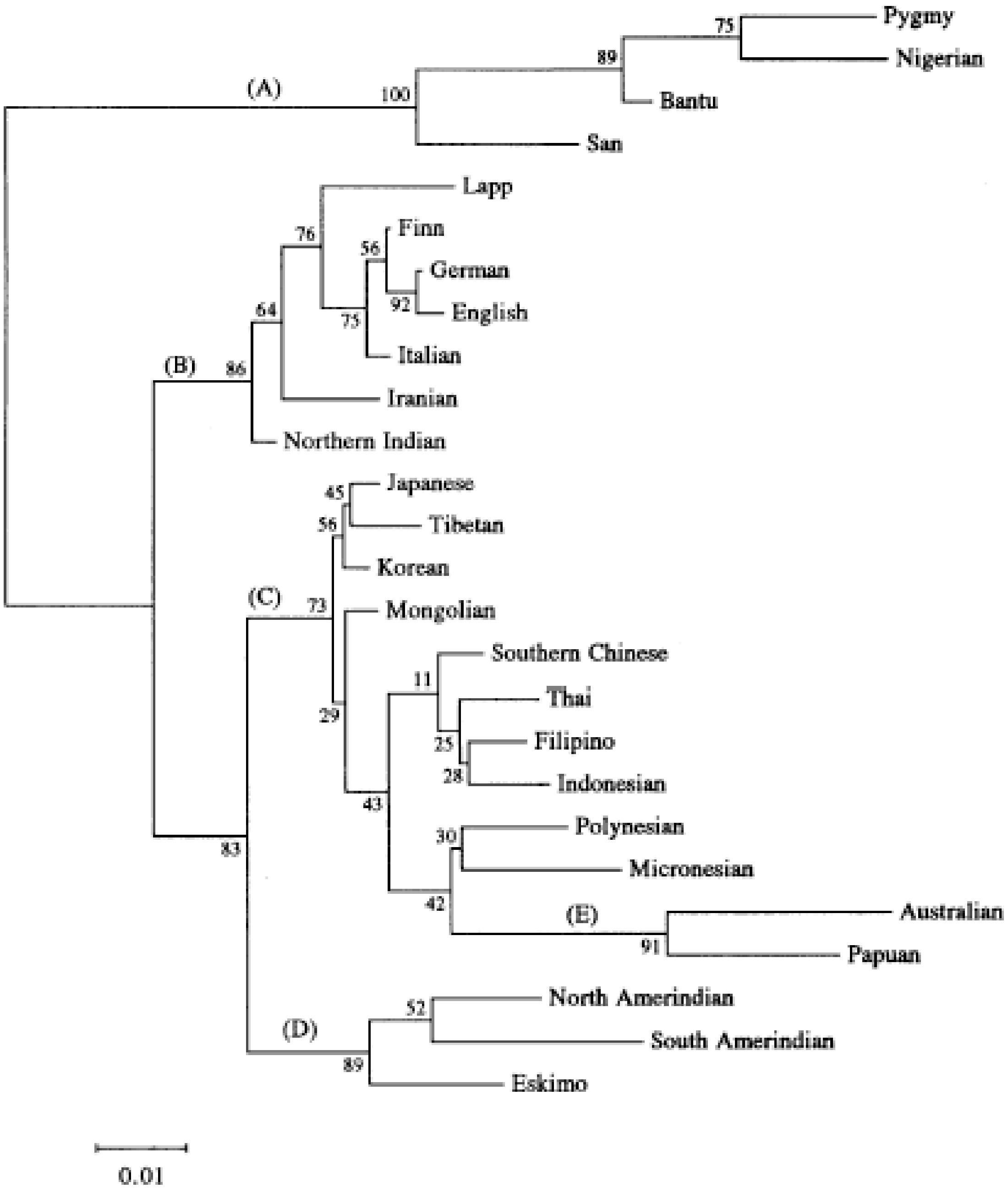
Phylogenetic Character Mapping or PCM (Avise, 2006)

Assignment of relevant characters on the phylogenetic picture of the species

Traps

Bootstrap religion


75
89
(A)
100
76
56
92
64
75
(B)
86
45
56
(C)
73
11
29
25
28
43
30
83
42
(E)
91
52
(D)
89
0.01


75
Pygmy
89
Nigerian
(A)
100
Bantu
San
Lapp
76
Finn
56
German
92
English
64
75
Italian
(B)
86
Iranian
Northern Indian
45
Japanese
56
Tibetan
Korean
(C)
73
Mongolian
Southern Chinese
11
Thai
29
25
Filipino
28
Indonesian
43
Polynesian
30
Micronesian
42
(E)
Australian
91
Papuan
83
North Amerindian
52
(D)
South Amerindian
89
Eskimo
0.01

Gene trees and species trees are not the same

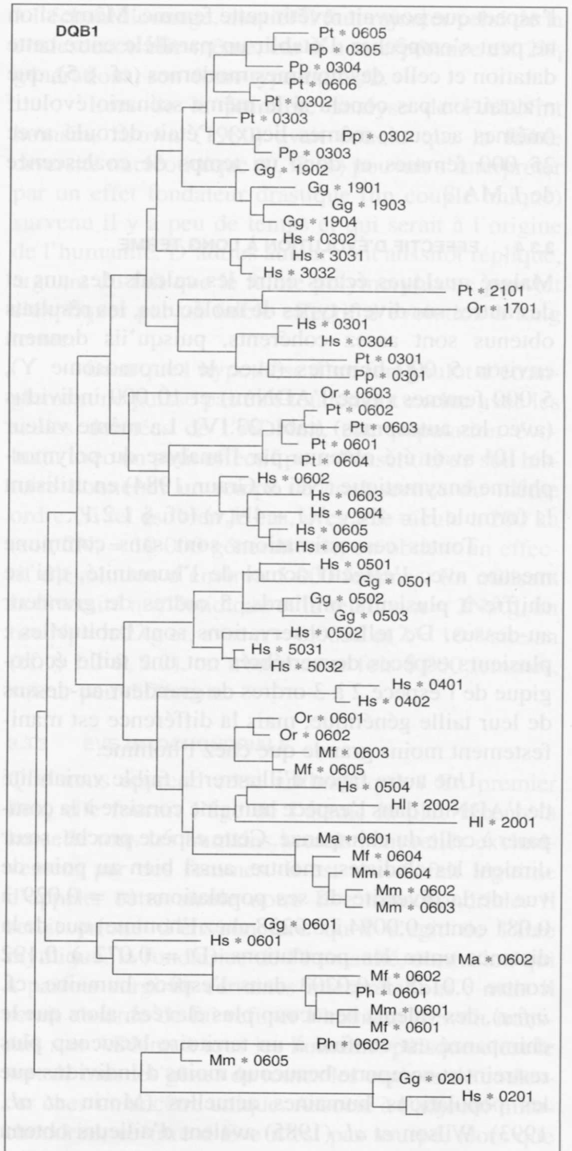


Figure 23.5 : Arbre phylogénétique de 64 allèles du locus *DQB1* chez 10 espèces de primates. L'arbre a été obtenu par la méthode du *neighbor-joining* sur des distances évaluées par la méthode de Kimura sur les 270 nucléotides de l'exon 2. Hs, homme ; Pp et Pt, chimpanzés ; Gg, gorille ; Or, orang-outan ; Hl, gibbon ; Ma, Mf, Mm et Ph, macaques et hamadryas.
(D'après Ayala & Escalante, 1996.)

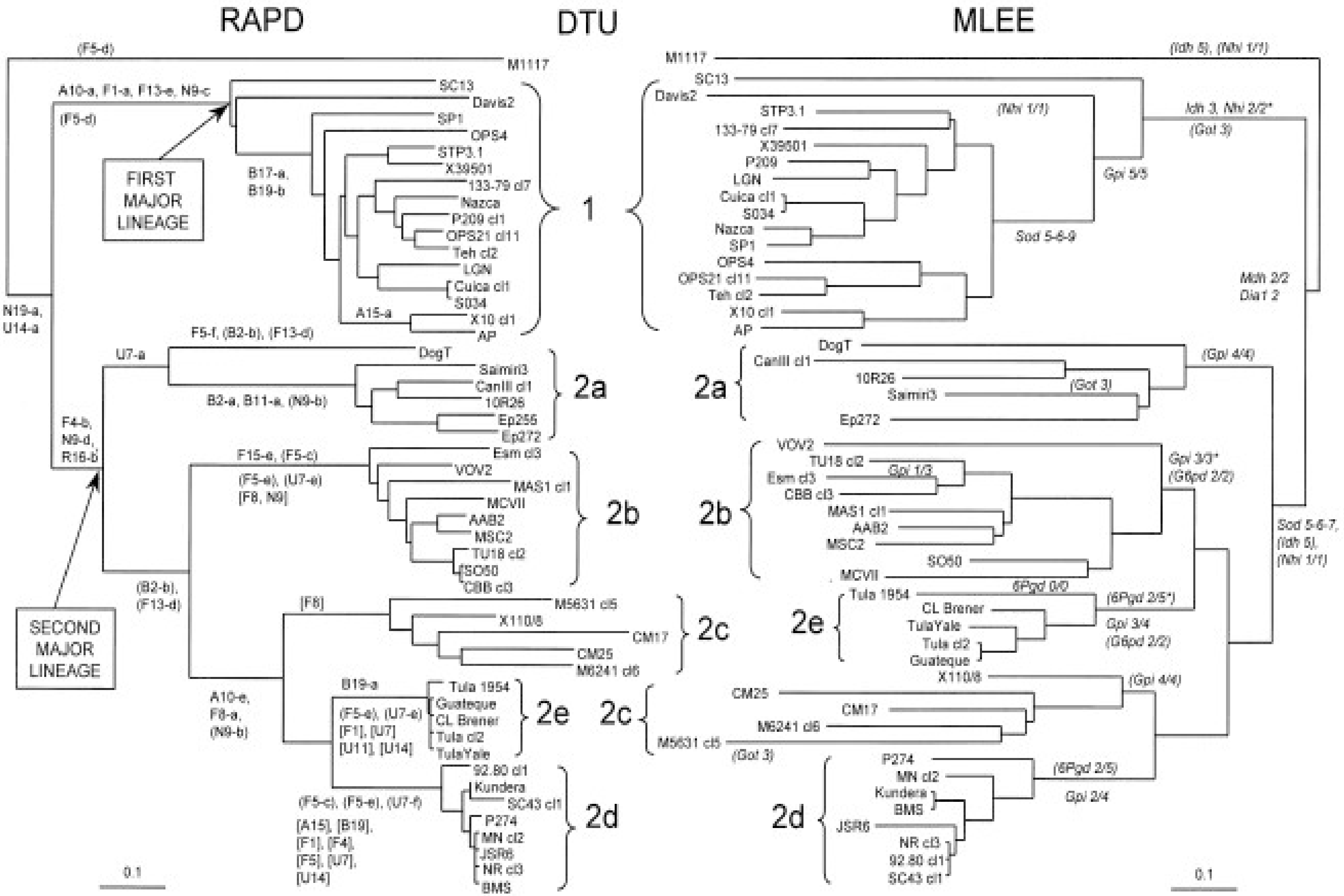
Typologist approach

A major concern for sequencers

Typologism

The approach that considers that each species is well represented by an ideal “«type”» – the specimen used for the first description – and that all variability around this type is to be ignored

Brisse, S., Barnabé, C. & Tibayrenc, M. 2000. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. *Int. J. parasitol*. 30: 35-44.



What are the limits of the entity under study?

What is a species?

What is a strain?

Tibayrenc, M. 2006. The species concept in parasites and other pathogens: a pragmatic approach? *Trends Parasitol.* 22: 66-70.

Recombination rate in pathogens: sexual vs clonal?

Highly relevant for data mining and phylogenetic character mapping

Recombining species

Plasmodium falciparum: highly recombining, although able of some clonal propagation

No linkage disequilibrium between characters: they are transmitted independently: no strain typing

Predominantly clonal species

Trypanosoma cruzi

Tibayrenc, M., Ward, P., Moya, A. & Ayala, F.J. 1986. Natural populations of *Trypanosoma cruzi*, the agent of Chagas'disease, have a complex multiclonal structure. *Proc. Nat. Acad. Sci. USA.* 83: 115-119.

Gaunt, M.W., Yeo, M., Frame, I.A., Tothard, J.R., Carrasco, H.J., Taylor, M.C., Mena, S.S., Veazey, P., Miles, G.A., Acosta, N., Rojas de Arias, A. & Miles, M.A. 2003. Mechanism of genetic exchange in American trypanosomes. *Nature* 421: 936-939.

Linkage disequilibrium (nonrandom association)

Between genetic characters

Between genetic and phenotypic characters

Reticulate evolution in pathogens

In most if not all pathogens, there is a combination of clonal propagation and genetic exchange

To be taken into account in phylogenetic studies and Phylogenetic Character Mapping

Relevance of population genetics

Challenge: population genomics

The ultimate cure against typologism

Relevant megaprogrammes in human genetics

Human diversity genome project (HDGP)

HapMap project

Couzin, J. 2006. *Science* 312: 1131

GENOMICS

The HapMap Gold Rush: Researchers Mine a Rich Deposit

Scientists are parsing a raft of new data on genetic variation for clues to disease and evolution

CAMBRIDGE, MASSACHUSETTS—For a conference on the next generation in genomics, the setting was just right: a pristine auditorium in a gleaming new building near the Massachusetts Institute of Technology (MIT). More than 200 people gathered here at the Broad Institute earlier this month to discuss the HapMap, a database cataloging human genetic variation. Begun in 2002, the map has been assembled primarily to boost the analysis of inheritance using pieces of DNA that are often transmitted as intact blocks.

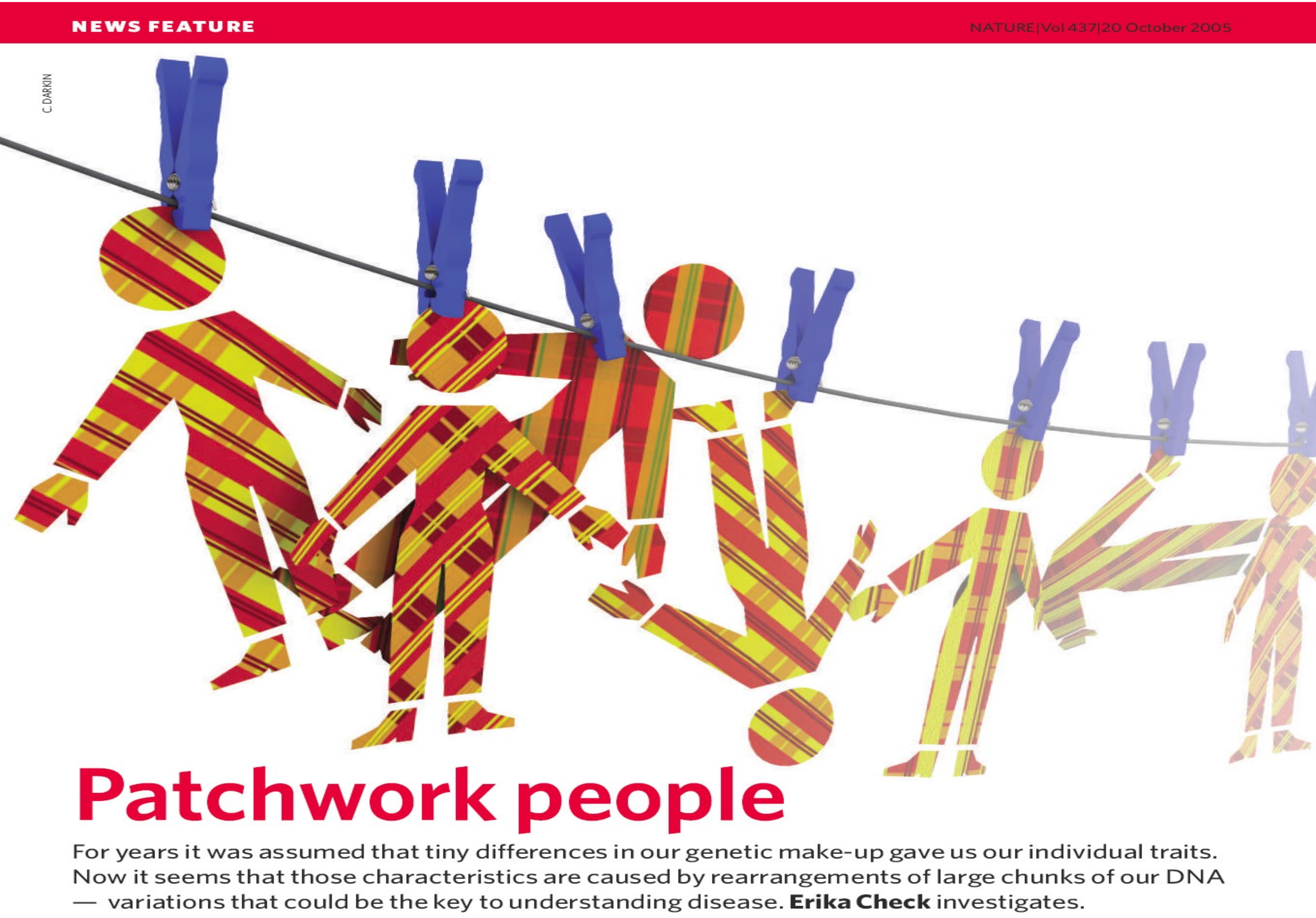
Nearly complete, the HapMap is now being tested for a number of uses: to find genetic variants behind common diseases, to examine the genome's architecture, and to study natural selection. The human HapMap has even inspired the launch of a parallel effort for *Plasmodium falciparum*, the deadly malaria parasite.

Five countries kicked in about $138 million to fund the human project, properly known as the International HapMap Project. One early challenge was to allow for the fact that haplotypes differ somewhat across populations. To include a sweep of variants, the HapMap gathered DNA from 270 individuals of African, Japanese, Chinese, and European ancestry.

Deciphering disease. Aided by the HapMap, researchers are finding gene variants that may help explain diabetes and other conditions.

Is DNA sequence so relevant?

Check E. 2005. *Nature* 437: 1084-1086



NEWS FEATURE

NATURE|Vol 437|20 October 2005

C.DARKIN

Patchwork people

For years it was assumed that tiny differences in our genetic make-up gave us our individual traits. Now it seems that those characteristics are caused by rearrangements of large chunks of our DNA — variations that could be the key to understanding disease. **Erika Check** investigates.

The abyss between genotype and phenotype

Integrated approach

Impact, on the transmission and severity of infectious diseases, of the genetic variability of

The host

The pathogen

The vector

A unique biological phenomenon (coevolution)

Durty propaganda:

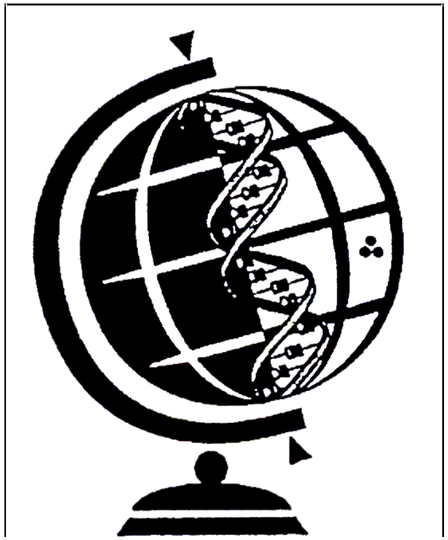
How to circulate bioinformatics and other advanced technologies?

International congresses Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases

(MEEGID)

I: CDC Atlanta 1996
II: Montpellier 1997
III: Rio de Janeiro 1998
IV: Dakar 1999
V: Hyderabad 2000
VI: Pasteur Institute Paris 2002
VII: Valencia 2004
VIII: Bangkok 2006
IX: Nairobi 2008

Meeting Announcement



9th International Congress on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases

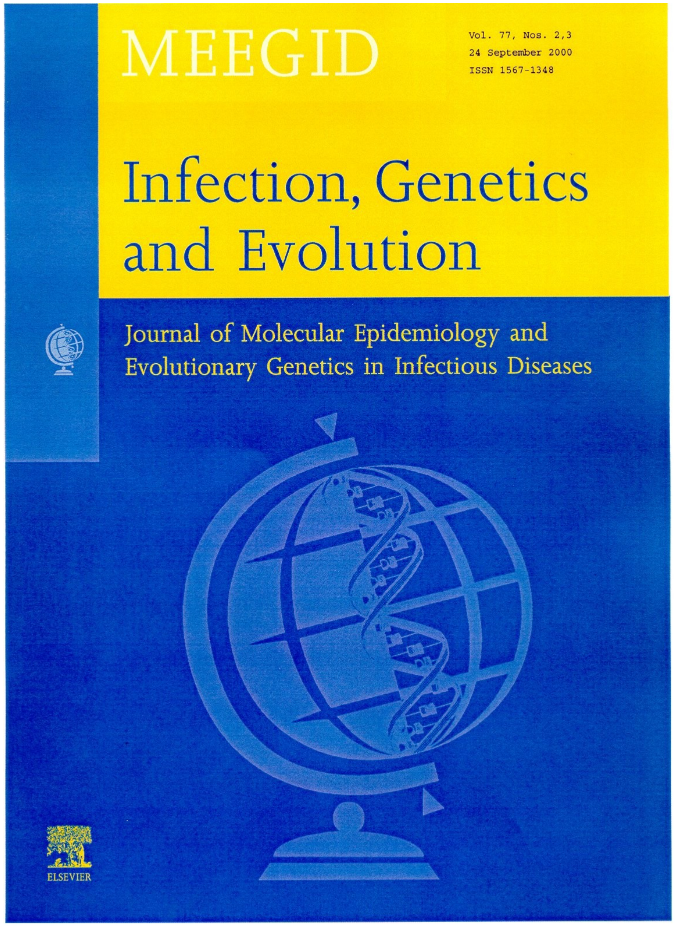
(MEEGID IX)

(coorganized by CDC, USA,
and IRD, Bangkok)

NAIROBI, KENYA, 30th October-1st November 2008

Infection, Genetics and Evolution (Elsevier)

http://www.elsevier.com/locate/meegid



MEEGID

Vol. 77, Nos. 2,3
24 September 2000
ISSN 1567-1348

Infection, Genetics and Evolution

Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases

ELSEVIER

Started 2001

30th issue presently being compiled

6 issues per year

2005 impact factor evaluated by Elsevier: 3.554

Bibliometric data for journals competing with *Infection, Genetics & Evolution*

Journal name	Cites to journal in 2005[a]	2005 Impact Factor	2005 Immediacy Index	Papers published in 2005	2005 Citation half-life	2005 uncited 2 years
Evolution	20,934	4.155	0.597	243	>10.0	8.3%
Infection, Genetics & Evolution	410	3.554	0.738	82		20.2%
Journal of Clinical Microbiology	38,704	3.537	0.498	1101	6.1	16.7%
International Journal for Parasitology	5,343	3.346	0.404	151	5.8	18.3%
European journal of human genetics	3,356	3.251	1.083	168	4.1	16.9%
Annals of human genetics	2,164	3.192	0.368	68	7.8	26.3%
Advances in Parasitology	897	3.154	0	9	9.5	6.3%
Development genes and evolution	1,525	2.549	0.632	68	4.8	15.3%
American Journal of Tropical Medicine and Hygiene	10,159	2.482	0.351	359	8.7	23.9%
Acta Tropica	2,276	1.8	0.817	131	5.3	36.9%
Parasitology	5,367	1.703	0.326	172	9	30.2%
Journal of Parasitology	5,878	1.524	0.202	272	9.3	43.3%
Journal of Medical Entomology	3,402	1.489	0.19	158	8.7	35.4%
Journal of Eukaryotic Microbiology	1,782	1.447	0.152	66	6	41.0%
Experimental Parasitology	2,877	1.306	0.225	138	>10.0	34.1%

Source: 2005 Journal Citation Reports, ScienceEdition ©ThomsonScientific, 2006 and data from Web of Science©, as well as other databases.

[a] To all papers ever published.

Genetics and Evolution in the broad sense: includes population genetics and biology, evolution, phylogeny, genomics, proteomics, **bioinformatics**

All infectious models: bacteria, viruses, parasites, fungi. Human, animal, plant diseases

South-South Collaboration

Latin America-Africa-Thailand Initiative for Scientific Partnership

(LAATISP)

Under the auspices of

TICA (Thailand International Cooperation Agency)

IRD

French Ministry of Foreign Afffairs

French Agency for Development

The « World CDC Belt »

Origin: The « Euro-CDC » Proposal

Tibayrenc, M. 1997. European Centres for Disease Control. *Nature* 389: 433-434.

Tibayrenc, M. 1997. Microbes Sans Frontières and the European CDC. *Parasitol. Today* 13 (12): 454.

Tibayrenc, M. 2001. A European centre to respond to threats of bioterrorism and major epidemics. *Bull. WHO* 79: 1094.

Tibayrenc, M. 2004. The European Commission pocket CDC: *encore un effort*!. *The Lancet Infectious Diseases*: 4: 12-13.

Tibayrenc, M. 2005. A hard lesson for Europeans: the Asean CDC. *Trends Microbiol.*13 (6): 266-268.

Tibayrenc, M. 2007. the « World CDC belt (W-CDC): An “utopian” vision for controlling major pandemics and infectious diseases in general. *ESCMID News* 1: 42-43.

- - US CDC
- - European CDC considerably magnified, involving former USSR and Turkey.
- - Asean CDC, involving all countries of this region of the World
- - A Latin American CDC
- - An African CDC +++

All centres should feature:

- Triple mission: advanced research, surveillance and control, training
- Powerful technologies
- Multidisciplinary approaches: advanced technologies, traditional know-hows, field research, human science
- Large, multinational teams

Utopia : what has not happened yet (the author).

Mankind beauty is diversity


AUSTRALIA