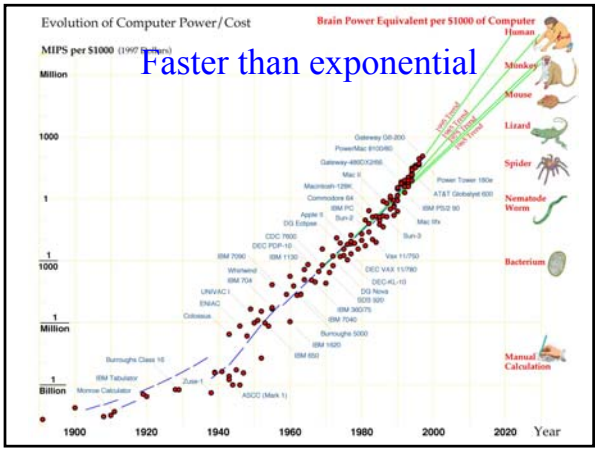
Net2: Bio-algorithms (Last week)

- **Biology to aid Computing … to aid biology**
- **Molecular nano-computing (DNA)**
- **Self-assembly (nano-I/O)**
- **Intra cellular network computing (oscillators)**
- **Genetic algorithms (multi-align)**
- **Neural nets (exons)**

Net3: Global integration

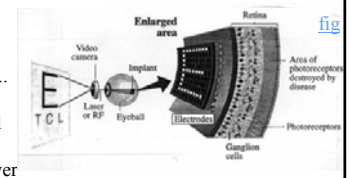
- **Multi-cellular models -- e.g. sensory integration**
- **Systems biology, simulation & integration**
- **Organ systems**
- **Multi-organism - Ecological modeling**
  - **predator/prey - host/parasite - HIV**
- **Global & socioeconomic considerations**
- **Education**
  - **Model evaluation & sharing**



Faster than exponential

The human neural net

"The retina's 10 million detections per second [.02 g] ... a risky extrapolation ... 100 trillion instructions per second to emulate the 1,500 gram human brain. ...Computer power for a given price roughly doubled each year in the 1990s, ... thirty more years at the present pace would close the millionfold gap." (Morovec99)



fig

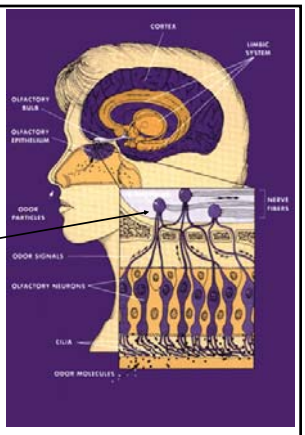
Edge & motion detection (examples)

1000 receptors,
one per cell,
+/-2sd olfactant concentration

$c_{threshold}$

span 6.8 log10 units

Olfactory integration: glomeruli



fig

Basic olfactory tasks

1) Odor memory and recognition.
2) Background elimination (one known + unknown thoroughly mixed)
3) Component separation. (few known odors mixed)
4) Odor separation (turbulent unknowns)

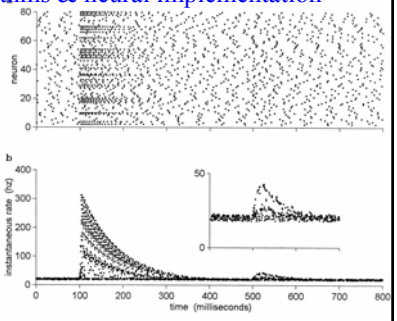
$cov_i = cov_{min} (c_t/c_{thresh,t})$ (1 or f_{it}) + $cov_{min} (c_b/c_{thresh,b})$ (1 or f_{ib})

$c_i = (c_{thresh} / f_{it})(cov_{iu}/cov_{min})$

t= target , b=background, u= unknown, receptors i=1 to 1000.
minimum coverage for concentration threshold.
f_{it} = fraction bound random 1 to 10^{-6} (log uniform pdf)

Hopfield 1999; PNAS 96:12506-11

Odor space and olfactory processing: Collective algorithms & neural implementation



80 adapting neurons, two sniffs: 100-500 msec has a mixed odor 50*x + 1000*y. At 500 msec 75*x + 1100*y. The sniff at 100 milliseconds strongly activates more than half the neurons, after which they adapt. The changed sniff at 500 milliseconds is almost invisible.
b) as in a), but the y-axis = firing rate at the time of each action potential. The second sniff is now clearly visible, and most spikes appear to belong to one of three patterns. A 20% spread in D was included to produce parameter-spread noise.

Hopfield 1999; PNAS 96:12506-11 & MATLAB

Olfaction code

```
% figlshare.m is a matlab script for one odor
% its two graphics panels are like Fig 1b and 1e in the paper
% by adjusting strength (below) other parts of Fig 1 can be recovered
% span of binding constants 10^6
% N odor receptor types
clear
N=2000
strength=.1 % for saturation, set strength=1
% strength=.003 is 3 times threshold
clf;

logtarget = 6*rand(N,1)-6; %the target at saturating concentration
target = exp(2.3*logtarget); % the signal due to the target
% when present at saturating strength

logbackground = 6*rand(N,1)-6; % the logbackground at saturating conc.
background = exp(2.3*logbackground); % actual signal from saturating
bkgnd.

y=log10(target*strength)+0.1*randn(2000,1); % log10 of signal due to
% target at (strength) + plus noise
```


Systems biology model sharing

Simulators	Director	Inst.	Features
ERATO,j	John Doyle	Caltech	SysBiolWorkbench&SBML
Gepasi,w	Pedro Mendes	Santa Fe	MCA, systems kinetics
JarnacScamp,wx	Herbert Sauro	Caltech	MCA, Stochastic
StochSim,w+	Dennis Bray	U.Camb.	Stochastic
BioSpice,u	Adam Arkin	LBL	Stochastic
DBSolve,w	Igor Goryanin	Glaxo	enzyme/receptor-ligand
E-Cell,u+	Masaru Tomita	Keio	metab. Net ODE
Vcell,j	Jim Schaff	U.CT	geometry
Xsim,u	J.Bassingthwaighte	Seattle	enzymes to body physiology
CellML,x+	Peter Hunter	U.Auckland	geometry, model sharing
GENESIS,u	James Bower	Caltech	neural networks
Simex,u+	Lael Gatewood	U.MN	Stochastic micropopulation disease spread

J=java, w = windows, u=unix, x=XML, + = source/community input

Net3: Global integration

- **Multi-cellular models -- e.g. sensory integration**
- **Systems biology, simulation & integration**
- **Organ systems**
- **Multi-organism - Ecological modeling**
  - **predator/prey - host/parasite - HIV**
- **Global & socioeconomic considerations**
- **Education**
  - **Model evaluation & sharing**

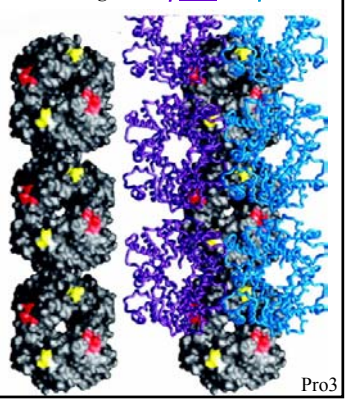
Design and strategy for the Cardiome Project

Adv Exp Med Biol 1997;430:325-39 Bassingthwaighte JB

The Physiome Project has the goal of providing the quantitative description of the integrated functions of the living organism ... a central scheme, a description of the spread of excitation and contraction through an anatomically detailed cardiac model with fiber directions. This will be augmented by the additions of regional blood flows, substrate uptake and metabolism, and energy production and utilization in serving contraction and ionic balances. Later stages will involve cellular regulation and responses to interventions. The organization of such projects is by the assembling of components whose linkages one to another are first minimized and then augmented to improve the approximation to reality.

(url)

Modeling SNP to 3D to cell morphology & pathology



Harrington et al. Crystal structure of deoxy-human hemoglobin β6 Glu --> Trp. Implications for the structure and formation of the sickle cell fiber. J Biol Chem. 1998, 273:32690-6. (Pub)

Pro3

Red Blood Cell Function

- Transport O_2 from lungs to tissues – using hemoglobin to carry the O_2
- Hemoglobin is maintained in its functional state (reduced) by the metabolic machinery
- Cell membrane separates the internal environment from the external environment
  - subject to physicochemical constraints
    - Electroneutrality, Osmotic balance
  - Cause of the imbalance:
    - impermeable polyions inside the cell
      - hemoglobin, organic phosphates

3D-structure to function to morphology to function

Dacrocytes; pernicious anemia

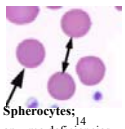
Echinocytes *(crenated);* hyperosmotic medium

Macrocytes 9-12µm; megaloblastic anemia

Acanthocytes; abetalipoproteinemia

Normal RBC 6-8 µm; 4e12 per L

Sickle cell; HbS



Spherocytes; enzyme deficiencies

www.wadsworth.org/chemheme/heme/cytocritique

From SNPs to pathogen resistance mechanisms

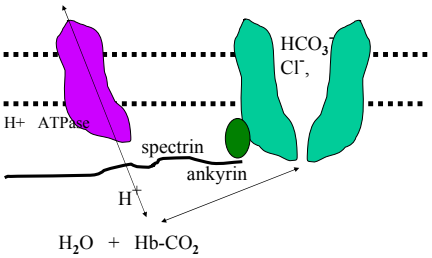
"Model...erythrocytes of [glutathione peroxidase] GPX1*2 heterozygotes should be more efficient in sheltering the cell membrane from irreversible oxidation and binding of hemoglobin caused by the oxidant stress exerted by Plasmodium falciparum... we observed a clear trend toward a dissociation between the HBB*A/*S and GPX1*2/*1 genotypes in the overall data."
Destro-Bisol et al. Hum Biol 1999; 71:315-32. (Pub)

From SNPs to toxicology mechanisms

"Drug induced [e.g. primaquine] oxidative hemolysis ... with certain enzymopathies, notably glucose 6 phosphate dehydrogenase deficiency... Others ... disrupt mitochondrial function and ... heme biosynthesis ... including ... alcohol & chloramphenicol"
Ammus & Yunis, Blood Rev 1989;3:71 82, Drug induced red cell dyscrasias. (Pub)

A possible molecular mechanism governing human erythrocyte shape.

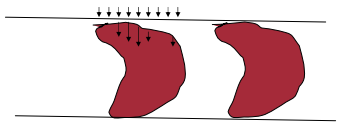
Band3 = 10% RBC membrane



Gimsa *Biophys J* 1998 Jul;75(1):568-9 (Pub)

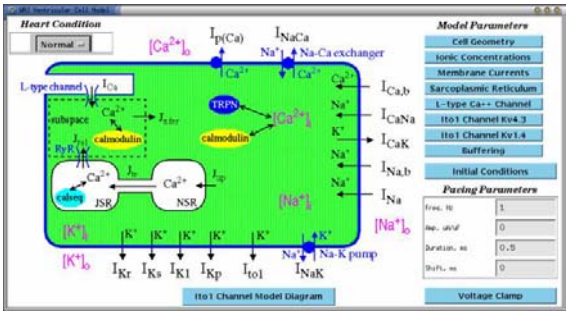
Red cell distortion and conceptual basis of diffusing capacity estimates: finite element analysis

Compute the uptake of CO across a two-dimensional geometric capillary model containing a variable number of equally spaced RBCs (circular or parachute shaped, with the same perimeter length). Total CO diffusing capacity (DLCO) and membrane diffusing capacity (DMCO) were calculated by a finite element method. DLCO calculated at two levels of alveolar PO_2 were used to estimate DMCO by the Roughton-Forster (RF) technique or by the morphometric random linear intercept method. Results: shape distortion of RBCs reduces diffusive gas uptake, & exaggerates errors in the RF technique at a high capillary hematocrit. Shape distortion error in morphometric estimates of DMCO are exaggerated at a low hematocrit.



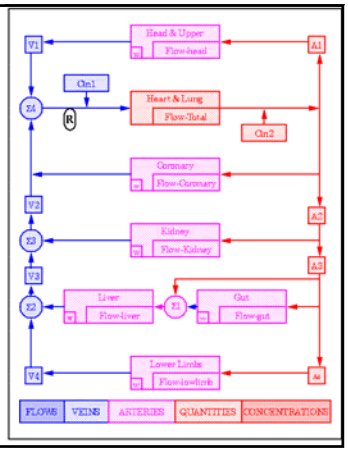
Adapted from Hsia CCW, et al. *J Appl Physiol* 1997 83:1397-404..

Action Potential Model for Canine Ventricular Cell



Greenstein et al. Role of the Calcium-Independent Transient Outward Current I_{to1} in Shaping Action Potential Morphology & Duration. Circ Res. 87:1026-1033, 2000

Whole body recirculation model(NSR) with four chambered heart, seven organs, four arteries and four veins



Multi-Organ System Failure

Is a, life-threatening complication of otherwise mild sickle cell disease ref

(as well as of injury in non-sickle cell individuals).

The system model is largely unknown. Will genomic data help prediction and/or prevention?

project

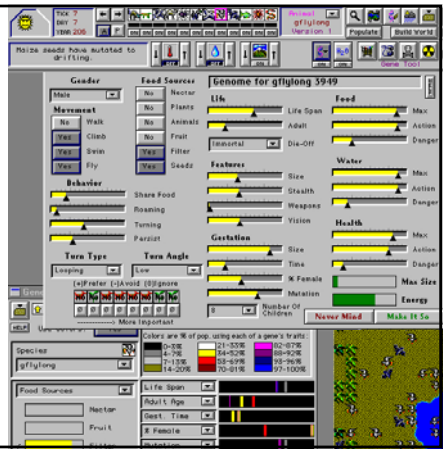
Net3: Global integration

- **Multi-cellular models -- e.g. sensory integration**
- **Systems biology, simulation & integration**
- **Organ systems**
- **Multi-organism - Ecological modeling**
  - **predator/prey - host/parasite - HIV**
- **Global & socioeconomic considerations**
- **Education**
  - **Model evaluation & sharing**

The Sims:

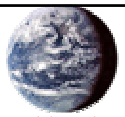
Sim City'87
Sim Earth'88
Sim Ant'91
Sim Life'92
Sim Farm'93
Sim Isle'95
…

Simulation software list



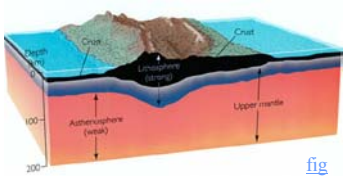
Integration of multiple inputs: Think globally; act locally:

Photomosaic , AIDS quilt Robert Silvers



Think globally; act locally

Lithosphere (0.2% C, 75% SiO_2) 110 C at 4 km
Diameter = 1.3e6 m = 5e22 g (5000 species / g soil)
Biosphere 3e15 g (dry wt. marine); 2e18 g (land)
Microbial hydrosphere 1.4e21 ml = 1e27 cells = 4e33 bp
Anthrosphere (23% C) = 6e23 cells = 4e32 bp.



fig

Biodiversity for nanoengineering

Tetracycline

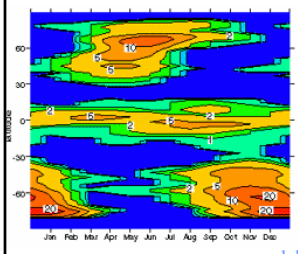
Coronene 

Protein Eng 2000 Feb;13(2):121-8 Protein engineering of cytochrome 8p450(cam) (CYP101) for the oxidation of polycyclic aromatic hydrocarbons.

Lett Appl Microbiol 2000 May;30(5):396-401 Microbial degradation & detoxification of high molecular weight polycyclic aromatic hydrocarbons by Stenotrophomonas maltophilia strain VUN 10,003.

J Am Chem Soc 2002 Oct 30;124(43):12664-5. Bioelectrochemical single-walled carbon nanotubes.

Modeling bio-effects on global warming



models

The equatorial Pacific, sub arctic pacific and Southern Ocean are high-nutrient low-chlorophyll (HNLC) areas which may support higher plant biomass if micro-nutrients such as Fe were added… No ocean fertilization study has been long lived enough to follow the effects of iron fertilization through the **food web,** and hence determine the **potential for long term carbon sequestration**.

<1% of photosynthetic biomass, phytoplankton ~50% carbon fixation.

Chisholm et al. (2001) Science 294(5541):309-1. Discrediting ocean fertilization.

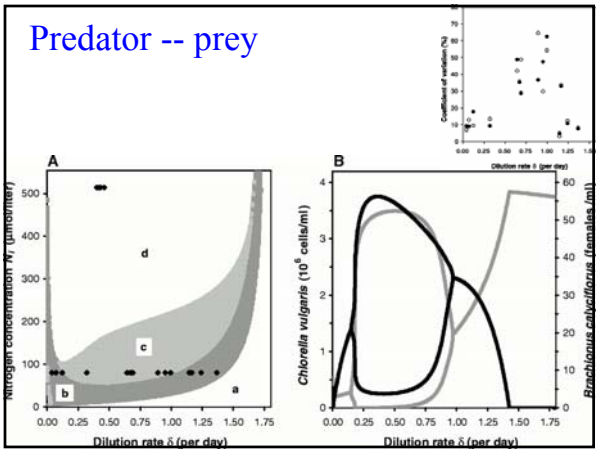
Crossing the Hopf bifurcation in a live predator-prey system.

$$dN/dt = \delta(N_i - N) - F_C(N)C$$
$$dC/dt = F_C(N)C - F_B(C)B/\varepsilon - \delta C$$
$$dR/dt = F_B(C)R - (\delta + m + \lambda)R$$
$$dB/dt = F_B(C)R - (\delta + m)B$$

$F_C(N) = b_C/(1+ K_C/N)$
$F_B(C) = b_B/(1 + K_B/C)$
N = concentration of nitrogen
ε = assimilation efficiency
δ = dilution rate
C = concentration of Chlorella
R = concentration of reproducing Brachionus
B = concentration of total Brachionus
m = demographic mortality of Brachionus
λ = decay of fecundity of Brachionus
b = maximum birth rates
K = half-saturation constants

Fussmann et al. Science 2000; 290:1358-60. (pub)

Predator -- prey



Human microbial ecology: Sorting out cause, effect & prescription (vaccine paradox)

Helicobacter pylori	**Stomach cancer**
	Stomach ulcers
Cytomegalovirus	**Coronary artery disease**
Porphyromonas gingivalis	**Alzheimer's**
Chlamydia pneumoniae	**Cervical cancer**
Human papilloma virus	**Liver cancer**
Hepatitis B, C virus	**Breast cancer**
EpsteinBarr Virus	**Nasopharyngeal carcinoma**

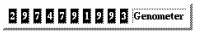
Weber, et al (2002) Microbial sequence identification by computational subtraction of the human transcriptome. Nature Genetics 30:141

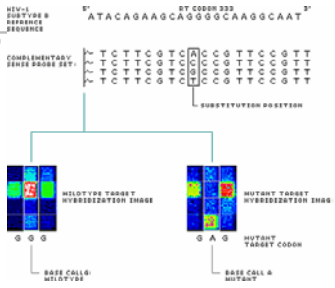
= www urls

HIV-1 resequencing

New nucleotide sequences processed in GenBank per month (above)

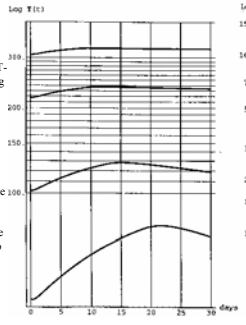
Today's total is:





HIV-AIDS models

Therapy T(0),Vs(0) in mm-3 =
306, 21 (5.8 years),
217, 31/mm3 (7.7 years),
100, 69/mm3 (8.4 years),
43, 156/mm3 (8.6 years).
The rates of exponential increase in Fig. 2a (.03, .02, .01, .005) are inversely correlated to starting CD4+ T-cell counts; decay rates in Fig 2b (all -.2) are not correlated to starting viral levels (different viral set-points would give different values for the parallel slopes) (1,2). The lack of correlation of viral decay rates is an indication of slower clearance of wild-type virus in the external lymphoid compartment. The time to the downward spike in Figure 2b is correlated to starting viral levels (1). The treatment parameters c1=2.0, c2=.17, c3=.15 and the resistance mutation parameter q=10-6 are the same in all four simulations.



D. Kirschner & G Webb Resistance, Remission, & Qualitative Differences in HIV Chemotherapy (Pub)

32

HIV economic modeling

HIV treatement model parameters

μ_T = mortality rate of uninfected CD4+ T cells	0.005/day
μ_{T1} = mortality rate of infected CD4 + T cells	0.25/day
k_s = rate CD4+ T cells are infected by sensitive virus	0.0005 mm^3/day
k_r = rate CD4+ T cells are infected by resistant virus	0.0005 mm^3/day
k_w = rate of virus loss due to the immune response	0.0062 mm^3/day
p_1 = production rate of uninfected CD4+ T cells	0.025/day
p_2 = production rate of infected CD4+ T cells	0.25/day
p_3 = production rate of virus in the blood	0.8/day
G_s = external lymphoid sensitive virus source constant	41.2/mm^3 day
G_r = external lymphoid resistant virus source constant	specified in text
V_0 = threshold value for remission	specified in figure legends
q = proportion of drug-resistant virus produced from wild type virus	specified in figure legends
C = half saturation constant of uninfected CD4+ T cells	47.0/mm^3
C_1 = half saturation constant of infected CD4+ T cells	47.0/mm^3
B = half saturation constant of external virus input	2.0/mm^3
B_1 = half saturation constant of CD4+ T-cell source	13.8/mm^3
S_1 = source of CD4+ T cells in absence of the disease	4.0/mm^3 day
S_2 = reduction constant of CD4+ T-cell source	2.8/mm^3 day
c_1 = treatment parameter for suppression of the rate of CD4+ T-cell infection by virus	specified in figure legends
c_2 = treatment parameter for suppression of the rate of virus contributed by the external lymphoid compartment	specified in figure legends
c_3 = treatment parameter for maximal suppression of virus contributed by the external lymphoid compartment	specified in figure legends
η_1 = treatment function for inhibition of the rate at which virus infects uninfected CD4+ T cells	specified in text
η_2 = treatment function for inhibition of the rate of virus influx from the external lymphoid system virus	specified in text

Vaccines for the 21st Century

http://books.nap.edu/html/vacc21/

Level I Most favorable: saves money & Quality-Adjusted Life Years(QALY)
Level II < $10,000 < Level III < $100K per QALY saved < Level IV

Level I candidate vaccines:

- Viral: CMV vaccine for 12 year olds, Flu vaccine for 20% of the US per year.
- Therapeutic vaccines: IDDM diabetes, MS, Rheumatoid arthritis
- Bacterial: Streptococcus B & pneumoniae vaccine for infants & 65 year olds.
- [HIV vaccines prominent already within NIH.]

"A **quantitative model** that could be used by decision makers to prioritize the development of vaccines against a number of disparate diseases" 1985 & 1999.

34

Role of genomics & computational biology in vaccine R&D?

DNA vaccines , Intracellular vaccines, RNAi, multiplexed…

Gaschen et al. (2002) Science 296(5577):2354-60 Diversity considerations in HIV-1 vaccine selection.

Malaria & Mosquito genomes

35

Net3: Global integration

- **Multi-cellular models -- e.g. sensory integration**
- **Systems biology, simulation & integration**
- **Organ systems**
- **Multi-organism - Ecological modeling**
  - **predator/prey - host/parasite - HIV**
- **Global & socioeconomic considerations**
- **Education**
  - **Model evaluation & sharing**

36

Human Genome Project Ethical, Legal & Social Issues (ELSI)

Fairness - Genetic non-discrimination
Privacy
Reproductive rights - cloning
Psychological stigmatization
Clinical quality-control
Safety and environmental issues - GMO & biowarfare
Uncertainties - testing minors
Conceptual & philosophical implications - diversity?
Commercialization of products - Who owns?

Human Genome Project Ethical, Legal & Social Issues (ELSI)

Clinton, June 26, 2000. The Genetic Nondiscrimination in Health Insurance and Employment Act of 1999, introduced by Senator Daschle and Congresswoman Slaughter, that will extend these employment protections to the private sector and finish the job of helping to extend protections to individuals purchasing health insurance, begun with the Health Insurance Portability and Accountability Act.

ELSI: Do Races Differ? Not Really, DNA Shows

Hb variants "evolved to help the ancestors of these groups resist malaria infection, but both prove lethal when inherited in a double dose. As with differences in skin pigmentation, the pressure of the environment to develop a group-wide trait was powerful, and the means to do so simple and straightforward, through the alteration of a single gene.

A founder effect explains the high incidence of Huntington's neurodegenerative disease in the Lake Maracaibo region of Venezuela, and of Tay-Sachs disease among Ashkenazi Jews.

But Dr. Naggert emphasized that medical geneticists had a much better chance of unearthing these founder effects by scrutinizing small, isolated and well-defined populations, like the northern Finns, the Basques of Spain, or the Amish of Pennsylvania, than they did by going after "races."

Dangers of model-free science

Lysenko: inheritance of acquired characteristics.
"His habit was to report only successes. His results were based on extremely small samples, inaccurate records, and the almost total absence of control groups. An early mistake in calculation, which caused comment among other specialists, made him extremely negative toward the use of mathematics in science. "

Dangers of ethics-free science

The 1979 release of Anthrax-836 spores in Sverdlovsk.
"In 1953 a leak…In 1956, Sizov found that one of the rodents captured in the Kirov sewers had developed a new strain more virulent than the original. The army immediately ordered him to cultivate the new strain…to install in the SS-18s targeted on western cities."

Alibek & Handelman "Biohazard" 1999 (Davis)

How can we improve our genome engineering tools preferentially toward defense and away from terrorism?

Genetically modified organisms (GMO)

Genetically modified organisms

Developing world needs:
Agri- vaccines, salt & drought tolerance

Terminators: Allergen dispersal vs reseeding

"Organic": no inorganic fertilizers means high animal load.

Many natural pesticides are carcinogens including estragole (basil), safrole (natural root beer), symphytine (comfrey tea), hydrazine (mushrooms) & allyl isothiocyanate (brown mustard); psoralen (celery) & aflatoxin(nuts & cereals).

Cloning & stem cells

Why do clones exhibit developmental defects?
Study epigenetic reprograming with expression profiles?

Can we increase fraction of stem cells without going through cloning?
"Radial glial cells that lacked a functional form of a transcription factor called Pax6 could not generate neurons. But when Pax6 was introduced into glial precursor cells, these cells started to produce neurons." (ref)

Net3: Global integration

- **Multi-cellular models -- e.g. sensory integration**
- **Systems biology, simulation & integration**
- **Organ systems**
- **Multi-organism - Ecological modeling**
  - **predator/prey - host/parasite - HIV**
- **Global & socioeconomic considerations**
- **Education**
  - **Model evaluation & sharing**

Models for education & decision-making

Improve our ability to deal with:

Uncertainty
Complexity
Quantitation
Exceptions (collect and cherish)
Comparisons of diverse entities
Translation & integration
Continuity over time

Measures of quality of structural & functional genomic data

Automate		Data quality	Model quality	Similarity search
X-ray diffraction	1960	resolution < 0.2nm	Σ|o-c|/Σo R < 0.2	DALI
Sequence	1988	discrepancy bp <0.01%	conserved proteins	BLAST
Function	1999	cc, t-test	AlignACE Map & specificity	Correlation

Biophilia & Consilience

Biophilia -- the connections that human beings subconsciously make with other living beings. (Cute animals, snake dreams, therapeutic greenery & natural sounds …)

Consilience - Long-separated fields come together and create new insights; e.g. chemistry & genetics created the powerful new science of molecular biology. Is all human endeavor, from religious feeling to financial markets to fine arts, ripe for explaining by hard science?

How might genomics & computational biology contribute?

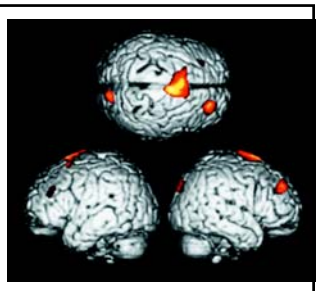
Kellert & Wilson 1993 The Biophilia Hypothesis.
E. O. Wilson 1999 - Consilience: The Unity of Knowledge

PET & MRI

Positron emission tomography

555MBq of ^{15}O butanol , scan for 60s; effective image resolution of 9mm (FWHM) .

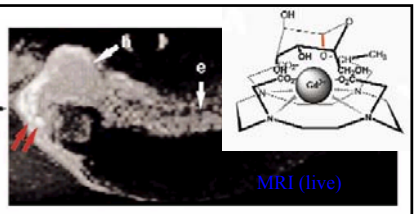
Significant activations for the contrast religious-recite vs. rest in religious subjects, rendered onto canonical T1-weighted image of SPM97d (P <0.001, uncorrected for multiple comparisons) For task comparisons, an ancova (analysis of covariance) model was fitted to the data for each voxel.

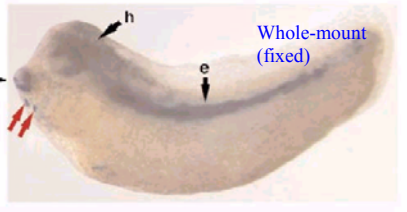


Azari et al. Eur J Neurosci 2001 13(8):1649-52. Neural correlates of religious experience.

MRI & gene expression

Suitable for intact, opaque organisms in 3D at cellular resolution (10 μ)





Louie et al. (2000) Nat Biotechnol 18(3):321-5 In vivo visualization of gene expression using magnetic resonance imaging.

Multi-thanks to

1999:
Jeremy Edwards, Tim Chen, Tao Wei, Bob Freeman, Dan Janse

2000: Jason Comander, Adnan Derti, Bob Freeman, Yonatan Grad, Haley Hieronymus, Dan Janse, Peter Kharchenko, Douglas Selinger, Allegra Petti, Nikos Reppas

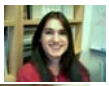
Teaching Fellows 2001

Douglas Selinger

Bob Freeman

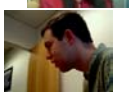
Hui Ge

Haley Hieronymus



Suzanne Komili

Sharyl Wong



Matt Wright

Zhou Zhu

Teaching Fellows 2002

Suzanne Komili

Woodie Zhao

Joon Lee

Lan Zhang

Tom Patterson

Jon Radoff

Yonghong Xiao

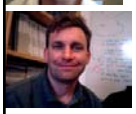
Gary Gao

Laksman Iyer

Teaching Fellows 2003

Lan Zhang

Woodie Zhao



Griffin Weber

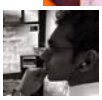
Tom Patterson

J Singh

Xiaoxia Lin

Mike Jones

ChihLong Liu

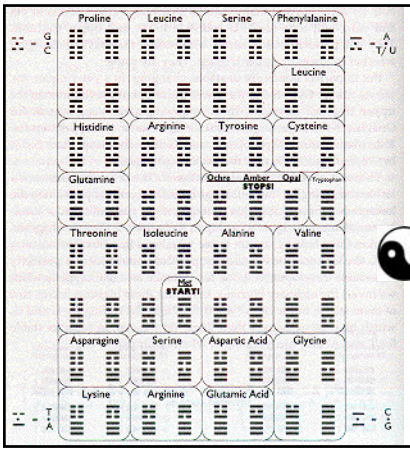


Faisal Reza

101 after '03

1) We will need Teaching Fellows if there is to be a course in Fall 2004. Please contact us.

2) A small number of projects based on need, merit, and interest may be selected for additional support, resources, and/or mentors.



Binary code

Over 5,300 yr BP, China's first emperor, Fu Xi, 64 hexagrams of the I Ching.

See Walter & Smith

Ryan, James (1996) "Leibniz' Binary System & Shao Yong's "Yijing"." Philosophy East & West 46:59-90.