The Next Revolution in Bioscience for Environment and Agriculture?

Dr Tan Tin Wee
Board Director, International Society for Computational Biology (ISCB.org)
Secretariat, S* Life Science Informatics Alliance (s-star.org)
Founding Secretariat, Asia Pacific Bioinformatics Network (APBioNet.org)
ExCo/Past President, Assn for Medical and BioInformatics Singapore (AMBIS.org.sg)
Assoc Professor, Dept of Biochemistry, National University of Singapore (NUS.edu.sg)

Is there a What’s Next?

- The Green Revolution (pre-Molecular Biology days)
- Biotech/recombinant DNA technology Revolution (1980s-1999)
- Genomic era (mid1990s-now)
- “Post-Genomic era” (now)

What's underlying the next revolution?

- Genomics
- Transcriptomics
- Proteomics
- Metabolomics
- Phenomics
- 'omics

The New Language (or "Calculus") of Biology
Computational Biology, BioInformatics
Scientific Computing Information Technology (IT)
Internet Networking (Cyberinfrastructure)

NSF's 21st Century Biology

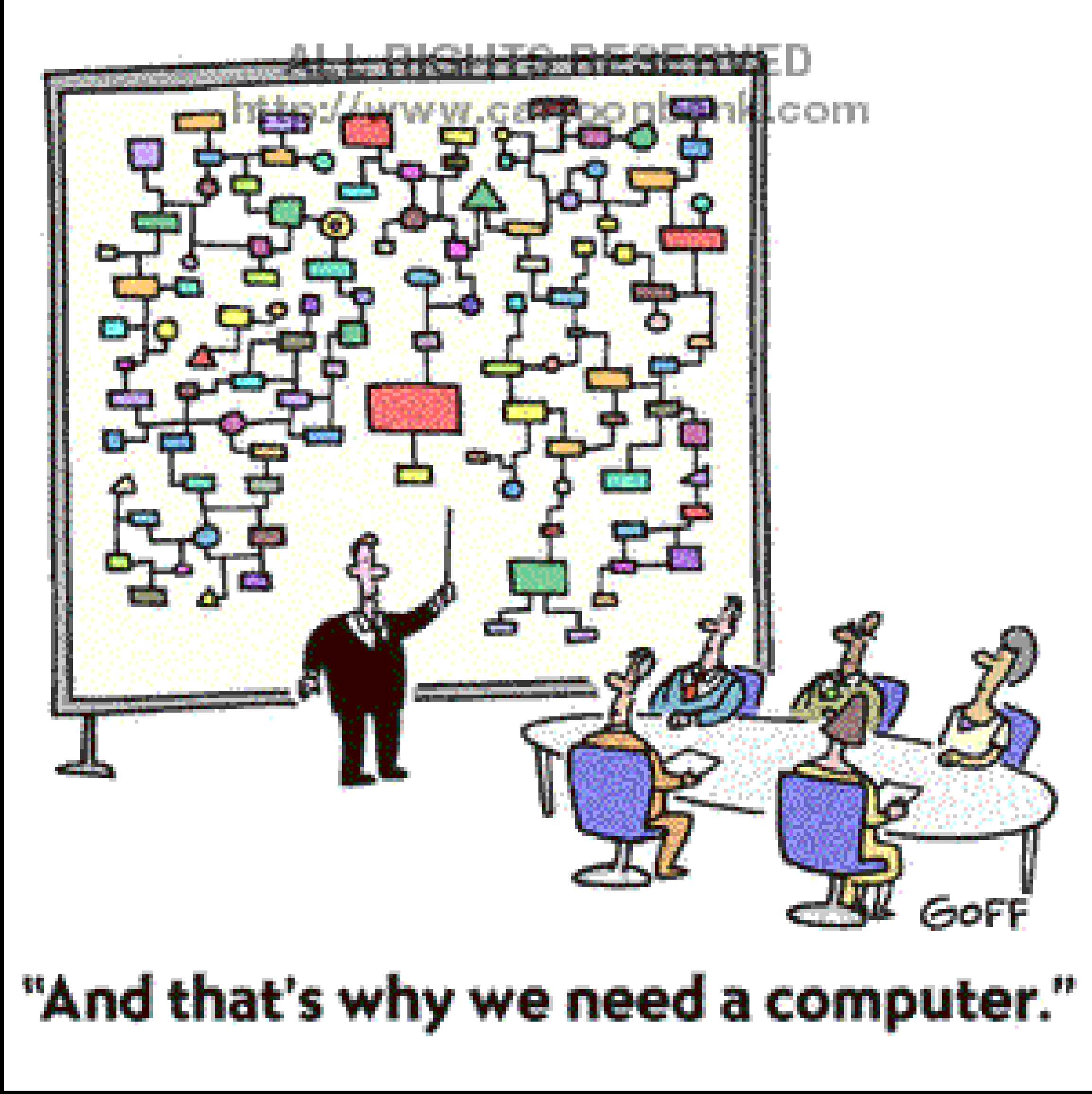
- Integrative
- Synthetic
- Predictive

To build a CIBIO CyberInfrastructure for 21st Century Biology

From NSF 's report:
"Building a Cyberinfrastructure for Biological Sciences" 2005 and Beyond: A roadmap for consolidation and exponentiation

Genbank data is doubling every 12-18months
Faster than Moore's Law of the IT world
After the Post-Genome deluge of data, we now need to analyse data at the Terabyte level

But the New Biology is not just about the volume of data
it is as much about the inherent complexity of biological information



BioImaging –
Ever Increasing Complexity of Biological Data

- linear DNA sequences through to complex movie clips of living cells labelled with quantum dots etc captured in confocal/ fluorescence/electron microscopy and MRI etc:
  - 1D sequence data,
  - 2D image of pixels, microarray data
  - 3D Xray structures and molecular models
  - 4D NMR spectroscopy
  - 6D space-time-spectral records of every voxel in a tomographic slice.

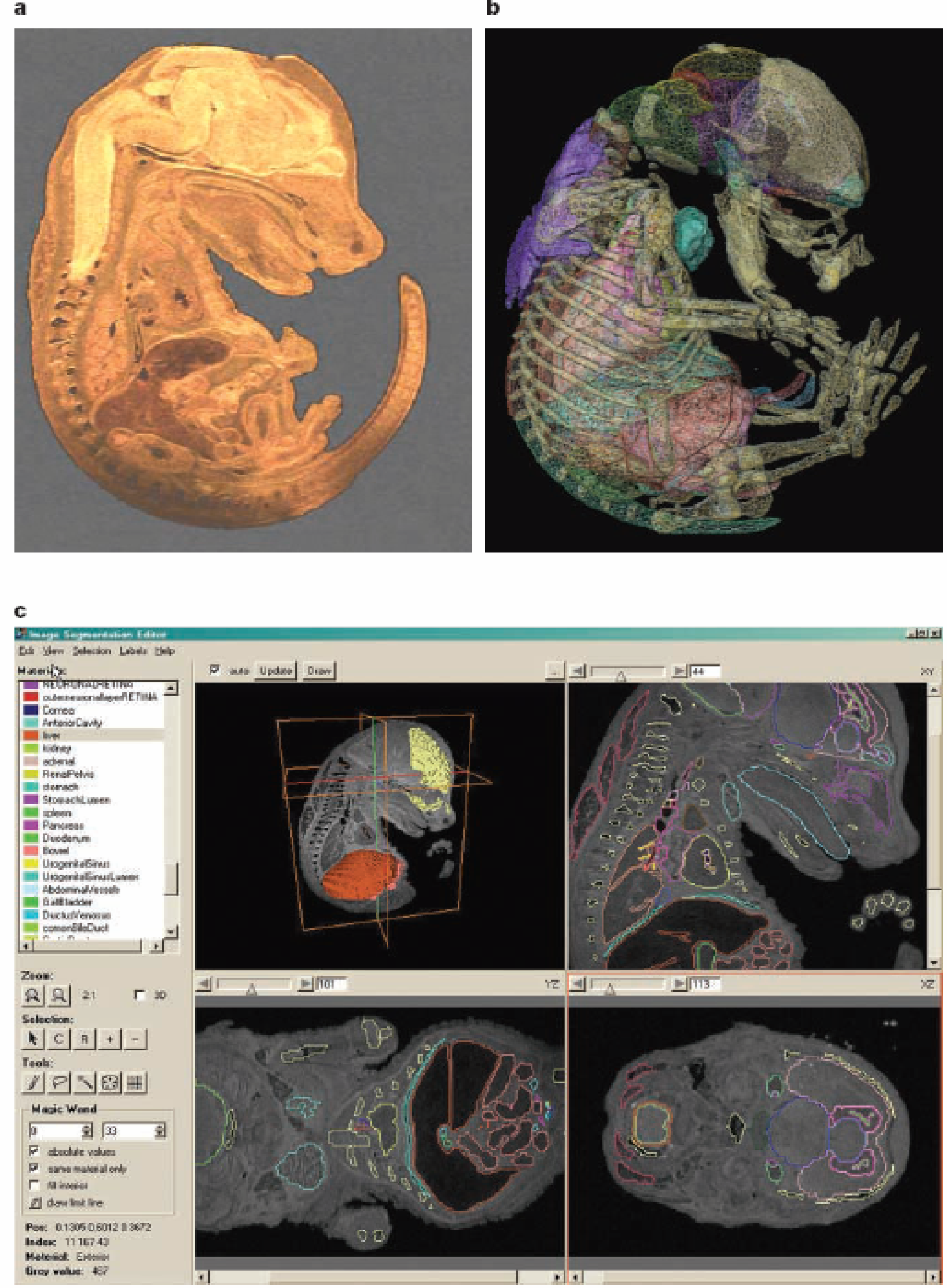


Figure 2 | The creation of volumetric atlases from microscopic magnetic resonance imaging data.
a | This panel shows a single slice of a three-dimensional (3D) magnetic resonance (MR) image of a fixed

Progression of Life Sciences in the Future

Tissue/Organ Physiology

"E-Cell" simulation

COMPUTATIONAL
INFORMATIONAL

Regulatory Networks/Circuits

Pathways

Interactions

Function

Structure

Sequence

Genomics

Transcriptomics

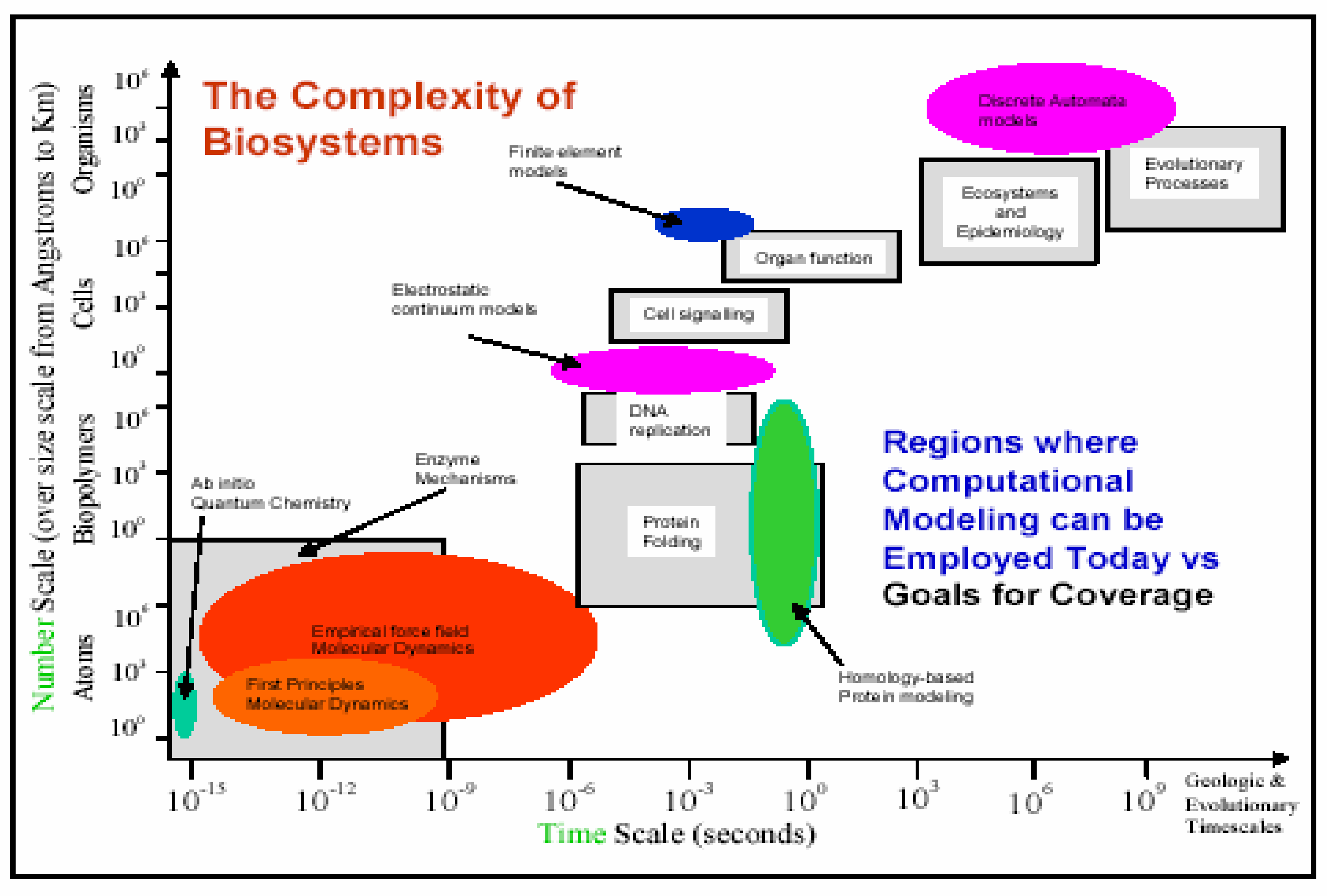
Proteomics

Metabolomics

Other 'omics

BioImaging

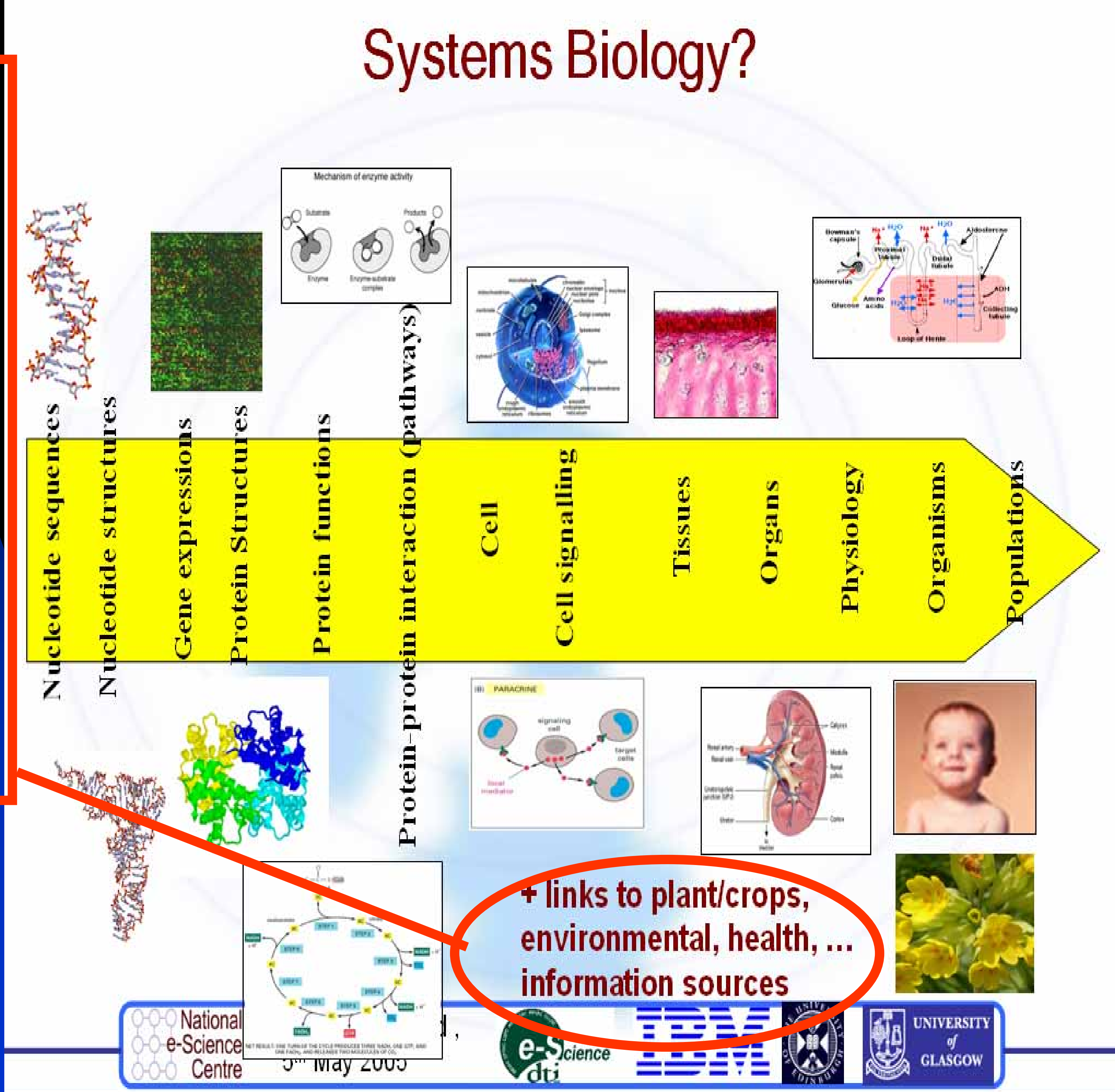
EXPERIMENTAL
OBSERVATIONAL



From John Wooley et al 2005

eScience,
Systems Biology
plant/crops…
Environmental…
All interconnected

Richard Sinnott
Director
UK e-Science
Centre
At the 2nd
Int'l Life Science
Grid Computing
Workshop
LSGrid 2005,
Biopolis, Singapore



2005/6/22

1991
"Towards a Paradigm Shift in Biology"

- The new paradigm, now emerging, is… that the starting point of a biological investigation will be theoretical. An individual scientist will begin with a theoretical conjecture, only then turning to experiment to follow or test that hypothesis. *Walter Gilbert Nature (1991)*

2000
What do some leading scientists think ?

- Biology is in the middle of a major paradigm shift driven by computing.

  Eric Lander

- Computing has changed biology forever; most biologists just don't know it yet.

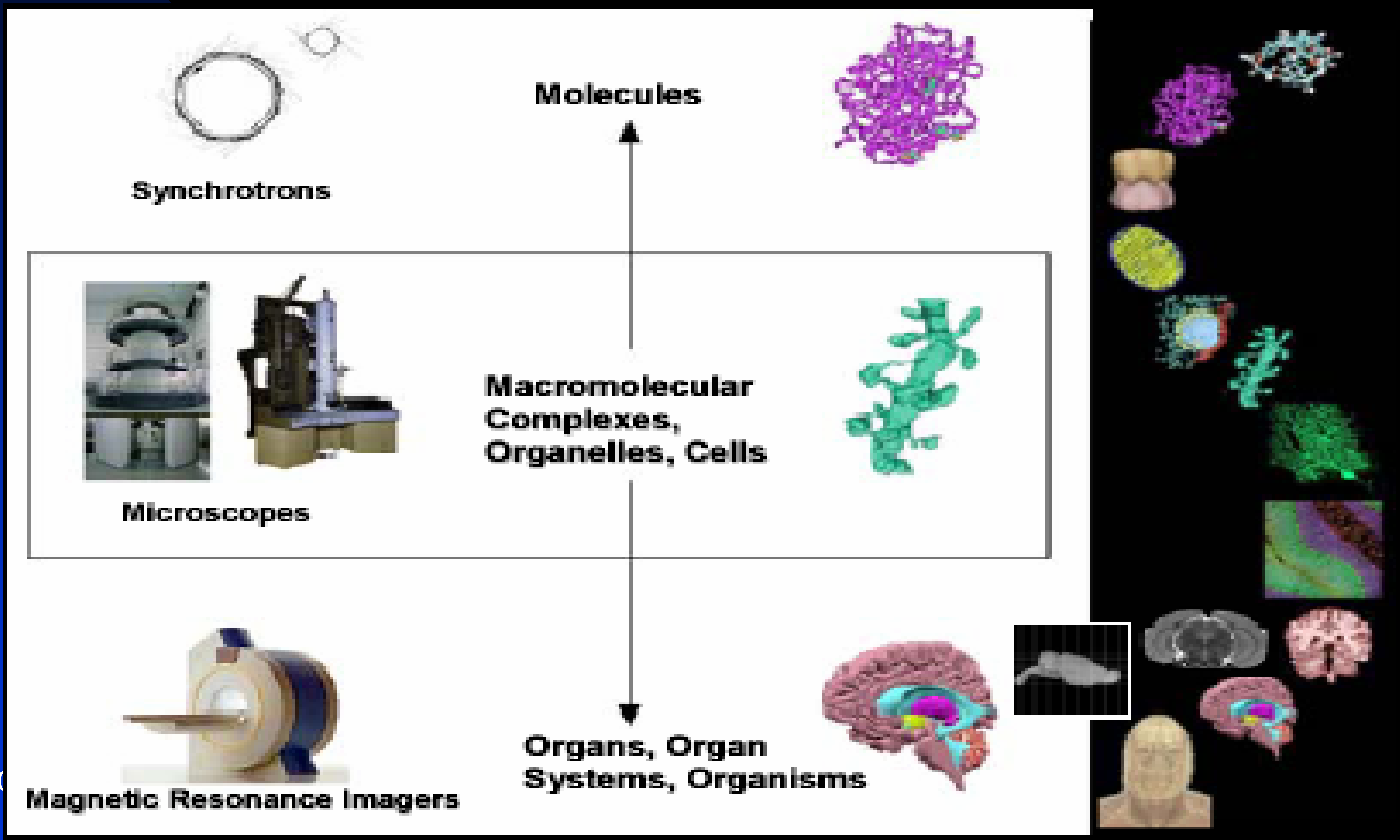
  M Levitt

- Computational Biology will be as essential for the next quarter century of biology as molecular biology was for the past quarter century.

  W McGinnis

The world of scientific computing and advanced IT reached the level of being fully applicable to a wide range of deep biological research themes

John Wooley, 2005



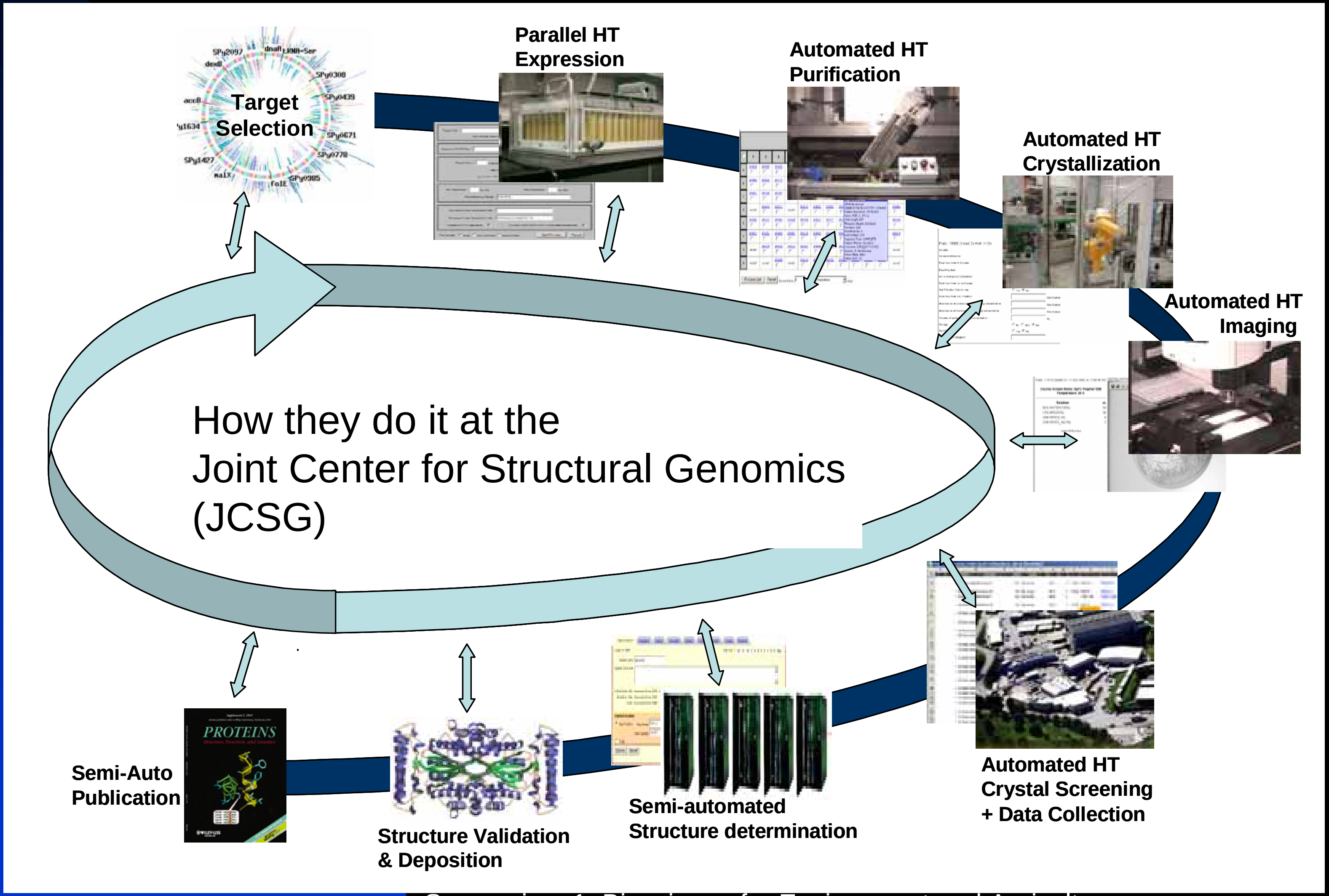
Ecological and Environmental research

- Environmental Sensors for measuring physical, chemical, biological, meteorological, spatial, ecological parameters
- 24by7 monitoring of environmental and population events
- New technologies, methodologies and infrastructure for the environmental sciences

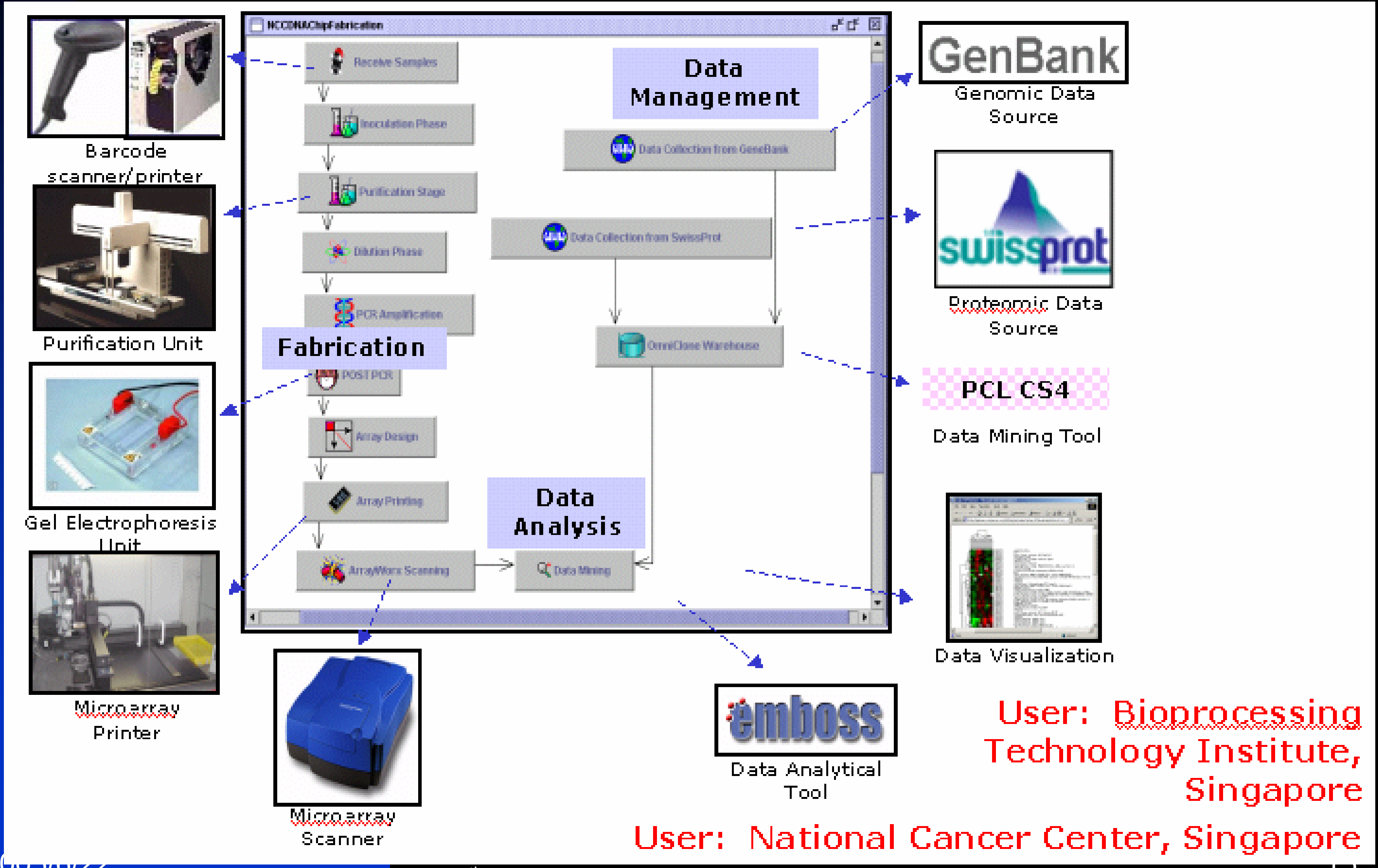
From
NSF Report CIBIO
2005

- Genes responsible for domestication of crops
- Mechanisms of polyploidization vs genome reduction
- Molecular mechanisms of symbiosis

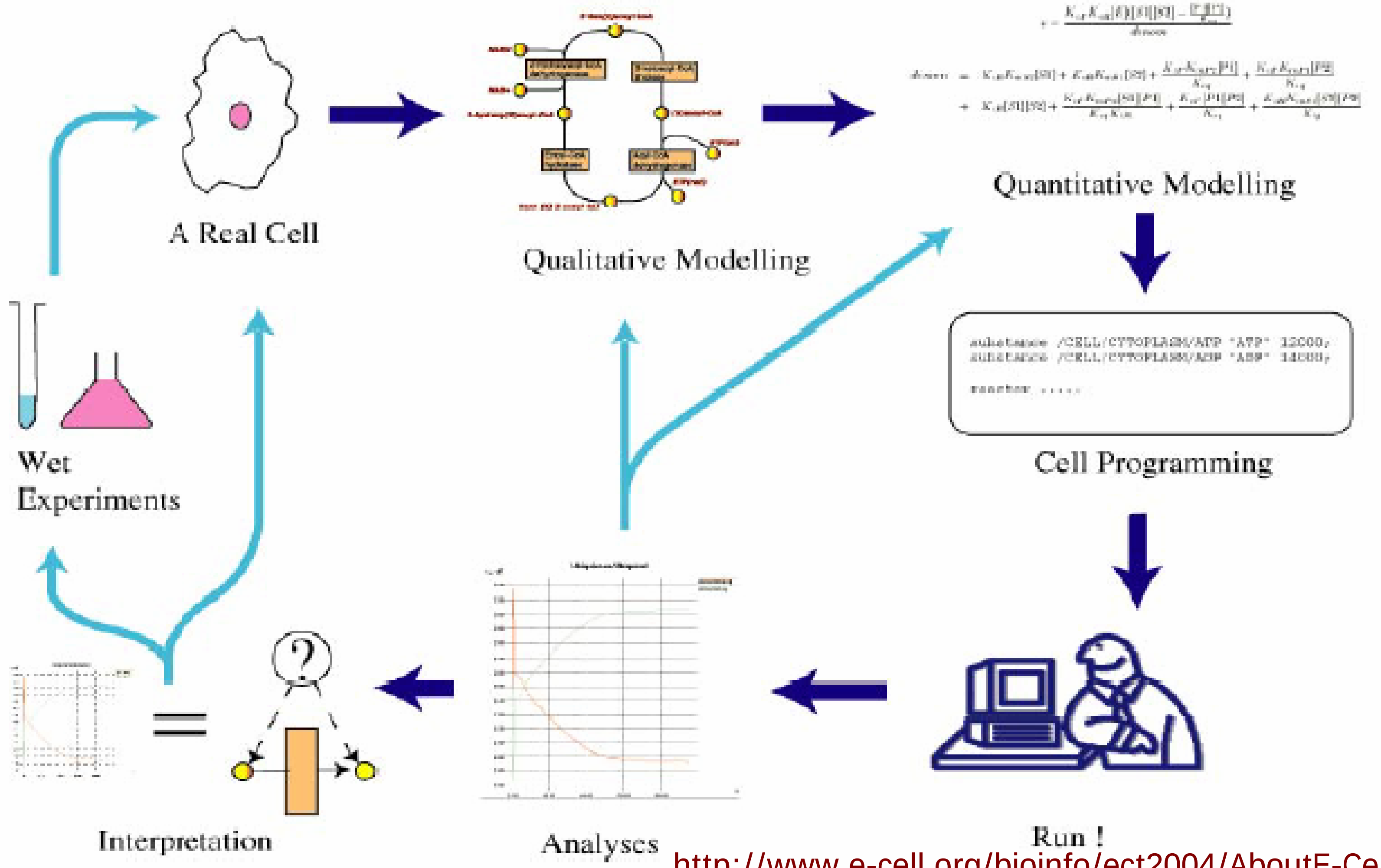
Workflow Integration for High Throughput research and Pipeline data handling architecture.



Workflow Integration in Laboratory Automation



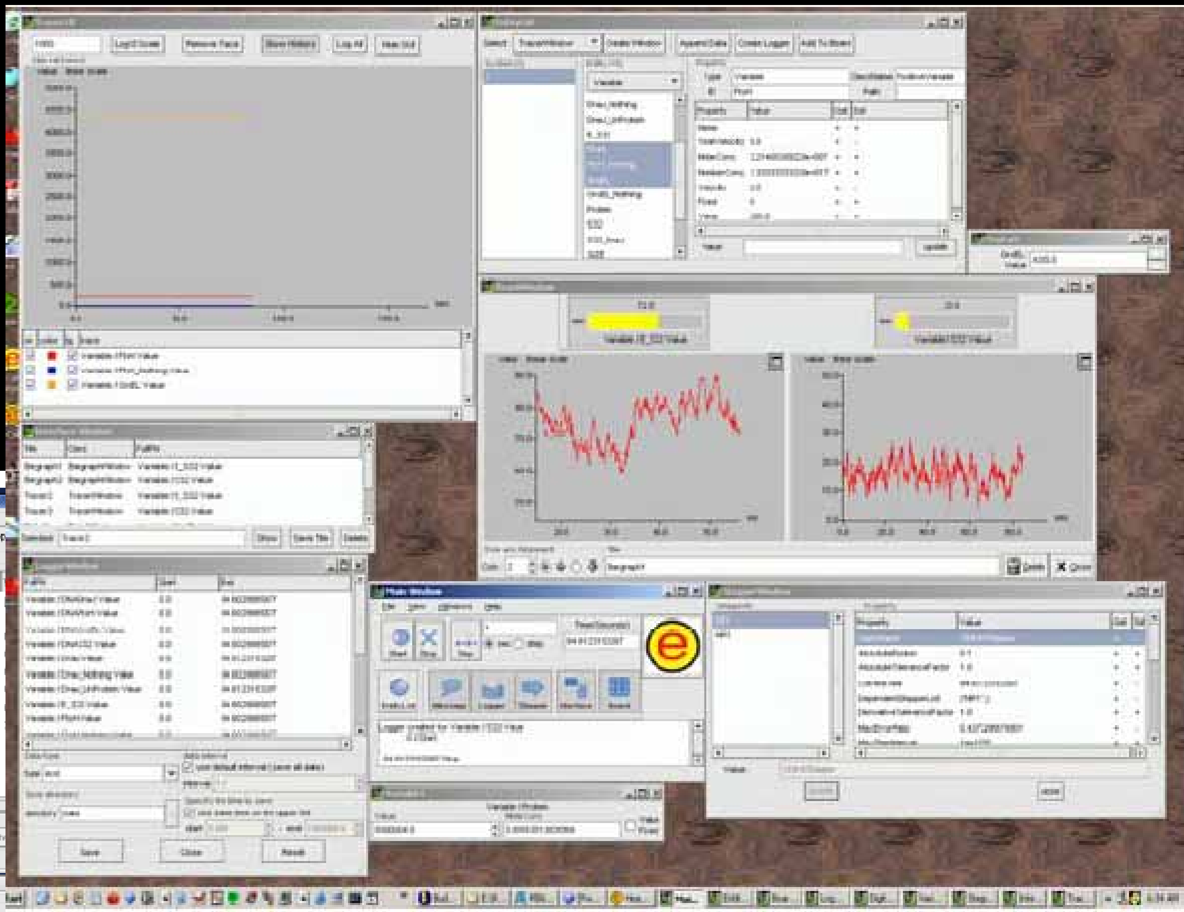
Cycle of the biological study with *in silico*

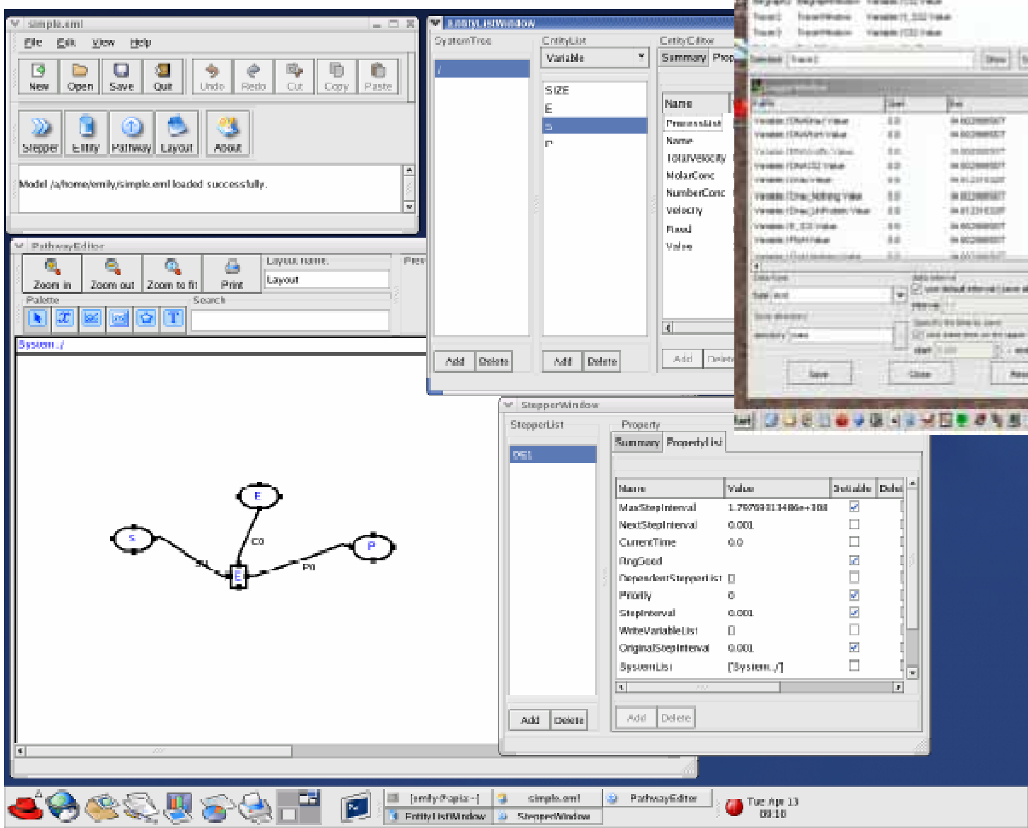


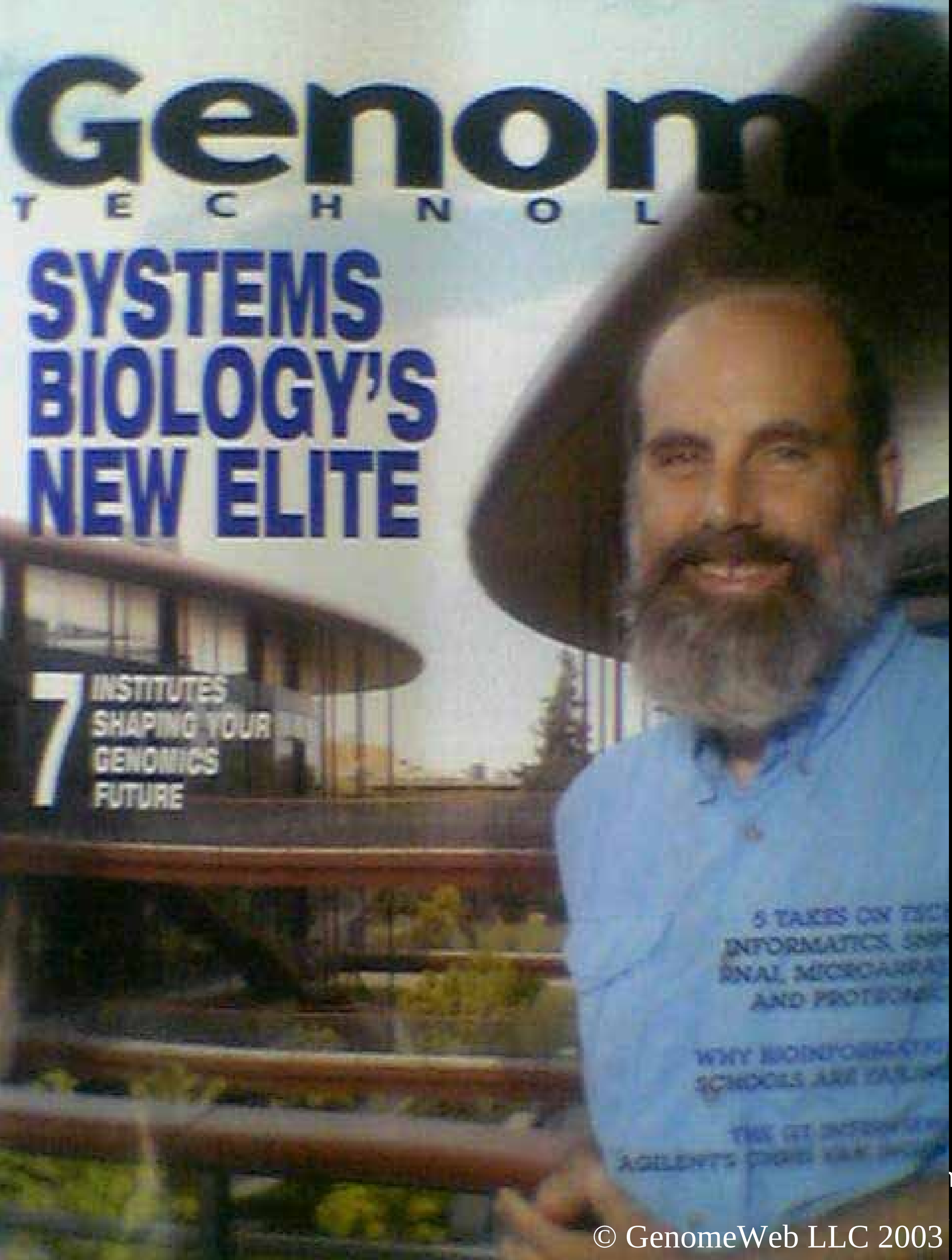
http://www.e-cell.org/bioinfo/ect2004/AboutE-Cell.pdf

E-cell

- Model builder
- Algorithm modules
- Simulation visualiser







- Stanford BioX (US$ 150M)
- MIT CSBi (US$10M/yr)
- Princeton Sigler Inst for Integrative Genomics ICAHN Lab (US$40M)
- Duke Institute for Genome Sciences and Craig Venter's TCAG (US$250M)
- UMichigan LSI (US$380M)
- QB3 UCaliforniaSF /Scruz/Berkeley (US$200M)
- Cornell LSI (US$140)
- UCSD JCSG
- NSF new CIBIO programme

Biology is Big Science these days

After the Genomes projects, industrial scale generation of data is no big deal.

Sophisticated bioinstrumentation from automated sequencers to microarray systems 24by7 churn out ever increasingly large scales of output, throughput and data generation.

Grid Computing tools are now available for supercomputing scale computing accessible from your desktop

APBioGrid 2002:

Asia Pacific’s role in Grid Computing for Life Science

The weekly GlobAl bioinformATiCs news serviCe

bioinform

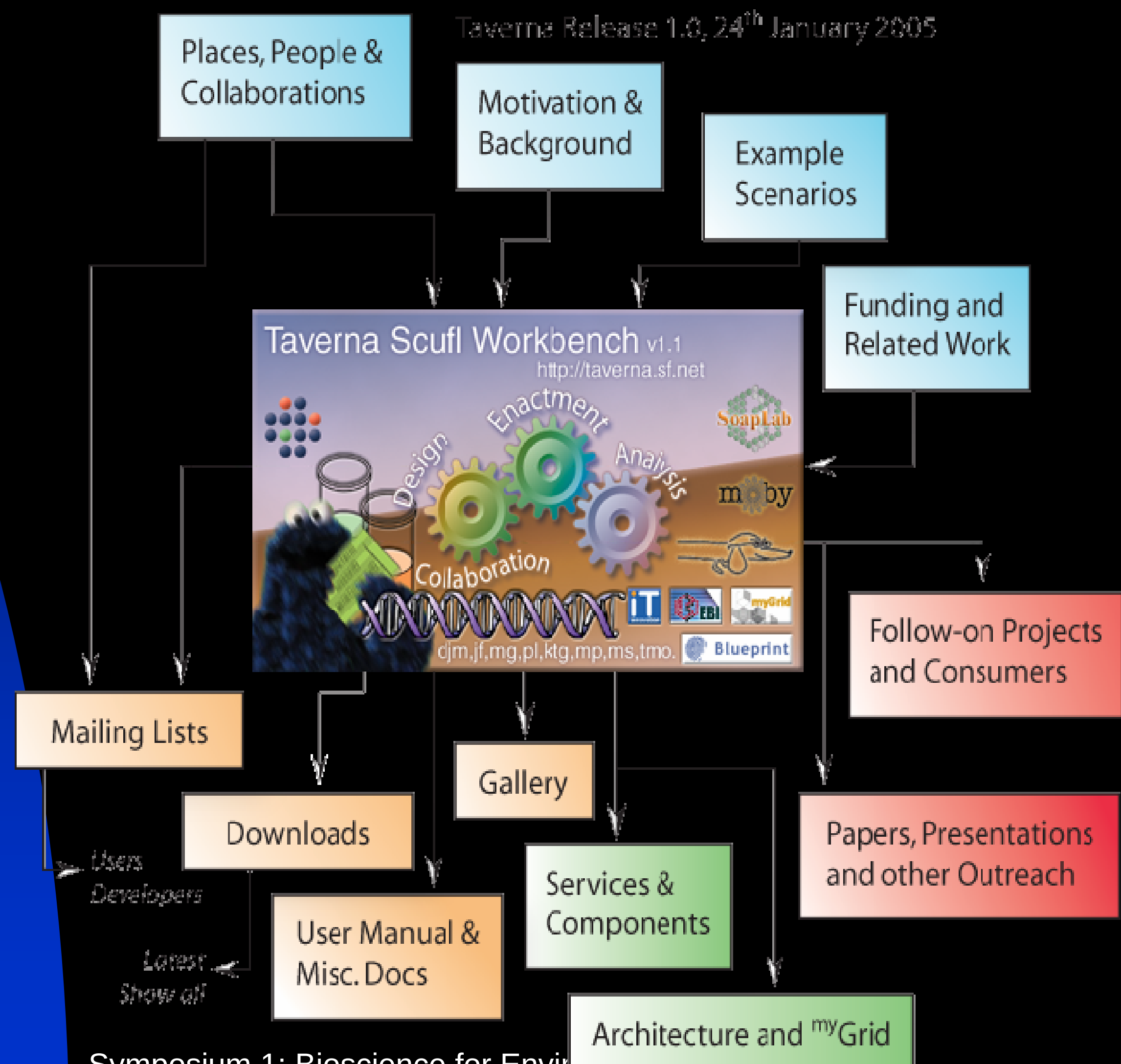
volume 6, number 26
July 1, 2002

in this issue

APBioNet Enlists Lion, Cray, KOOPrime For Grid-Based Interoperability Project

THE ASIA Pacific Bioinformatics Network is bringing together a number of industry partners to build a grid-based system to share bioinformatics applications and workflows throughout the Asia Pacific region. The testbed project, a brainchild of Tan Tin Wee, secretariat of APBioNet and an associate profes-

Taverna-FreeFluo-myGrid workflow integration from UK eScience



Research Process Integration for Life Sciences

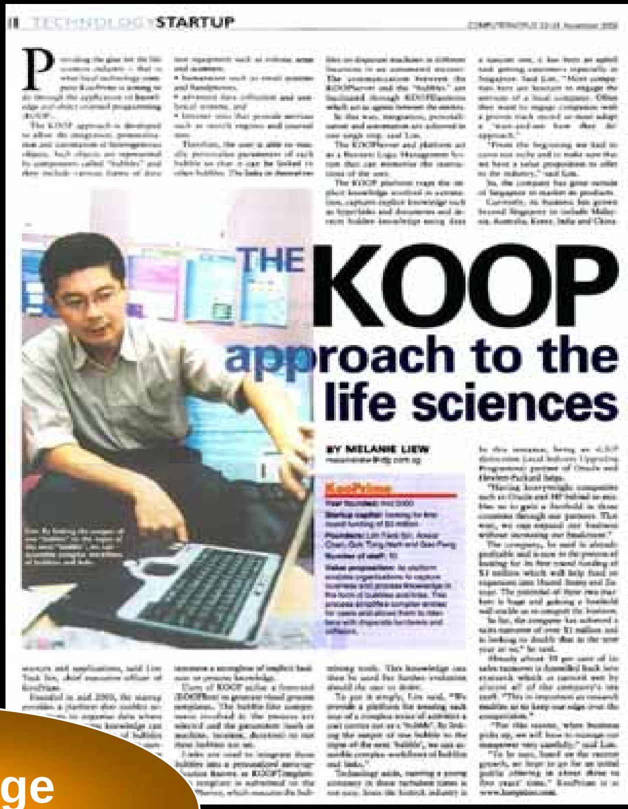
Data Management

Legacy Systems
(e.g. HR, CRM, Analytical)

Application Management

Knowledge Management

Equipment

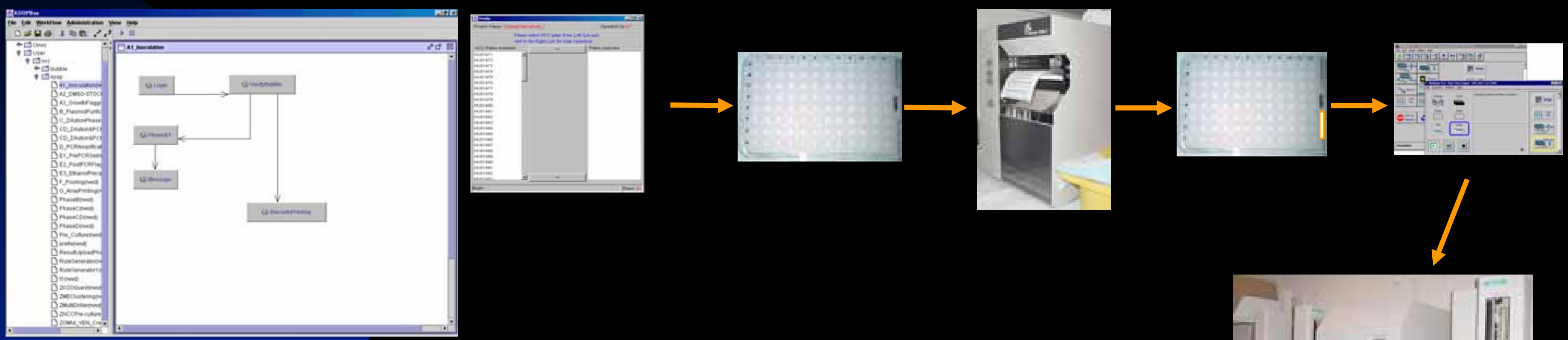

TECHNOLOGY STARTUP

THE KOOP approach to the life sciences

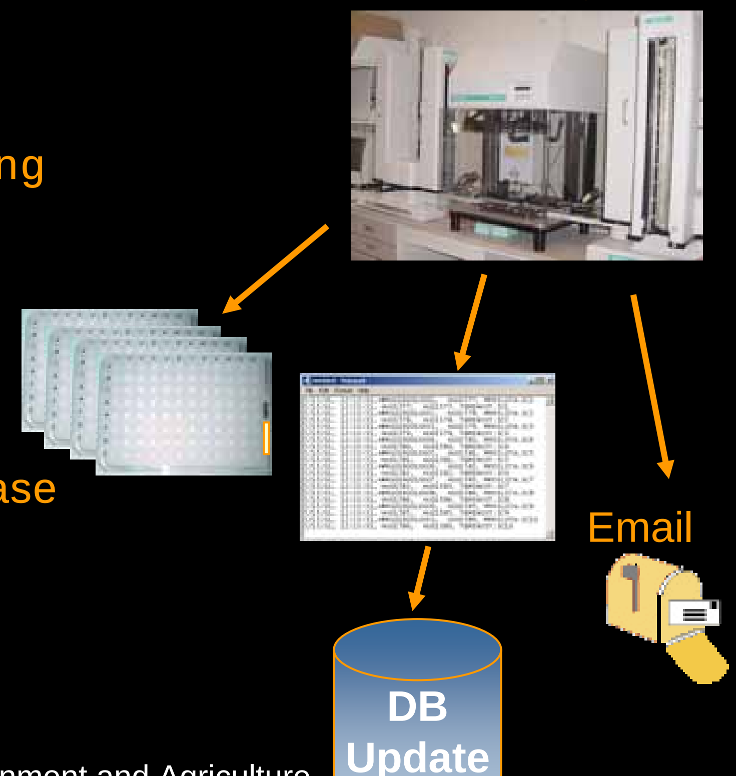
BY MELANIE LIEW

www.kooprime.com
KOOPrime Pte Ltd 71A, Tanjong Pagar Road Singapore 088492 Tel: 327 2433 Fax: 327 2438

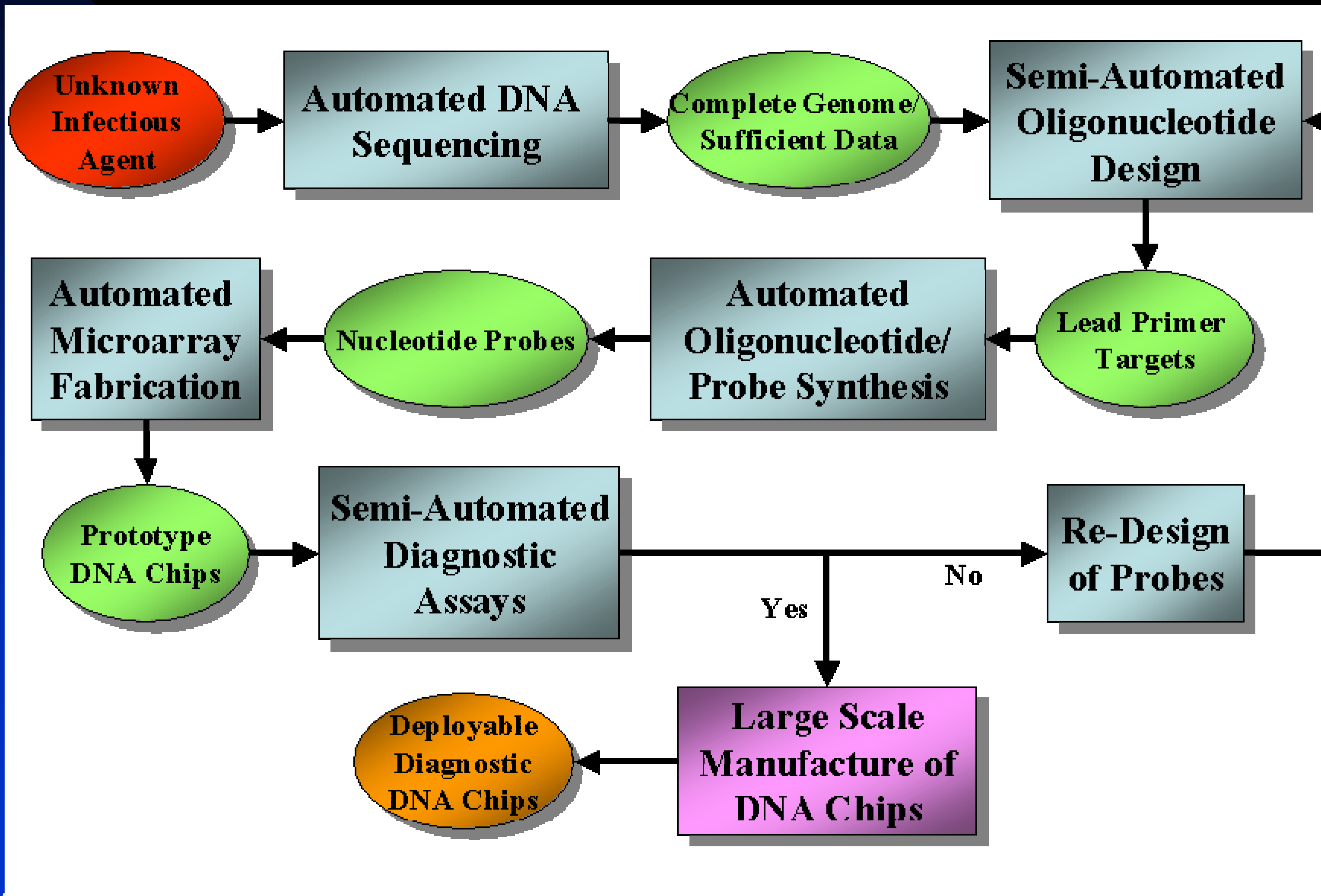
Laboratory Integration



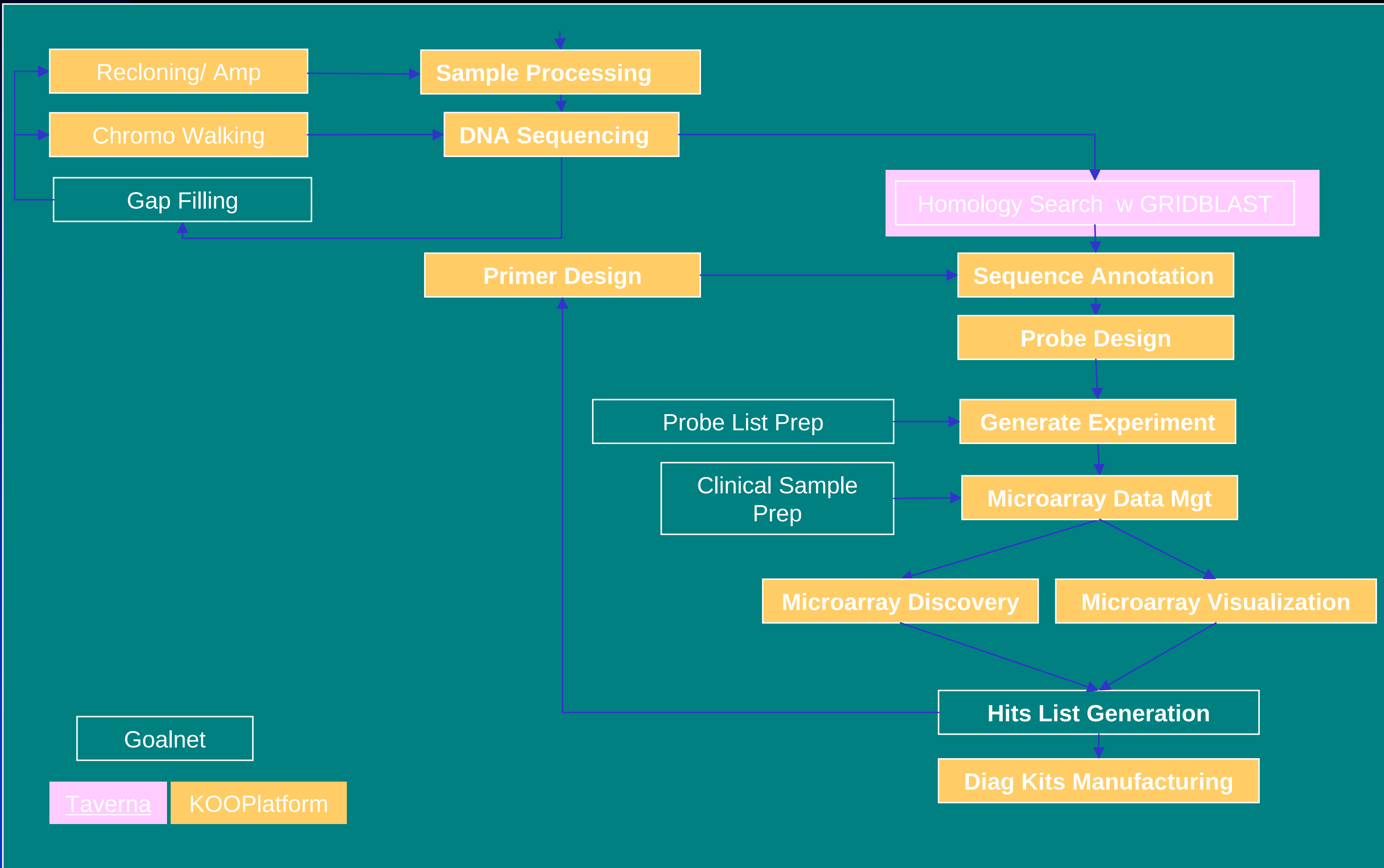
- Allows users to select vendor plates for processing
- Generate in-house plates from vendor plates
- Print barcodes for each selected plate
- Start up legacy dispenser software
- Auto-import output files of dispenser into database
- Email user if there is any error in processing

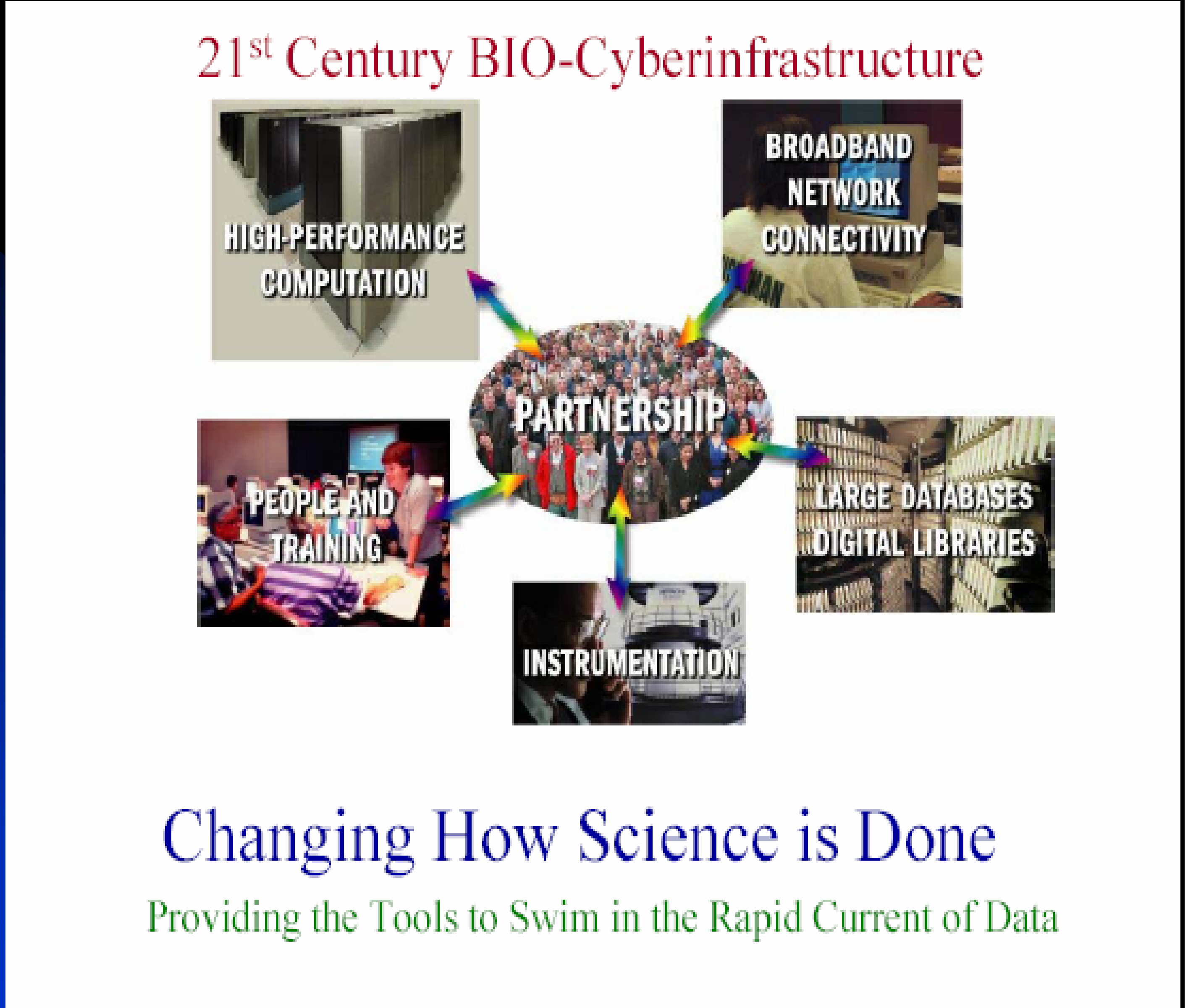


Pipeline from R&D to Bio-Manufacturing of Diagnostics for Bird Flu and other emerging pathogens



A Typical Bio-Manufacturing Workflow





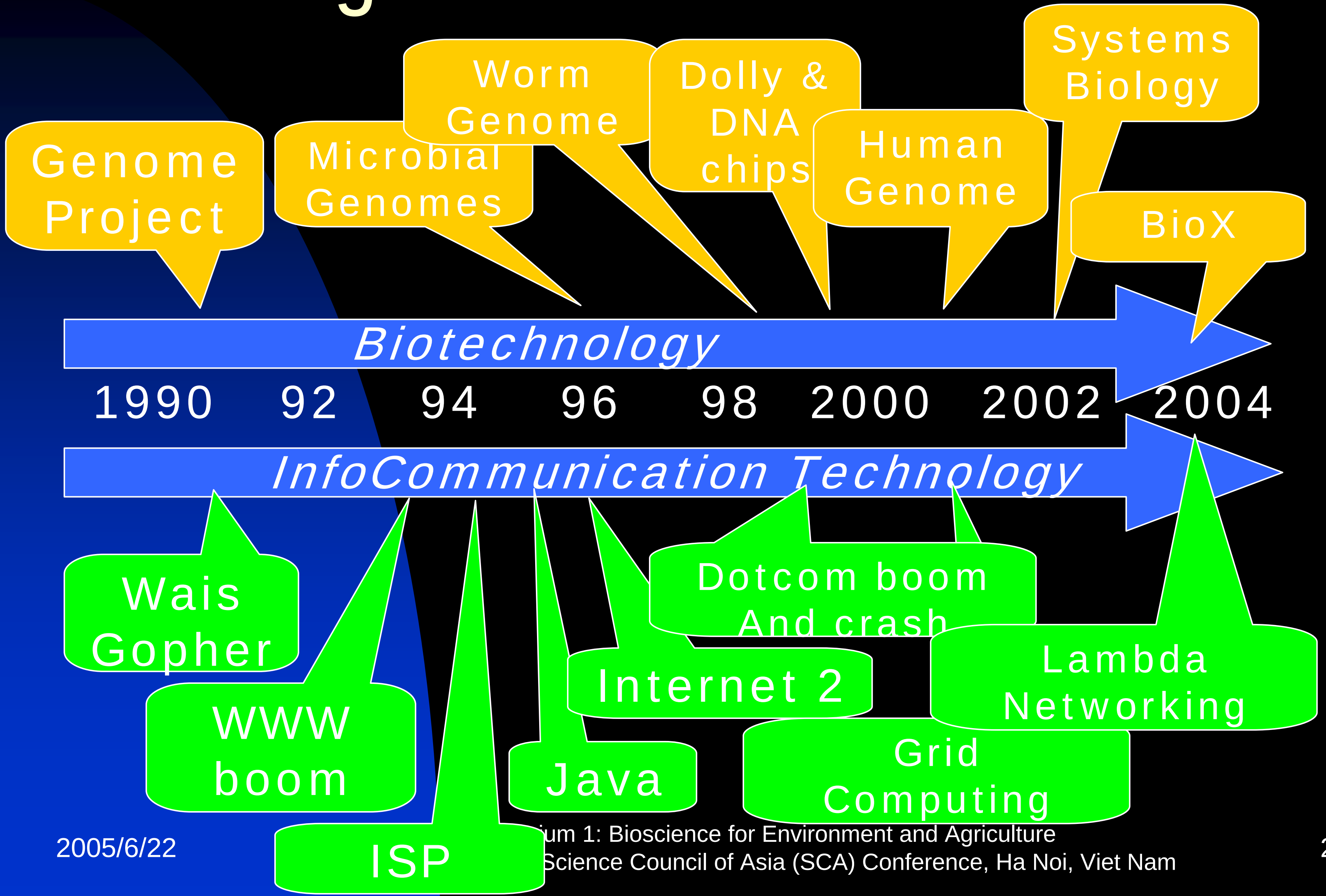
From John Wooley et al 2005

The Language of Life Sciences

- "Calculus, in managing the infinitely small but large scale of events filled with redundancy, has been the language of the physical sciences."
- "Biology has high information content, along with individuality, historicity and contingency… the biological sciences as a research discipline are said to be an information science. As such, information technology is the language of the life sciences, managing the discrete, non-symmetric, largely non-reducible, unique nature of biological systems and observations."

John Wooley 2005

Biotech and InfoComm Technologies: Parallel Growth



Framework of Bioinformatics Development in Asia Pacific from 1991-2005

POLICY

RESEARCH

EDUCATION & Manpower Training

Compute INFRASTRUCTURE

DATA INFRASTRUCTURE

NETWORK INFRASTRUCTURE

Planning

Coordination

Collaboration Cooperation

Training a new generation of Biologists for the Environment and Agriculture

- We must hook our individual computers in the worldwide network that gives us access to daily changes in the database and also makes immediate our communication with each other. The programs that display and analyze the material for us must be improved – **and we must learn to use them more effectively.** *Walter Gilbert Nature 1991*
- **We must equip our students with the skills to carry tomorrow's research today!**

APBioBox APBioKnoppix collaboration to deliver free Computational Tools

- A/P Tan Tin Wee,
  National University of Singapore
- Adjunct Professor Shoba Ranganathan, NUS and Chair Professor, Macquarie University, Sydney
- Ong Guan Sin, Consultant programmer, Singapore Computer Systems Pte Ltd
- Funded by International Development Research Centre of Canada, under their PAN Pan Asia Networking ICT grants

CGIAR's New Generation Challenge

- Programme to bring about a change in agricultural research and development
- Equip scientists with enabling skills
- Capacity building
- Crop Genomics and Informatics training programme

S* Life Science Informatics Alliance

- Stanford*Karolinska*Uppsala* SouthAfricanBioinformaticsInstitute (SANBI)*UCSD*NUS*USydney collaboration
- Online bioinformatics education since 2001 free of charge
- More than 1000 students have taken part
- Collaboration with AVIST

Conclusion

- Life Sciences, including environmental and agricultural sciences, have gone beyond simple sequence analysis of the 1980s and genome analysis of the 1990s.
- Computational aspects of biology will continue to increase
- Database building, knowledge structuring and knowledge organisation and integration will continue
- Imaging, Modeling and Simulation will emerge
- Workflow integration and Pipelining will continue to accelerate high throughput, high efficiency research
- In additional to observational/experimental, all biologists, **including environmental and agricultural researchers,** must be informationally and computationally competent in the 2000s
- “We must act now for the future”
  Kiyoshi Kurokawa, 2005

Thank you SCA!

Thank you my hosts, Ha Noi, Viet Nam