# Clinical Genomic Systems and related Clinical Standards

Yves A. Lussier
Columbia University

© 2002 – Yves A. Lussier, Columbia University

# Preview of Take Home Points

- Assessment of the post-genomic impact on standards, data structures and clinical information systems is intrinsically difficult
- Informatics Standards and Systems for individualized medicine are already available!
  - but, diffusion of new knowledge is slow
  - Molecular pathology will first be impacted and its information systems.
  - PACS vendors may have a unique advantage over other health IT systems.

© 2002 – Yves A. Lussier, Columbia University

# Outline

- Introduction
  - Medicine, Healthcare and Wellness
    - Genomic era
    - Post-Genomic Paradigm Shift
- Post-Genomic Impact
  - D A T A, Information and Knowledge
  - Standards
  - IT
- Future Directions

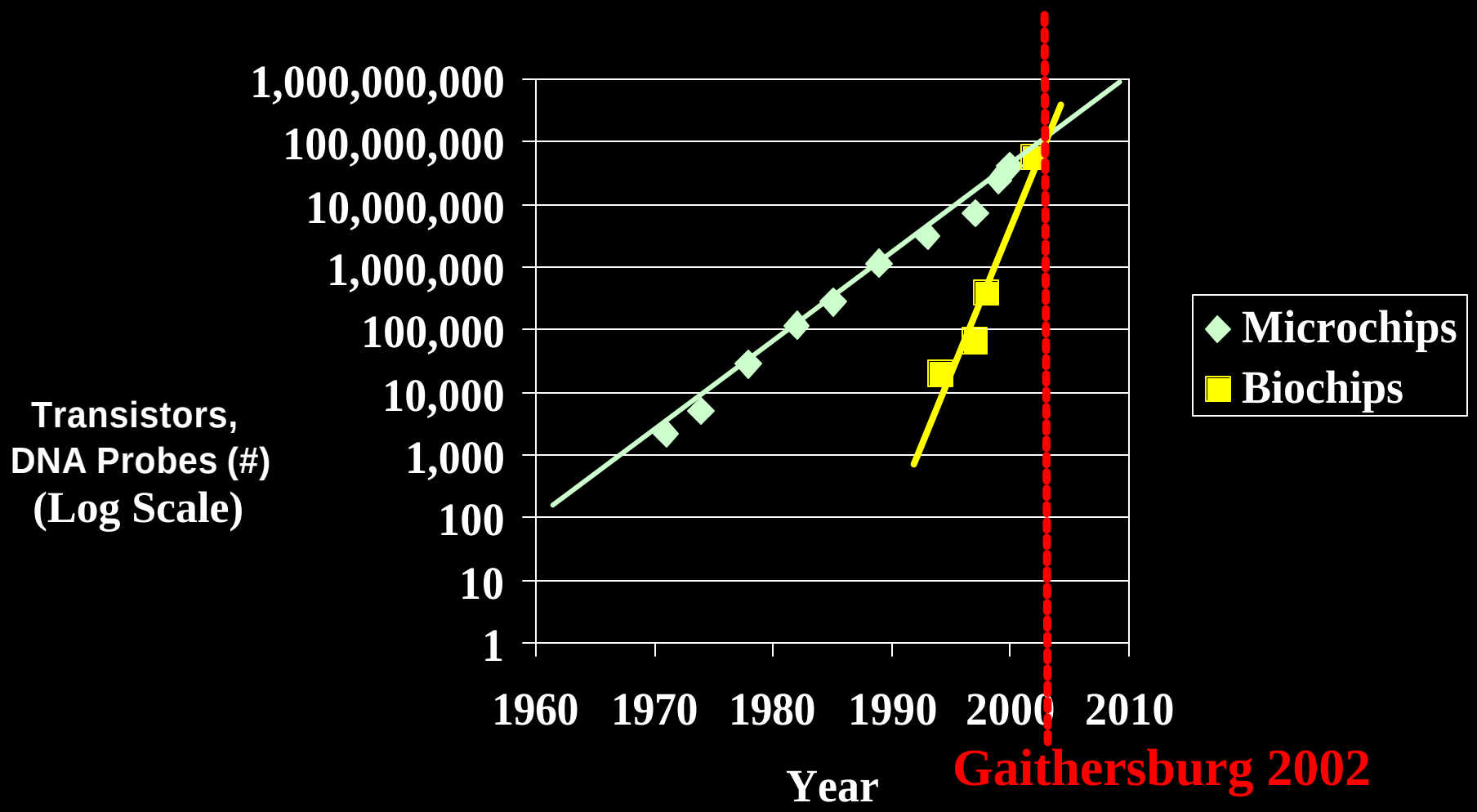
© 2002 – Yves A. Lussier, Columbia University

# Molecular Bio

- 100 trillion cells / human
- 2 sets of 23 chromosomes / cell
- 3 billion base pairs / cell
  - 1 Billion words, 800 bibles
- 35,000 Genes
- >100,000 proteins?
- 99.9% similar

© 2002 – Yves A. Lussier, Columbia University

# Columbia Knowledge Technology Lab




© 2002 – Yves A. Lussier, Columbia University

# Outline

- Introduction
  - Medicine, Healthcare and Wellness
    - Genomic era
    - Post-Genomic Paradigm Shift
- Post-Genomic Impact
  - Data, Information and Knowledge
  - Standards
  - IT
- Future Directions

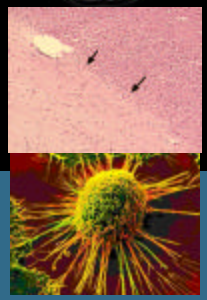
© 2002 – Yves A. Lussier, Columbia University

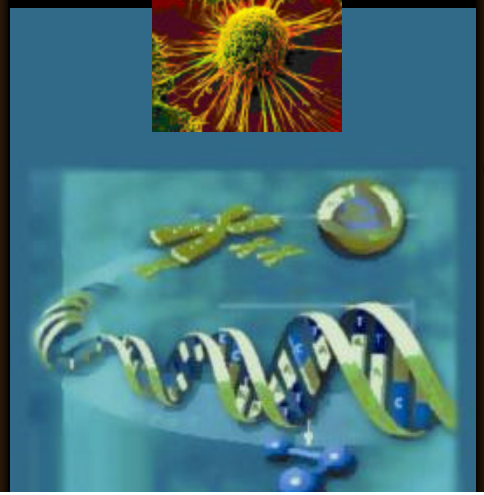
# Information Value Chain

| | | | | | |
|---|---|---|---|---|---|
| Storage Size | **Data** | > | Information | > | Knowledge |
| **Value** | Data | < | Information | < | **Knowledge** |

© 2002 – Yves A. Lussier, Columbia University

# D A T A





| meters | Measure | |
|---|---|---|
| $10^{0}$ | Organism | ICD-10, SNOMED |
| $10^{-1}$ | organs | 1 |
| $10^{-2}$ | | SNOMED |
| $10^{-3}$ | | |
| $10^{-4}$ | Tissue | |
| $10^{-5}$ | Cells | SNOMED |
| $10^{-6}$ | Organelles | SNOMED, ??? |
| $10^{-7}$ | Virus | ? |
| $10^{-8}$ | DNA | - |
| $10^{-9}$ | bases | Gene Ontology<br>Quaternary code |

# OSI-7 Reference Model

| | |
|---|---|
| **Application** | Controlled Biomedical **Terminology** |
| | HL-7 |
| | Presentation |
| | Session |
| | Transport |
| | Network |
| | Data Link |
| | Physical |

© 2002 Yves A. Lussier, Columbia University

# Biomedical

## D A T A, information, knowledge

**Genomic D A T A**

1 probe ➔ 1kb,

6 k probes ➔ 6MB

6 M probes ➔ 6 GB

- **PACS Clinical D A T A**
  - 15 MB, chest X-Ray
  - 30 MB – one mammogram
  - 73 MB, 140 images, CT chest
  - 500 MB, 32 images, Biplane rotation angiogram

**Genomic Information**

Image analysis & Statistics

- Can be re-interpreted, keep all versions

4-10 bytes/probe

- 4-10 kb per 1000 probes

**Clinical Information**

- 1-2 kb per study
- human interpretation of data
- Can be re-interpreted, keep all versions
- Can have different authors for the same image interpretation

**Genomic Knowledge**

dozen of bytes

- **Clinical Knowledge**
  - Smaller – several relationships ➔ dozen of bytes

© 2002 – Yves A. Lussier, Columbia University

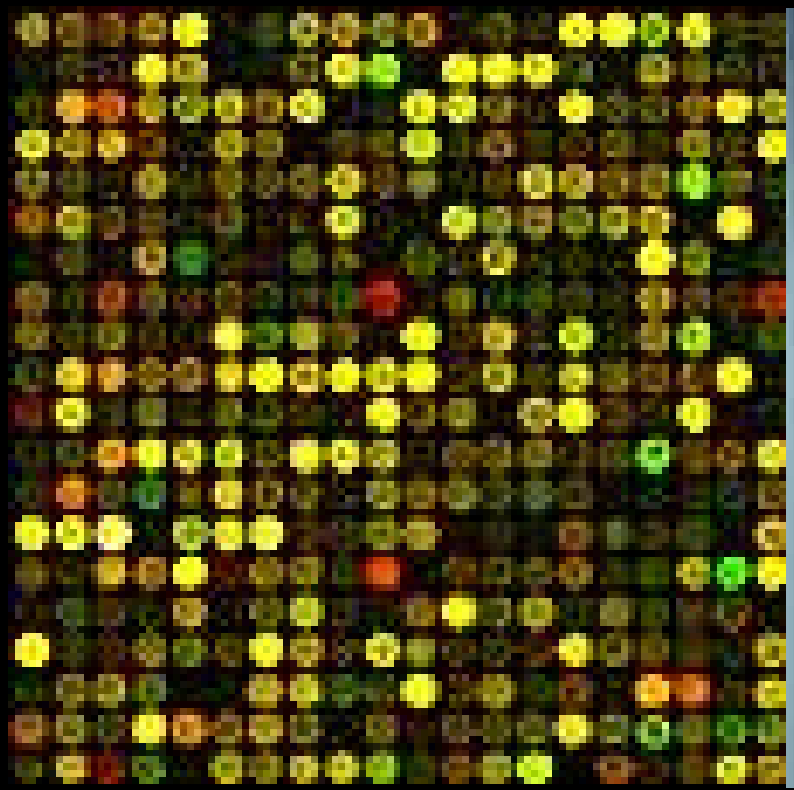
# OSI-7 Reference Model

(Open System Interconnection, ISO)

| | |
|---|---|
| **Application** | **Knowledge Representation** |
| | Controlled Biomedical Terminology |
| | HL-7 |
| | Presentation |
| | Session |
| | Transport |
| | Network |
| | Data Link |
| | Physical |

© 2002 – Yves A. Lussier, Columbia University

# PACS for Clinical Genomics



© 2002 – Yves A. Lussier, Columbia University

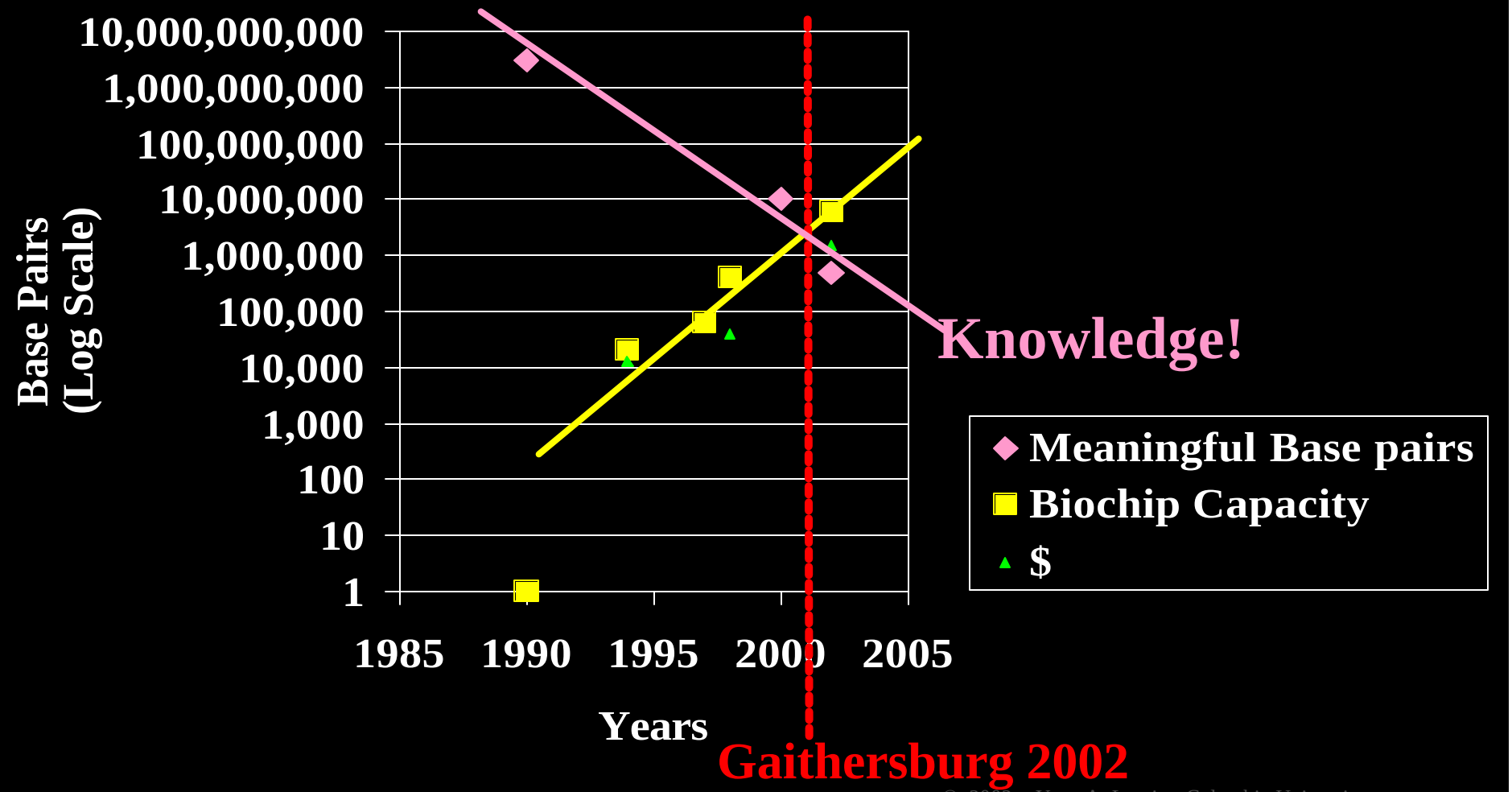
# Clinical Genomics Does not Follow Moore's Law

- Failed Predictions
  - Genome in 2001, not 2005
  - 100,000 genes predicted: 30,000 found
  - 10,000,000 SNPs?
- Paradigm shift 2001/ 12
  - **SNPs goup in 500,000 haplotypes**

*Science* **294**: 1719-1723 (2001)..* Perlagen

© 2002 – Yves A. Lussier, Columbia University

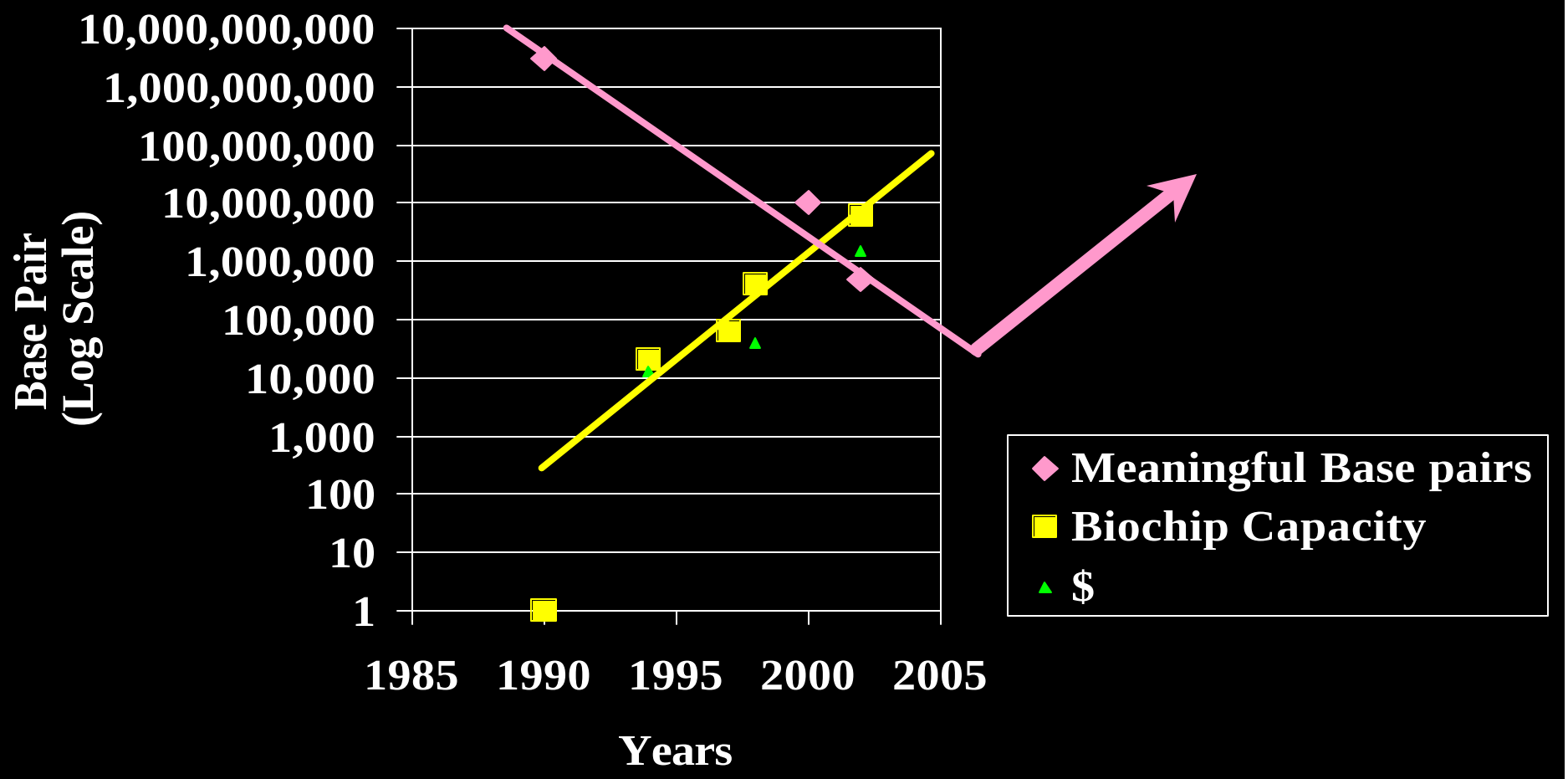
# Functional Genome Paradigm Shift




© 2002 – Yves A. Lussier, Columbia University

# Good Bye SNPs,

## Welcome mTAtions (<1% freq.)`& Environment




© 2002 – Yves A. Lussier, Columbia University

# Projections
# Synthetic < 2010

2002

- 500k meaningful bp

2005

- 100k meaningful bp
  - Other paradigm shift = PH
- Computing DNA for simple tests = cheaper, simpler, no biochips!
  - Testing one complex common disease, as easy as a pregnancy test!

© 2002 – Yves A. Lussier, Columbia University

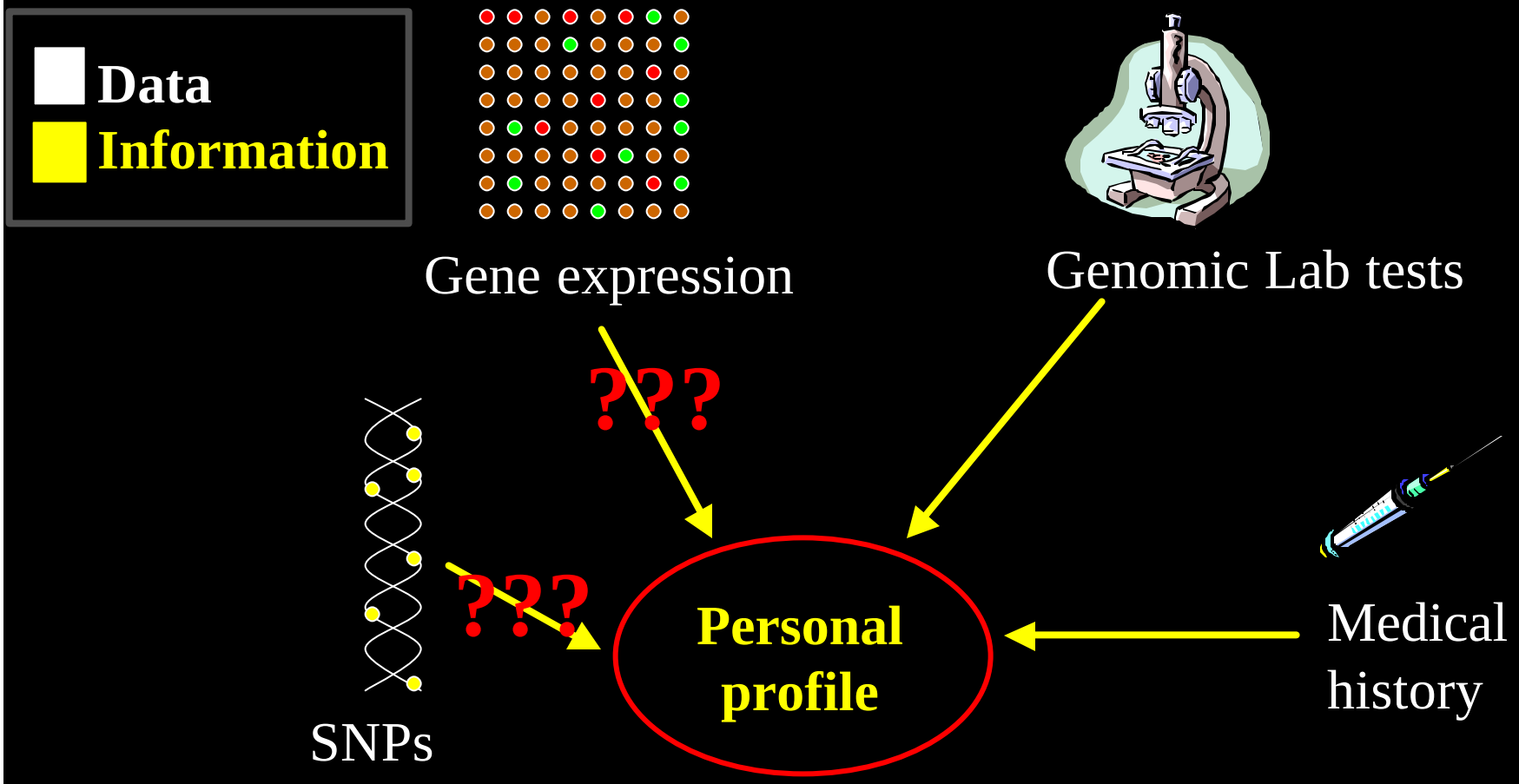
# Pharmacogenomics

- Cytochrome P450

  Home Test
  - Proton Pump Inh.: Omeprazole
  - Caffeine,
  - NSAIDs: Naproxen, Ibuprofen
  - Oral Hypoglucemic: Glyburide, tolbutamide
  - Antiepileptic
  - Angiotensin II inh.: Losartan
  - Beta Blockers: metoprolol
  - Antidepressive

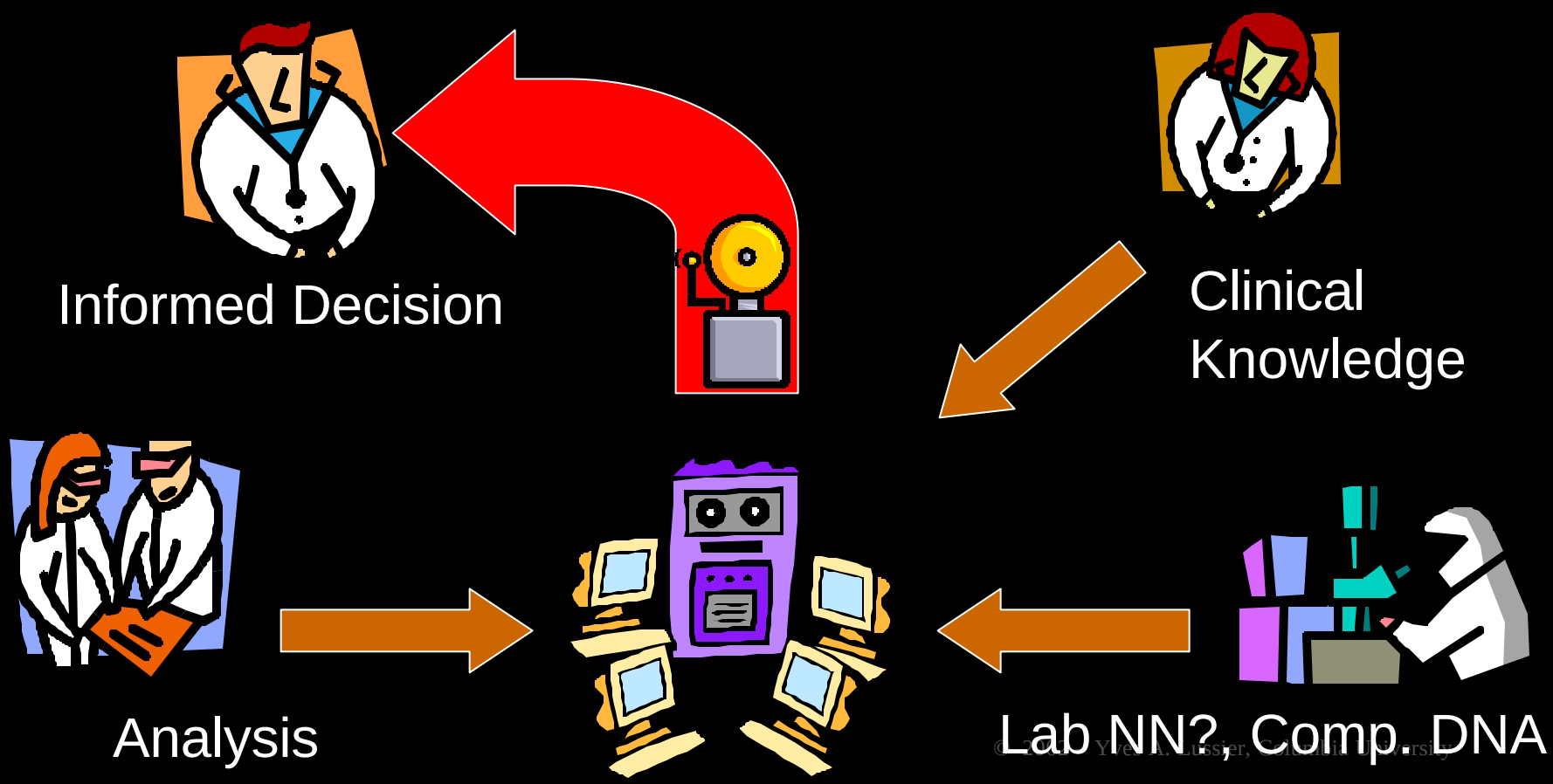
© 2002 – Yves A. Lussier, Columbia University

# Clinical Deployment of Genomic Data




© 2002 – Yves A. Lussier, Columbia University

# Clinical-Genomic Knowledge Deployment for Individualized Medicine




© 2006 Yves A. Lussier, Columbia University

# Strategies "*à la carte*"

- Use biomedical standards for computable information and knowledge technologies HL-7
  - terminology (GO, ICD-10,…)
  - knowledge (Arden, GLIF)
- High Throughput Data Technologies ? ➔ Team up
  - Advanced Terminology Informatics Systems: Health Languages (US), Apelon (US), 3M medical (US), Purkinje (Canada),
  - Molecular Pathology Systems
  - PACS vendors
  - Pursue Clinical Genomic Informatics Research

© 2002 – Yves A. Lussier, Columbia University

# Take Home Points "*du jour*"

- Assessment of the post-genomic impact on standards, data structures and clinical information systems is intrinsically difficult
- Informatics Standards and Systems for individualized medicine are already available!
  - but, diffusion of new knowledge is slow
  - Molecular pathology will first be impacted and its information systems.
  - PACS vendors may have a unique advantage over other health IT systems.

© 2002 – Yves A. Lussier, Columbia University

# Projections
# Renewed Complexity > 2010

2010

- 40,000 disease
  - Several variant for common disease
  - > 5 million mutations

2020

- 40k disease
  - > 1 million modifiers!
  - Mutations, mutations, mutations

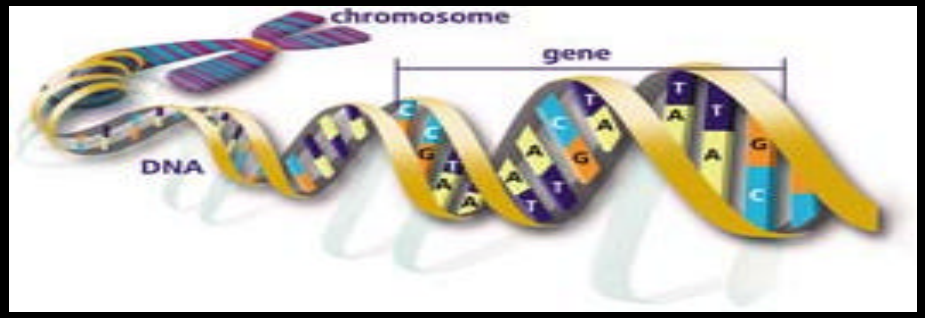
© 2002 – Yves A. Lussier, Columbia University

# Acknowledgements


- Blackford Middleton
- Jim Cimino
- Stephen Johnson


© 2002 – Yves A. Lussier, Columbia University

# Columbia Knowledge Technology Lab




Yves A. Lussier

212.305.0939

Lussier@dmi.columbia.edu

© 2002 – Yves A. Lussier, Columbia University