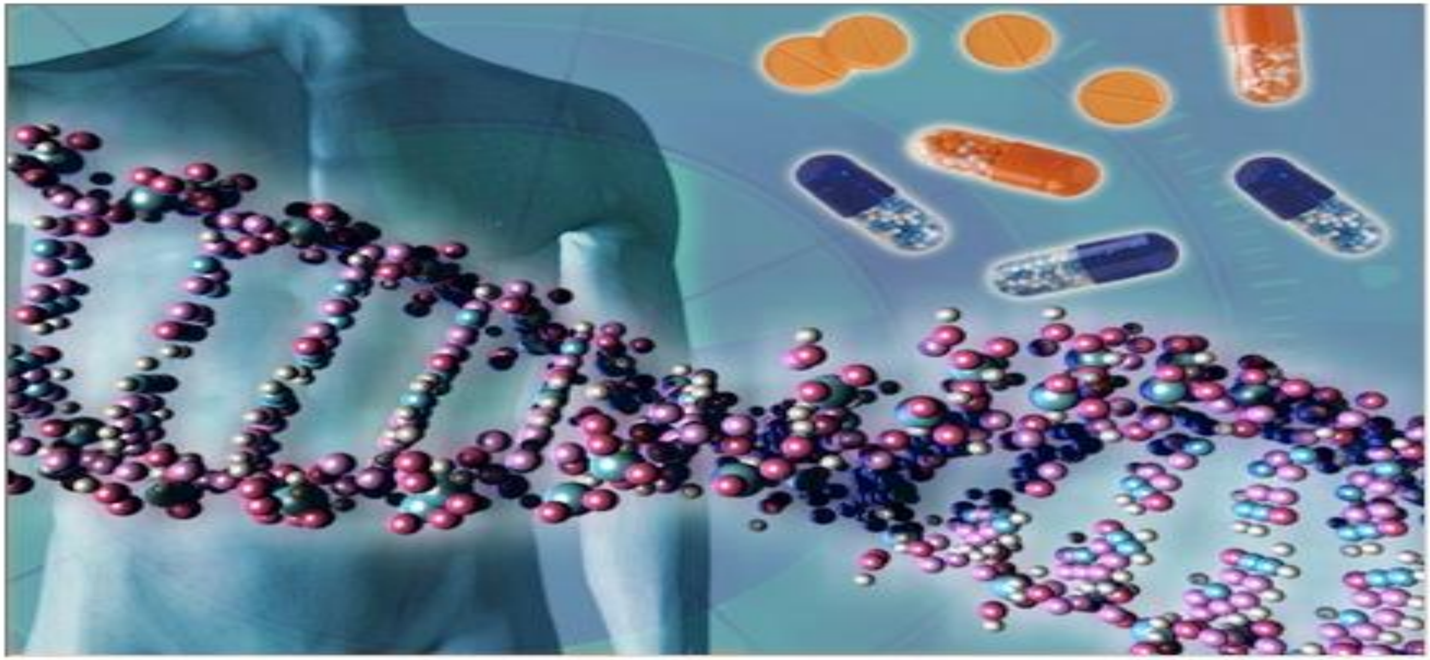
VirginiaTech
Invent the Future

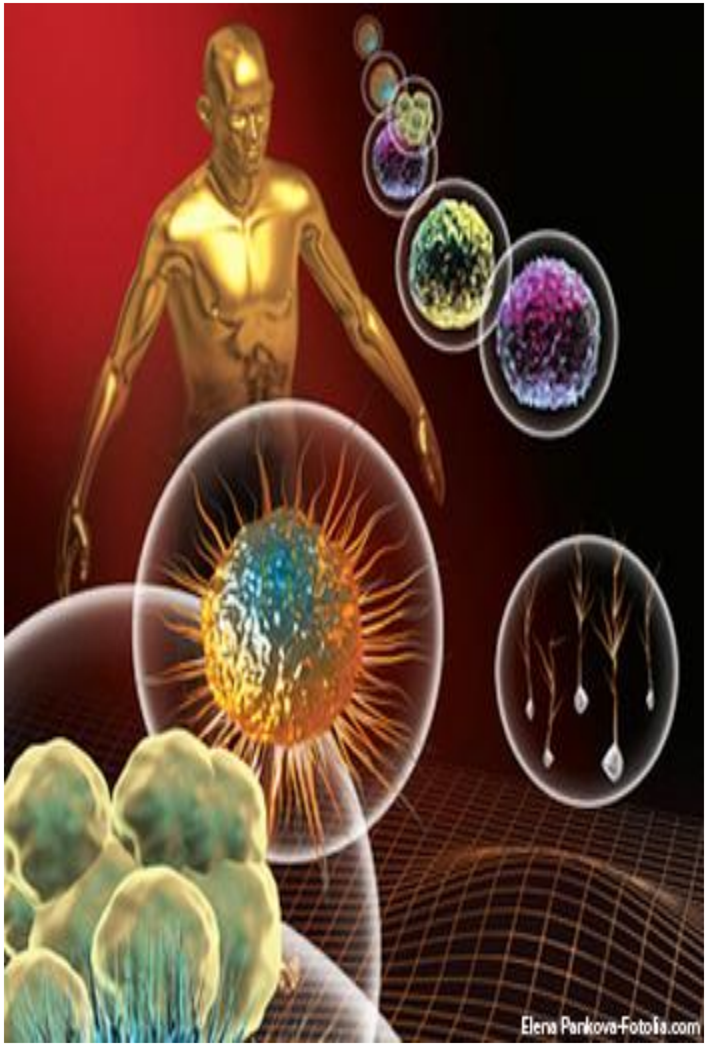
Institute for
Critical Technology *and* **Applied Science**

Genome Sequencing



Mohamed N. Seleem

From Gene to Genome



Elena Pankova-Fotolia.com

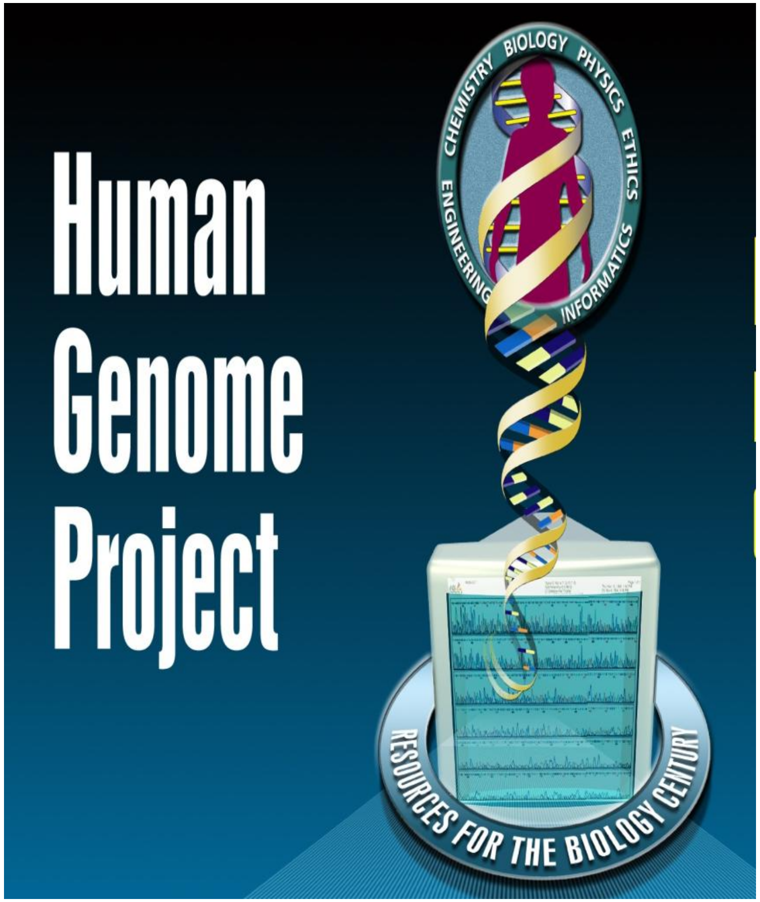
•Sequencing of the whole genome of the Organism

- •Sequence must be annotated
  - – Location of genes (locationofgenes)
  - – Location of transcribed regions (coding region)
  - – Location of promoters, start codons and terminators
  - – Function of other DNA sequences
  - – Translated Protein and assigned function

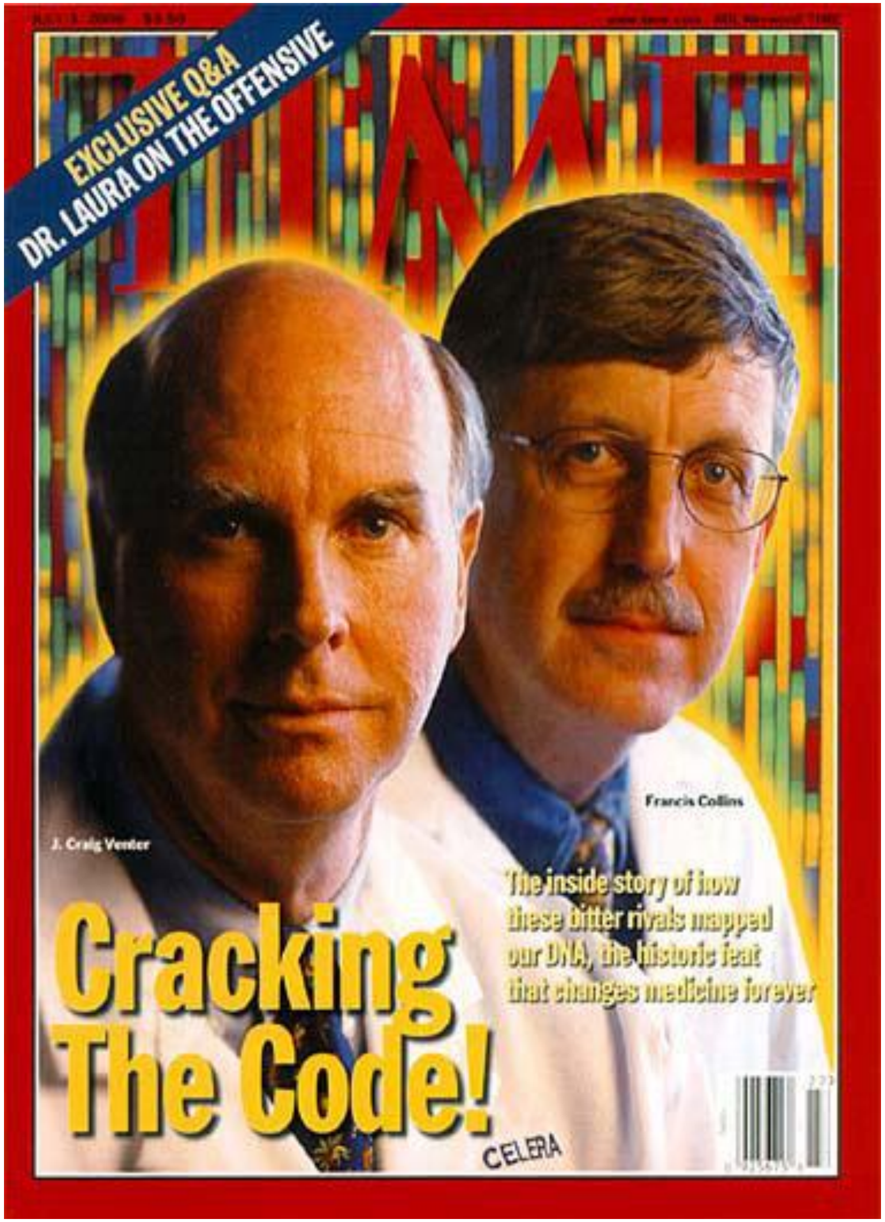
Human Genome Project

1990–2003

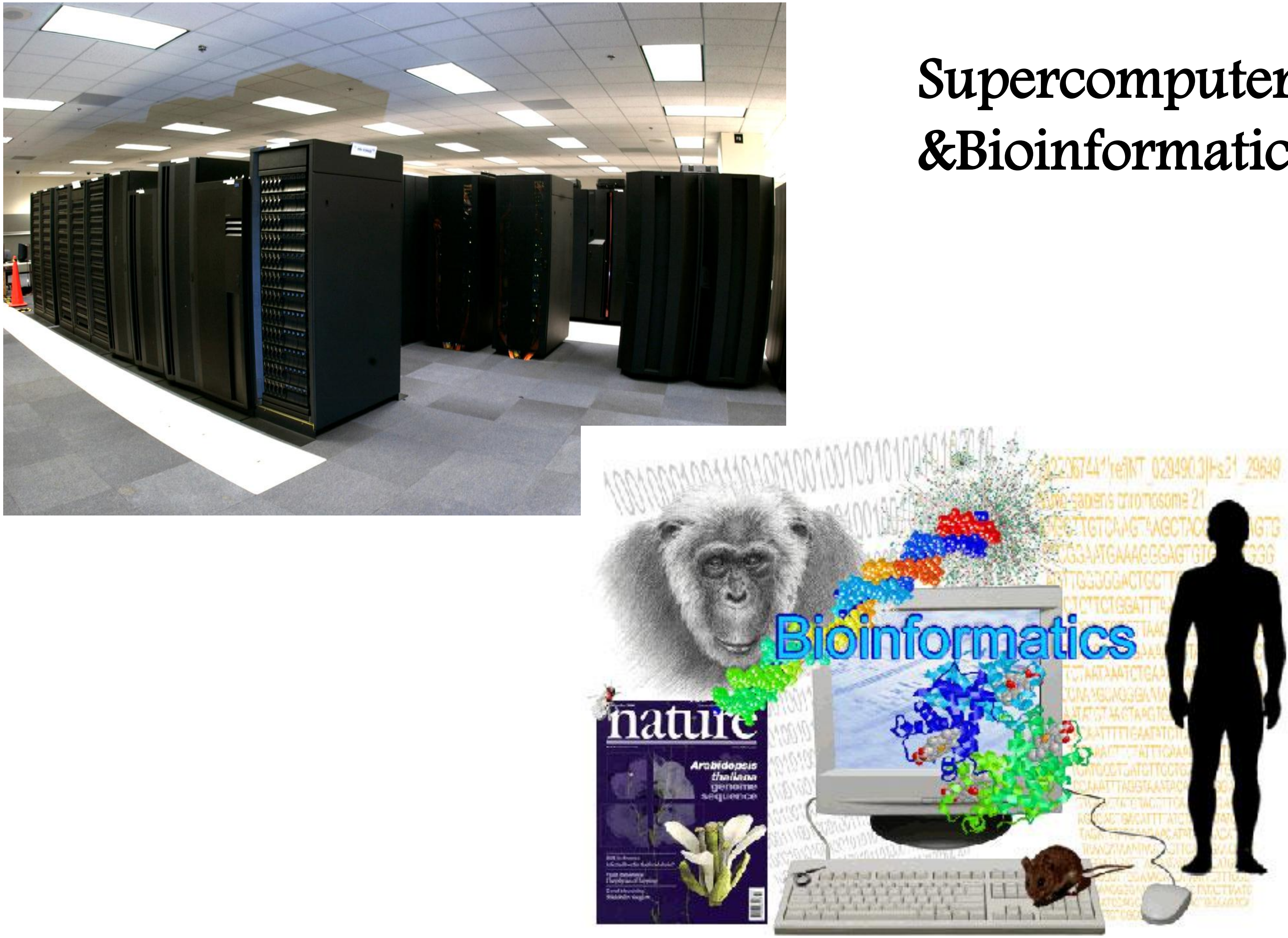
$3 billion

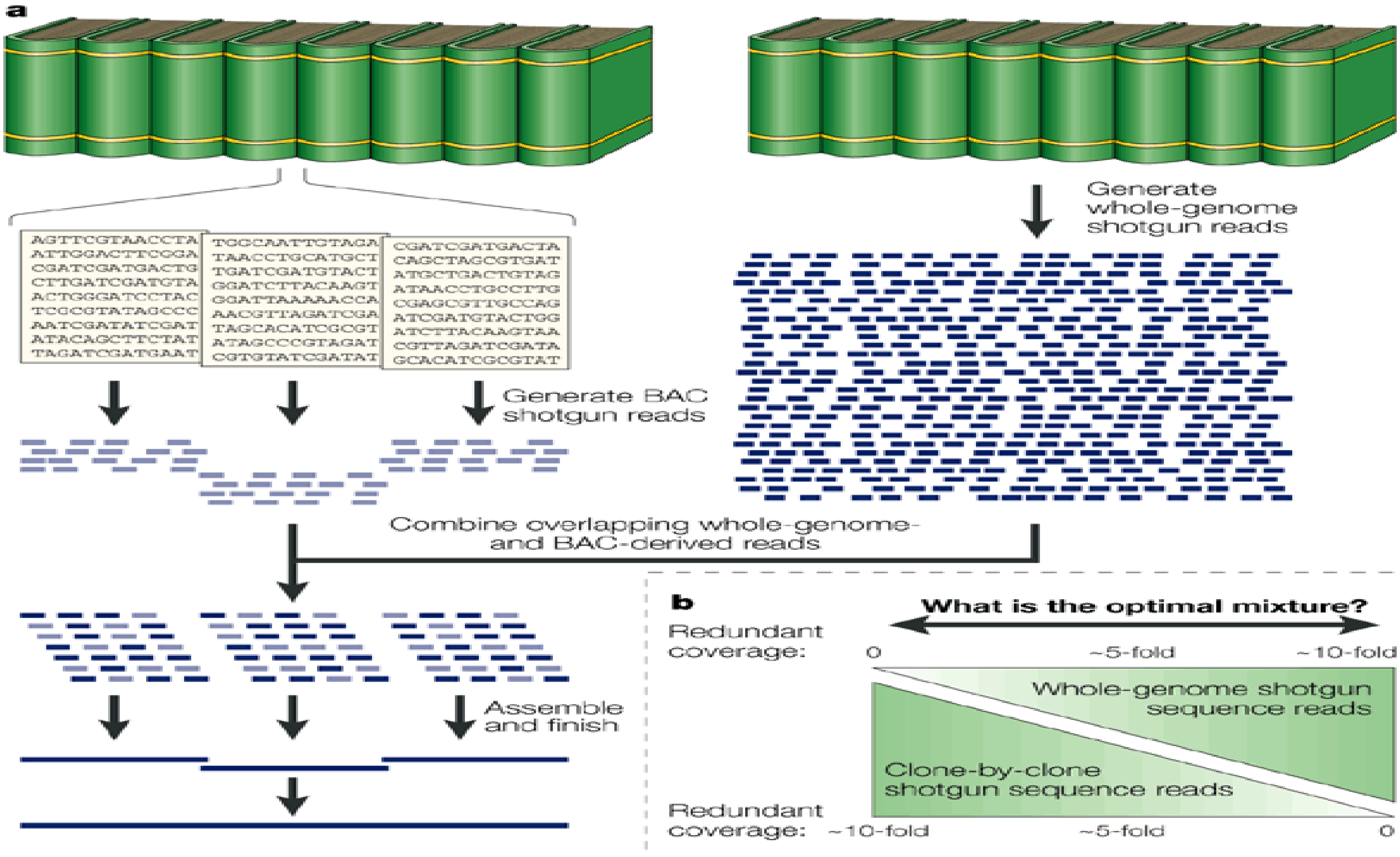


Francis Collins
J. Craig Venter



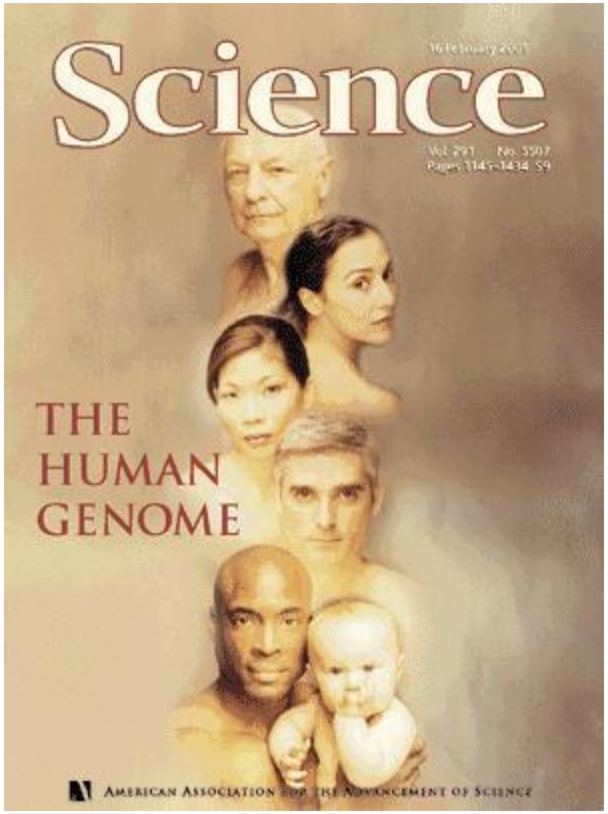
Supercomputers &Bioinformatics

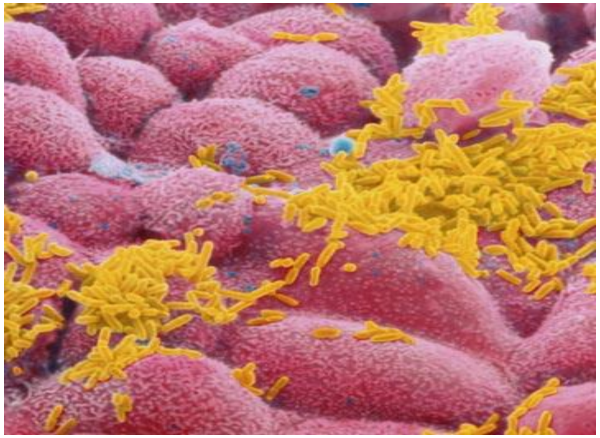




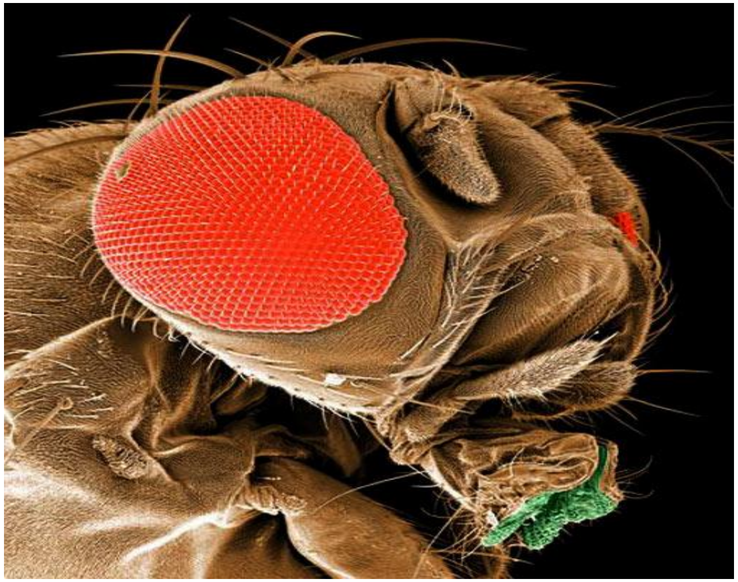
Nature Reviews | Genetics

Draft 2000
Complete 2003

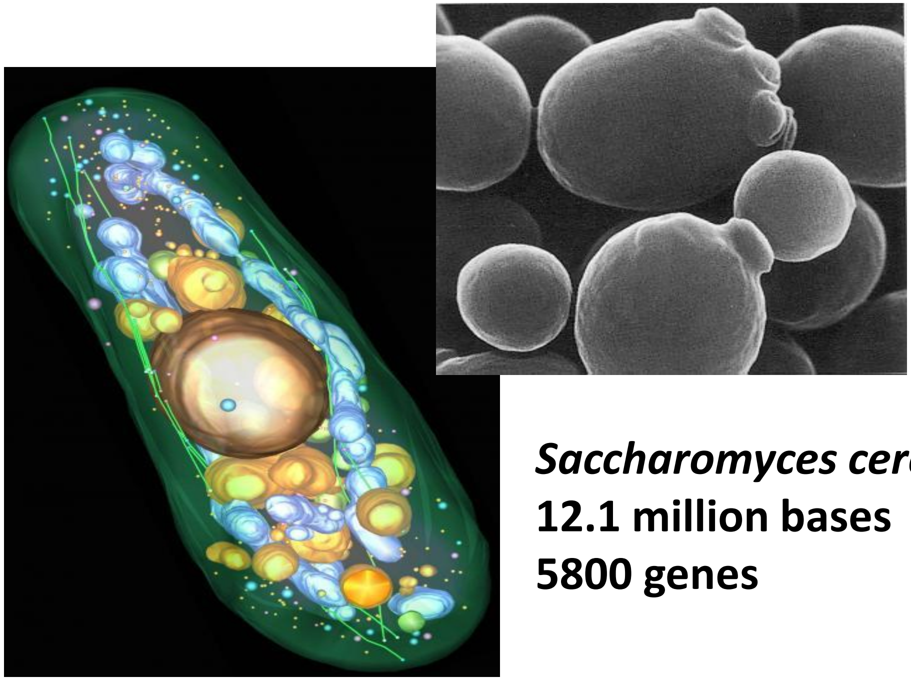




H. influenzae
1.8 million bases
1700 genes



Drosophila melanogaster
137 million bases
13,700 genes

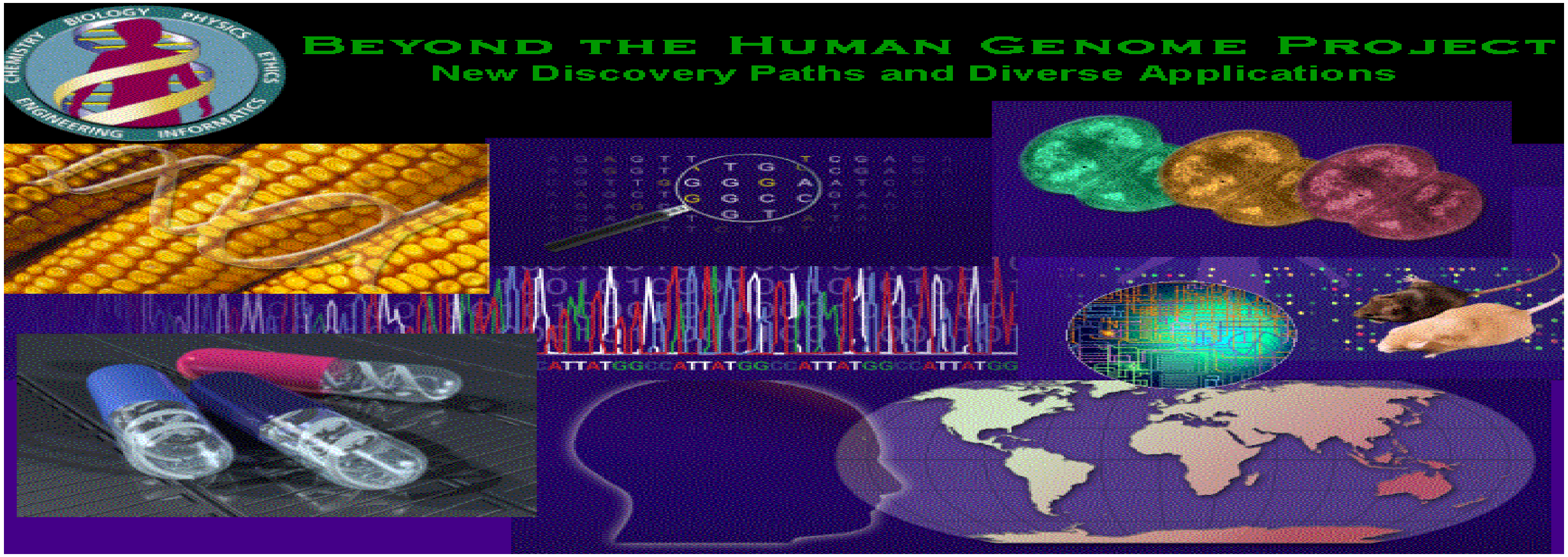


Saccharomyces cerevisiae
12.1 million bases
5800 genes

Caenorhabditis elegans
97 million bases
19,000 genes

***Homo sapiens* (human)**
3.2 billion base
~25,000 genes
2% only code for protein
100,000 proposed earlier
40,000 after first draft

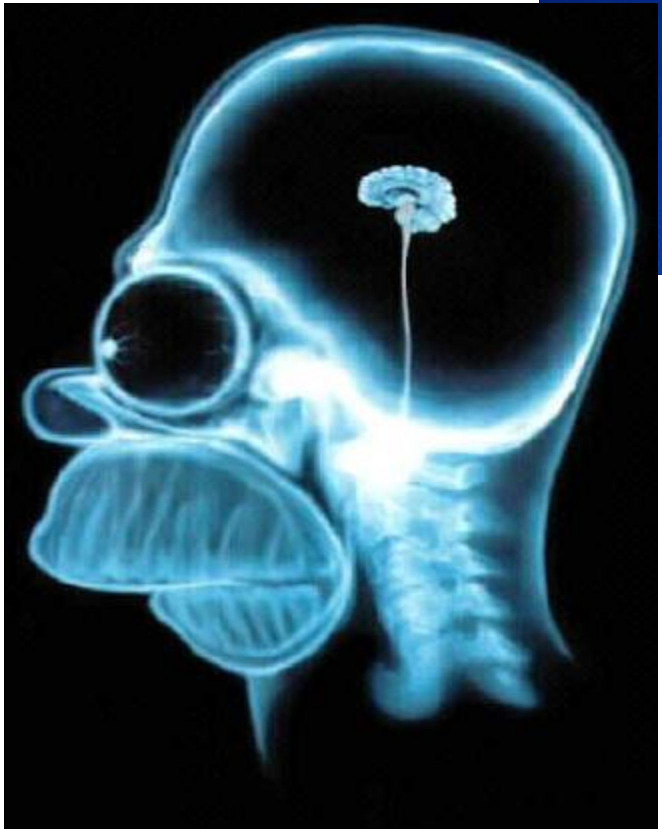
***Oryza Sativa* (Rice)**
430 million base
60,000 genes



$1,000 Genome by 2015-2020

Every Child Genome

Sequencing is just letters



How to use it?