Computational Genomics

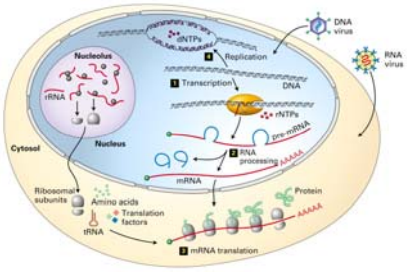
Introduction to cell biology, genomics, development, and probability

Eric Xing

Lecture 1a, January 18, 2007

Reading: Chap. 1, DTM book

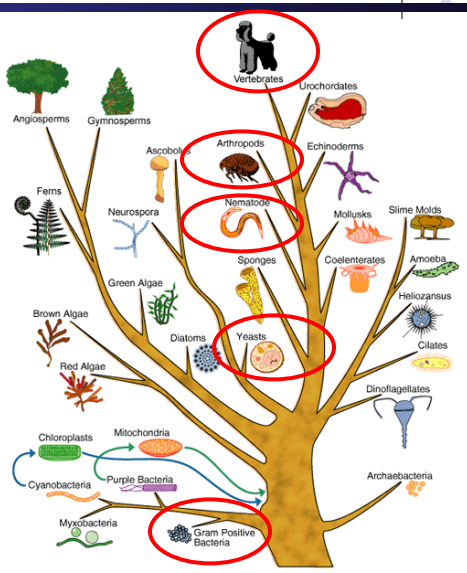
Introduction to cell biology, functional genomics, development, etc.



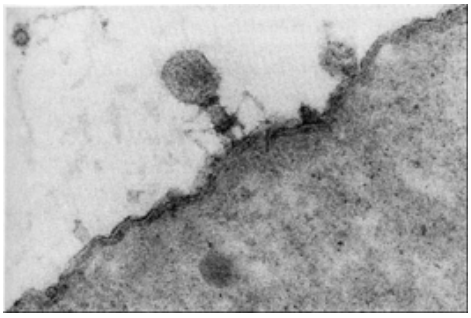
- What system to use?
- How informative it is?
- How much we can generalize?
- How much shall we believe?
- How much we can validate our results? How?

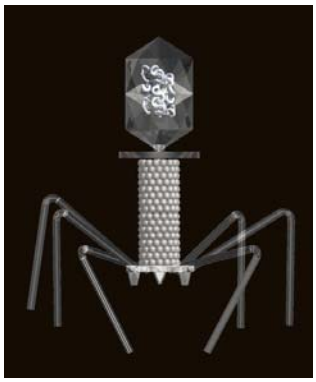
- How to avoid being **naïve** when doing compbio?

Model Organisms

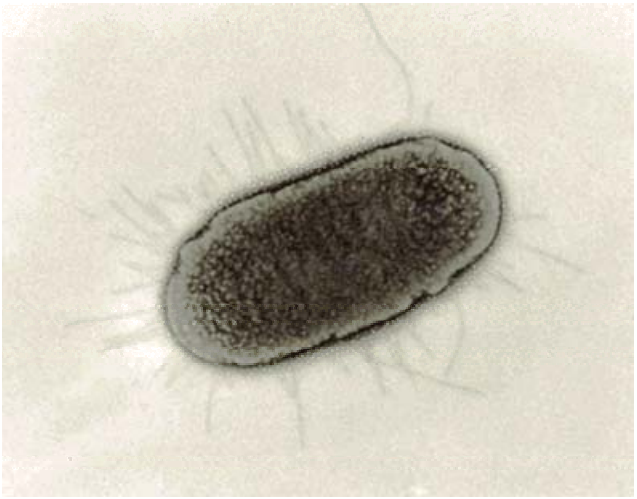


Bacterial Phage: T4

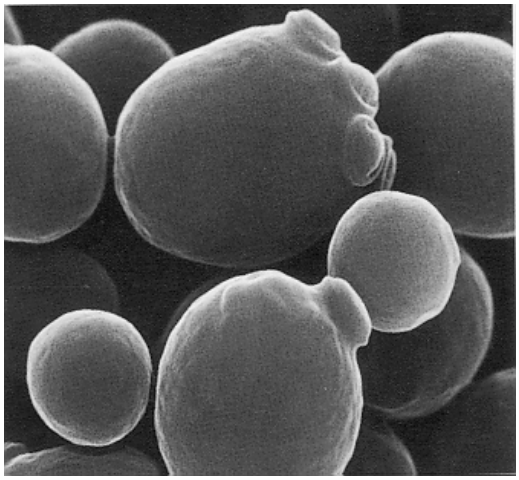




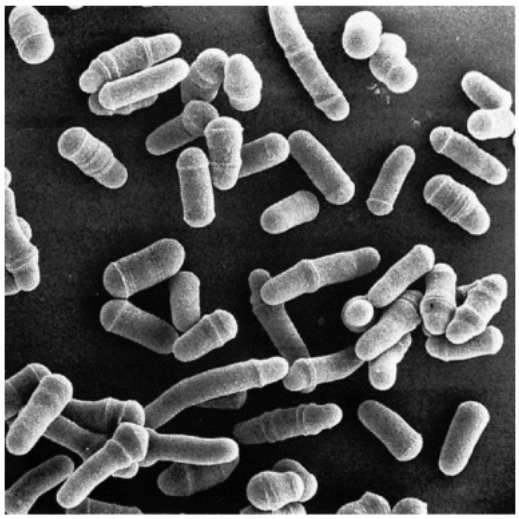
Bacteria: E. Coli



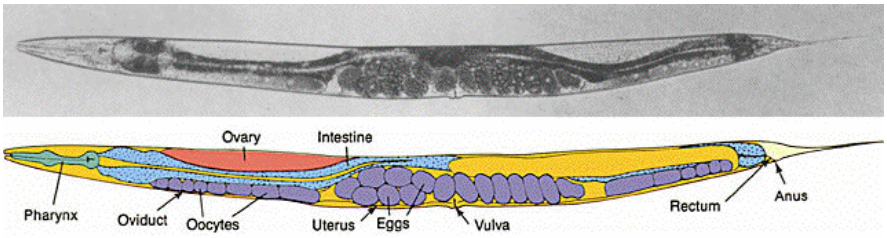
The Budding Yeast: *Saccharomyces cerevisiae*



The Fission Yeast: *Schizosaccharomyces pombe*



The Nematode: *Caenorhabditis elegans*



- **SMALL: ~ 250 µm**
- **TRANSPARENT**
- **959 CELLS**
- **300 NEURONS**
- **SHORT GENERATION TIME**
- **SIMPLE GROWTH MEDIUM**
- **SELF- FERTILIZING HERMAPHRODITE**
- **RAPID ISOLATION AND CLONING OF MULTIPLE TYPES OF MUTANT ORGANISMS**

The Fruit Fly: Drosophila Melanogaster

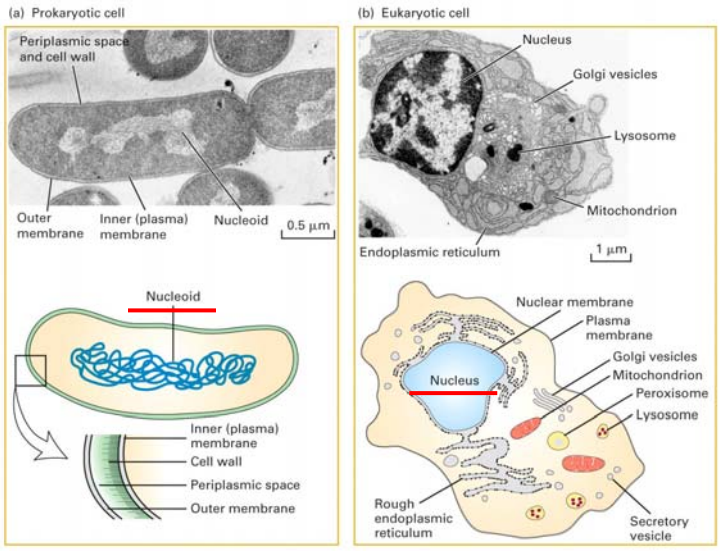
Normal

Ubx mutant

The Mouse

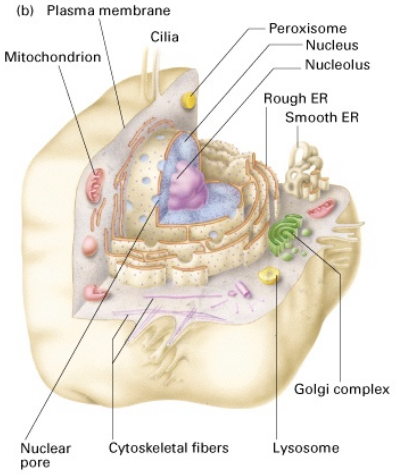
transgenic for human growth hormone

Prokaryotic and Eukaryotic Cells

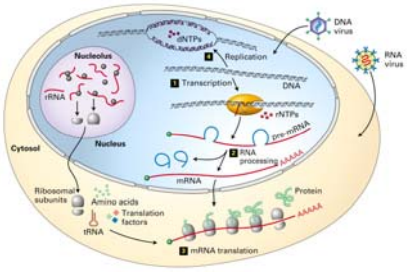


A Close Look of a Eukaryotic Cell

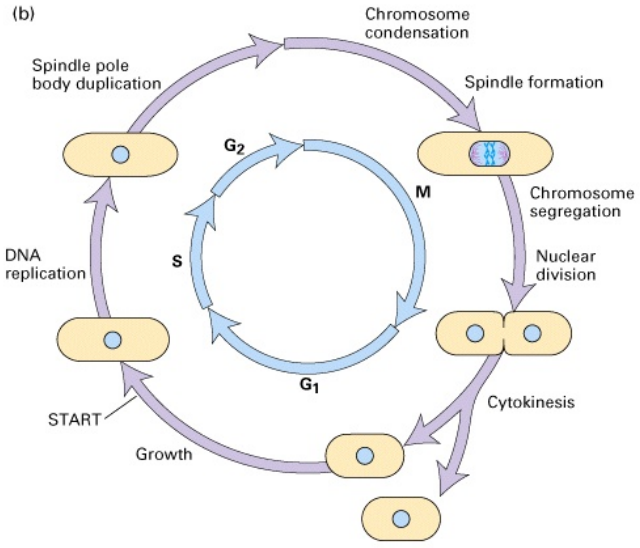
The structure:



The information flow:

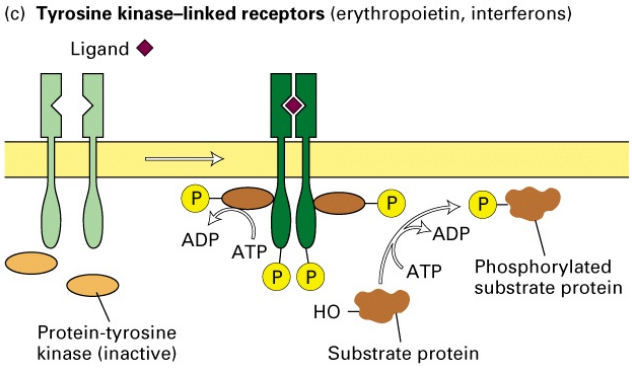


Cell Cycle

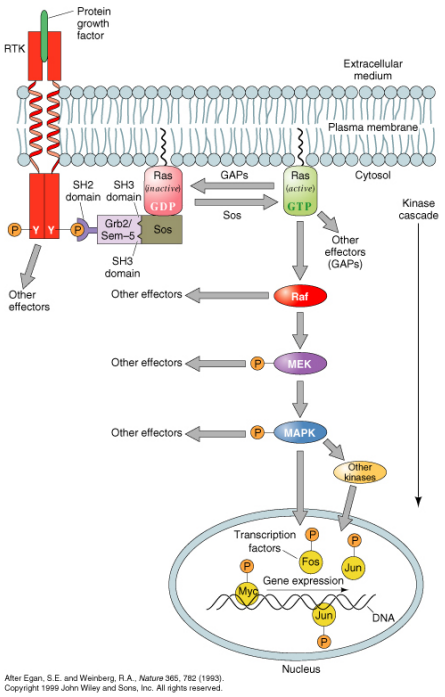


Signal Transduction

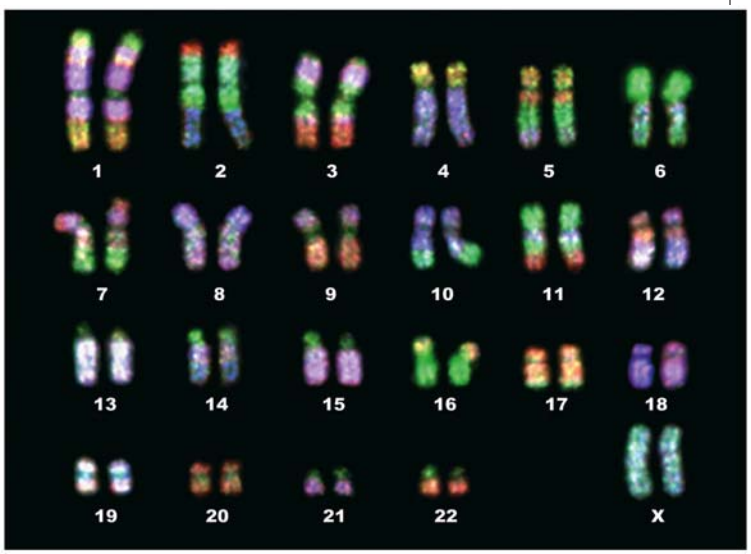
- **A variety of plasma membrane receptor proteins bind extracellular signaling molecules and transmit signals across the membrane to the cell interior**



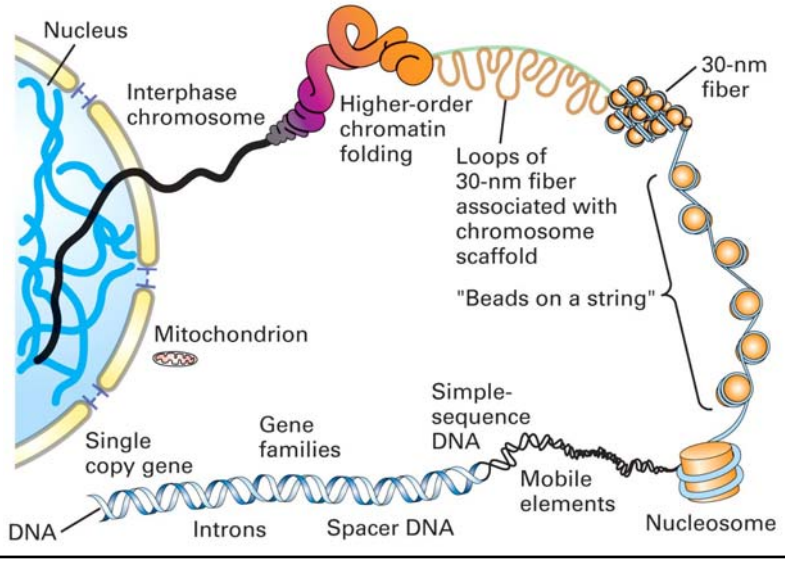
Signal Transduction Pathway



The Genome



A Multi-resolution View of the Chromosome



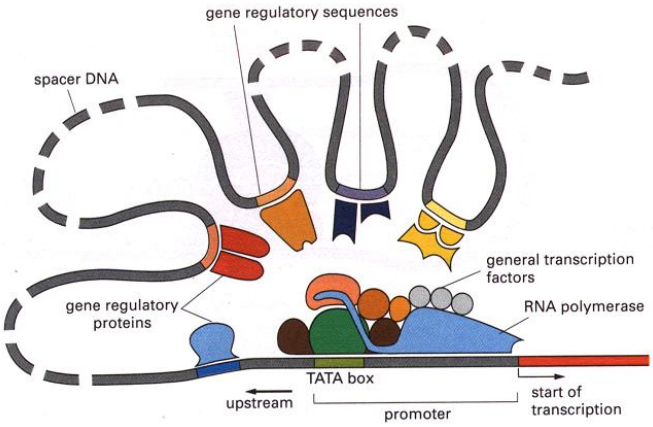
DNA Content of Representative Types of Cells

Organism	Number of base pairs (millions)	Number of encoded proteins	Number of chromosomes
PROKARYOTIC			
Mycoplasma genitalum (Bacterium)	0.58	470	1
Helicobacter pylori (Bacterium)	1.67	1590	1
Haemophilus influenza (Bacterium)	1.83	1743	1
EUKARYOTIC			
Saccharomyces cerevisiae (yeast)	12	5885	17
Drosophila melanogaster (insect)	165	13,601	4
Caenorhabditis elegans (worm)	97	19,099	6
Homo sapiens (human)	2900	30,000 TO 40,000	23
Arabidopsis thaliana (plant)	125	25,498	10

Functional Genomics

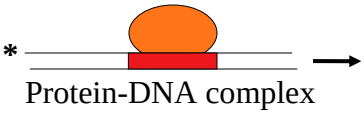
- The various **genome projects** have yielded the complete DNA sequences of many organisms.
  - E.g. human, mouse, yeast, fruitfly, etc.
  - Human: 3 billion base-pairs, 30-40 thousand genes.
- Challenge: **go from sequence to function**,
  - i.e., define the role of each gene and understand how the genome functions as a whole.

Regulatory Machinery of Gene Expression



Classical Analysis of Transcription Regulation Interactions

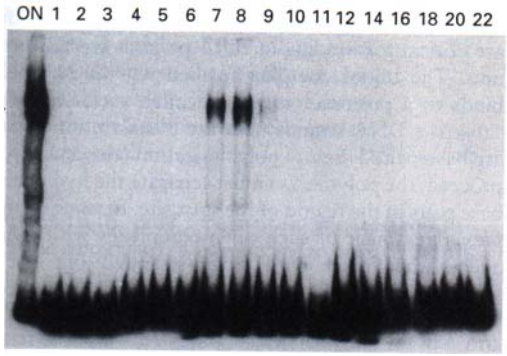
"Gel shift": electorphoretic mobility shift assay ("EMSA") for DNA-binding proteins



Protein-DNA complex

Free DNA probe

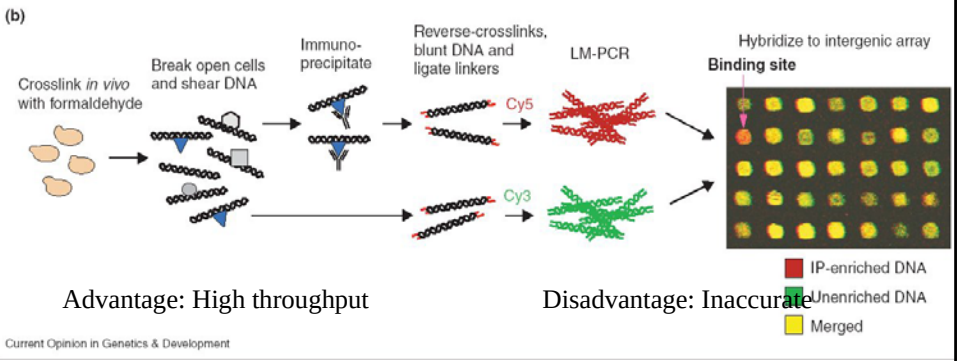
Advantage: sensitive



Disadvantage: requires stable complex; little "structural" information about which protein is binding

Modern Analysis of Transcription Regulation Interactions

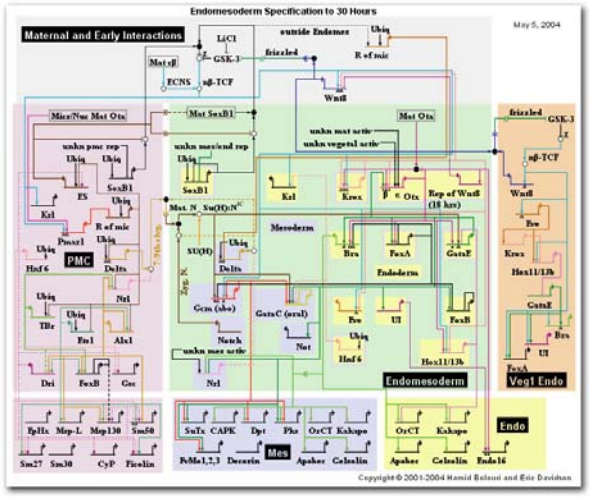
- **Genome-wide Location Analysis (ChIP-chip)**



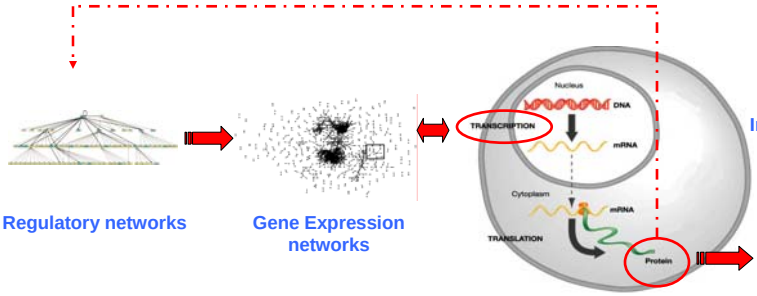
Advantage: High throughput

Disadvantage: Inaccurate

Gene Regulatory Network



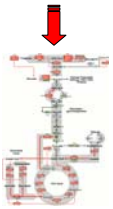
Biological Networks and Systems Biology



Regulatory networks

Gene Expression networks

Protein-protein Interaction networks



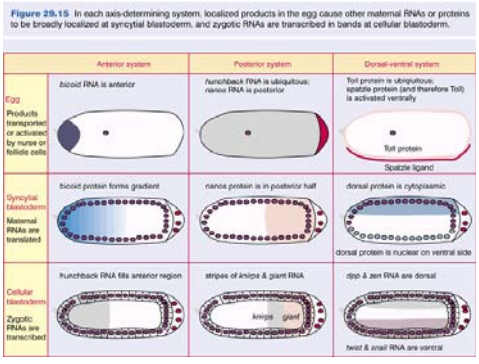
Metabolic networks

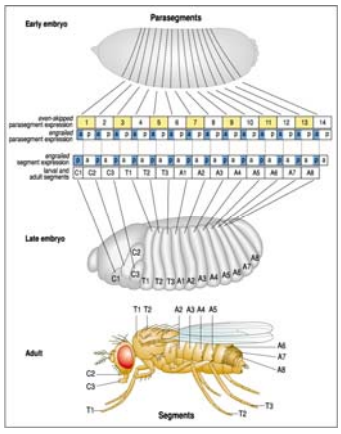
Systems Biology:

understanding cellular event under a system-level context

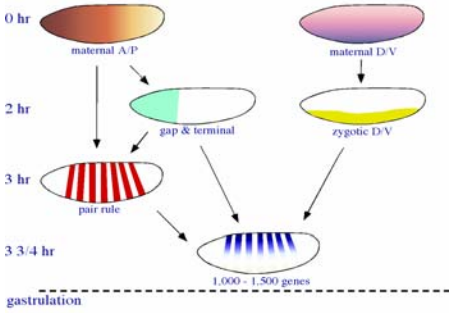
Genome + proteome + lipome + …

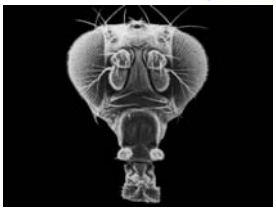
Gene Regulatory Functions in Development

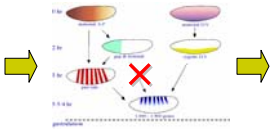




Temporal-spatial Gene Regulation and Regulatory Artifacts



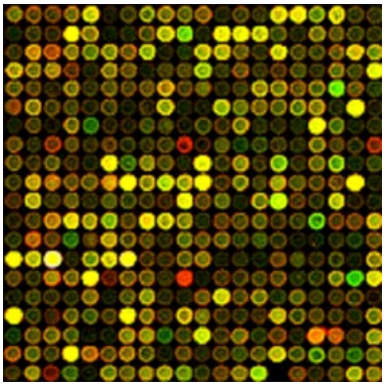




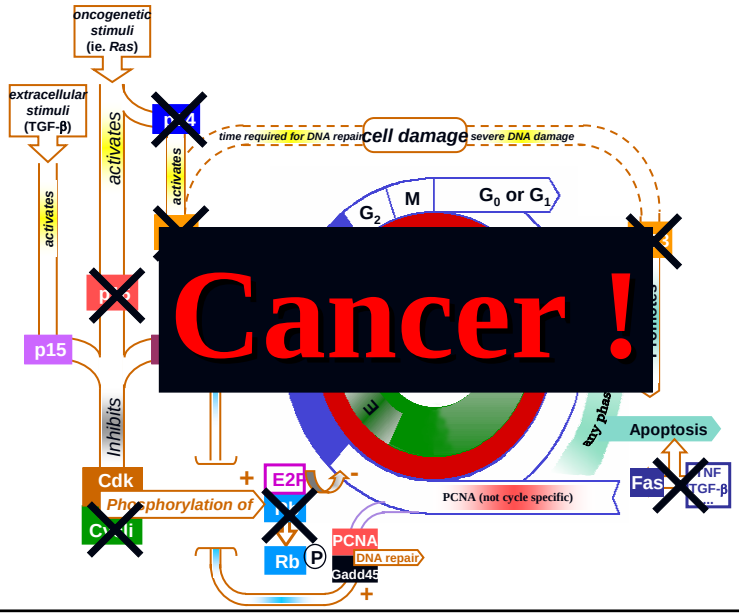
A normal fly

Hopeful monster?

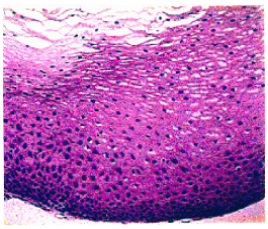
Microarray or Whole-body ISH?



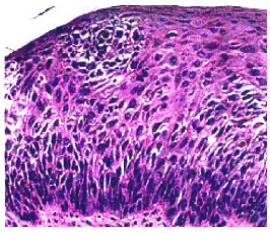
Gene Regulation and Carcinogenesis



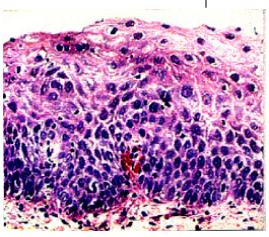
The Pathogenesis of Cancer



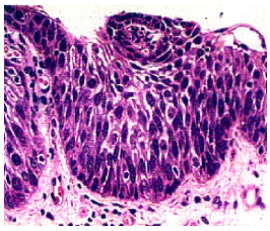
Normal



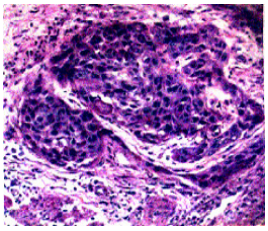
BCH



DYS



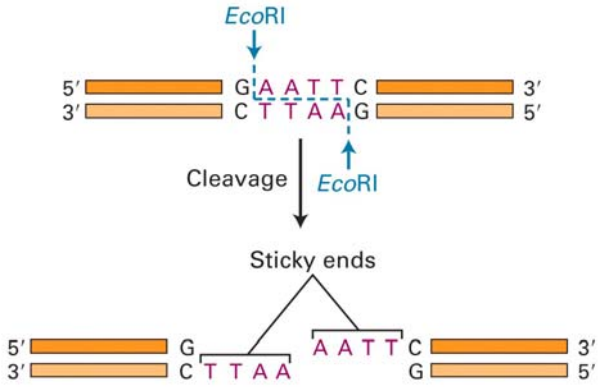
CIS



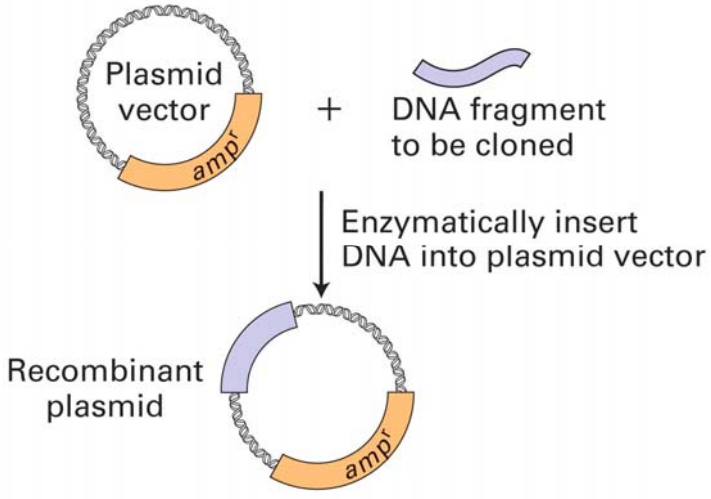
SCC

Genetic Engineering: Manipulating the Genome

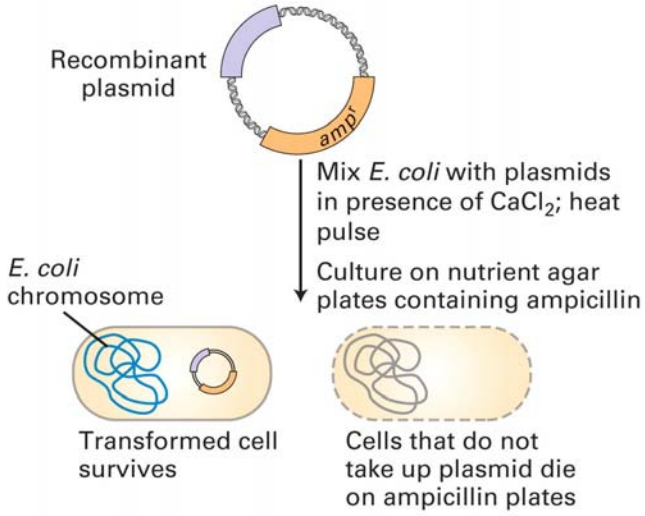
- Restriction Enzymes, naturally occurring in bacteria, that cut DNA at very specific places.



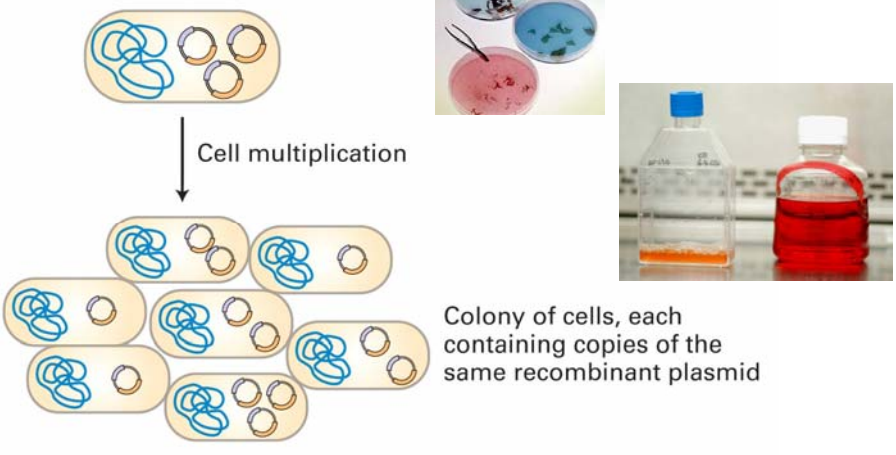
Recombinant DNA



Transformation

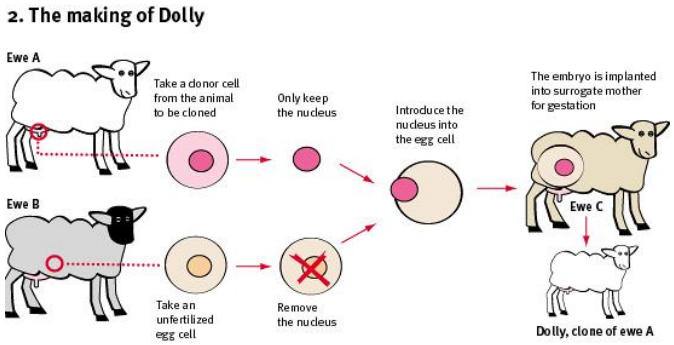


Formation of Cell Colony



How was Dolly cloned?

- Dolly is claimed to be an exact genetic replica of another sheep.



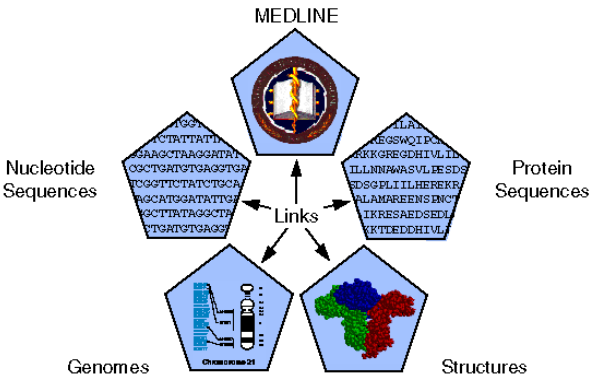
- Is it exactly "exact"?

Definitions

- Recombinant DNA: Two or more segments of DNA that have been combined by humans into a sequence that does not exist in nature.
- Cloning: Making an exact genetic copy. A **clone** is one of the exact genetic copies.
- Cloning vector: Self-replicating agents that serve as vehicles to transfer and replicate genetic material.

Software and Databases

- NCBI/NLM Databases Genbank, PubMed, PDB
  - DNA
  - Protein
  - Protein 3D
  - Literature



Entrez