# Computational Biology

Lectures

Jaap Kaandorp

Location and time

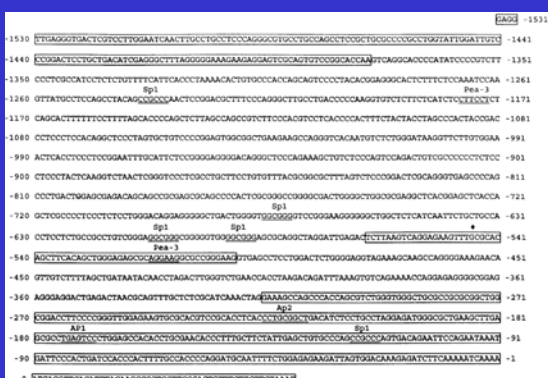
Mondays: 11.15 - 13.00, room B341

Tuesday: 14.15 -16.00, room P016

# Computational Biology

Part I blueprint of life

Jaap Kaandorp

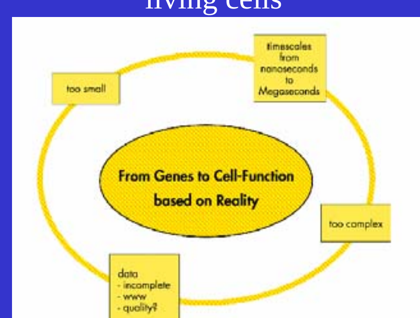


## A possible definition of Computational Biology/Bioinformatics /Systems Biology (source Wikipedia)

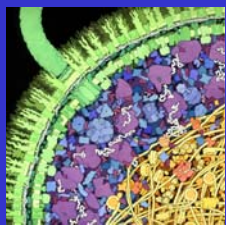
- Bioinformatics or computational biology is the use of techniques from applied mathematics, informatics, statistics, and computer science to solve biological problems. Research in computational biology often overlaps with systems biology. Major research efforts in the field include sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, and the modeling of evolution. The terms bioinformatics and computational biology are often used interchangeably, although the latter typically focuses on algorithm development and specific computational methods

## From cells to molecules

## The silicon cell: towards computing living cells



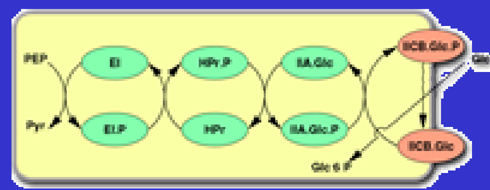
## Major challenge how to model the cell (for example E.coli) and biochemical pathways?



Detail of the E. coli cell

Overview of relevant biochemical pathways

## Case study: the Phosphoenolpyruvate dependent phosphotransferase system (PTS) responsible for the uptake of glucose in E. coli




**Diagrammatic view of the PTS pathway in E. coli**
**Problem: part of the pathway is membrane-bound**

## Computational Biology
Lab sessions
Yves Fomekong Nanfack

Location and time

Monday: 13.00 - 16.00, room P127

Wednesday: 11.00 -14.00, room P126

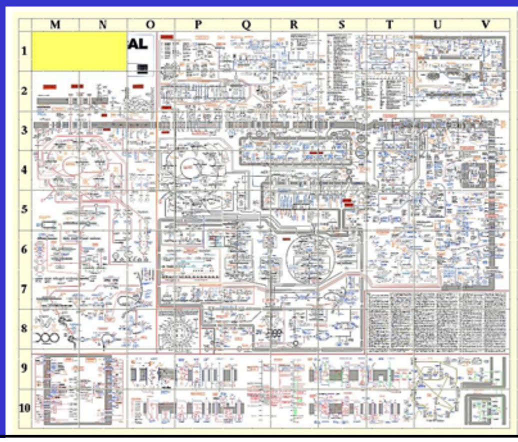
## Lab sessions
start on 5th of February (today!)

## Computational Biology
Overview I

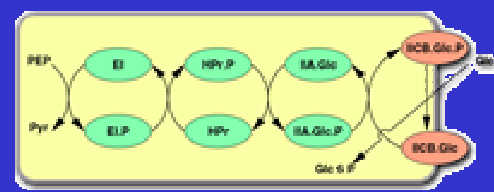
- Monday 5-2-07 Blueprint of life
- Tuesday 6-2-07 cancelled!
- Monday 12-2-07 Blueprint of life, molecular biology databases
- Tuesday 13-2-07 Sequence analysis I
- Monday 19-2-07 Sequence analysis II
- Tuesday 20-2-07 Sequence analysis III
- Monday 26-2-07 Biochemical networks

## Computational Biology
Overview II

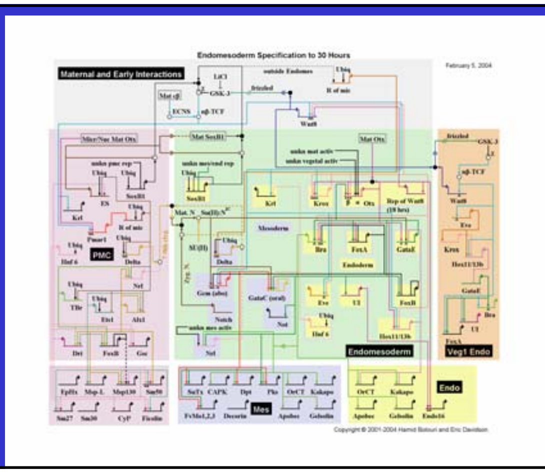
- Tuesday 27-2-07 metabolic networks I
- Monday 5-3-07 metabolic networks II
- Tuesday 6-3-07 Gene regulatory networks I
- Monday 12-3-07 Gene regulatory networks II
- Tuesday 13-3-07 Systems Biology I
- Monday 19-3-07 Systems Biology II
- Tuesday 20-3-07 Systems Biology III



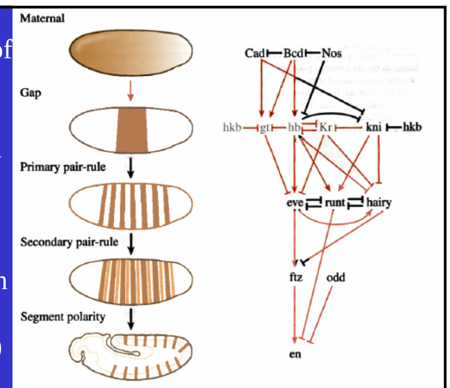
## Case study: the Phosphoenolpyruvate dependent phosphotransferase system (PTS) responsible for the uptake of glucose in E. coli




**Diagrammatic view of the PTS pathway in E. coli**
**Problem: part of the pathway is membrane-bound**




Development of morphogen gradients in time (left) and regulatory network for segment specification in *Drosophila* (Carroll,2001)




## Lab assignments

- Introduction solving simple biological problems
- Sequence analysis
- Enzyme kinetics
- Solving non-linear ODE in biological problems
- Compartimental modelling

Lab assignments, slides lectures etc
Computational Biology
Blackboard

Questions (in English) regarding the practical assignments:

Yves Fomekong Nanfack
(yvesf@science.uva.nl)

## Computational Biology course material I

- All slides etc on blackboard

## Computational Biology course material II (optional) books (in order of being advised)

- Systems Biology in Practice, E Klip et al., Wiley-VCH, Berlin, 2005
- Computational modeling of genetic and biochemical networks, J.M. Bower & H. Bolouri, MIT Press, Cambridge, 2001
- Post-genome informatics, M. Kanehisa, Oxford University Press, 2000
- (Introduction to Bioinformatics, A.M. Lesk, Oxford University Press, 2005 not advised)

## Computational Biology, to do:

- Lab assignments
- exam

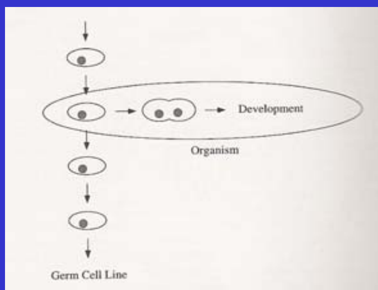
## Computational Biology

- Final grade = ½ (lab_assigments + exam)

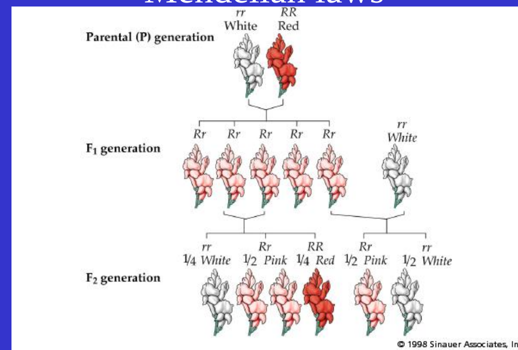
## Life is a complex sytem for information storage and processing

- Vertical information transfer: generation to generation
- Horizontal information transfer: ontogenesis

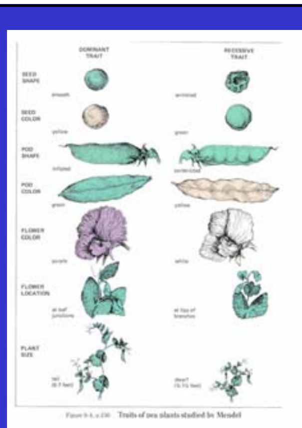
## Life is a complex sytem for information storage and processing



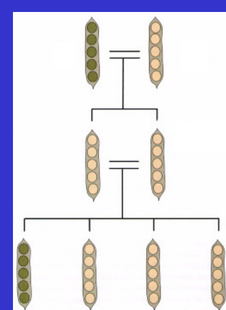

## Mendelian laws



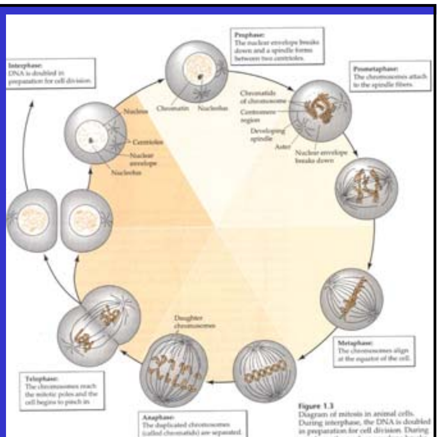

## Mendelian laws II



## Mendelian laws II



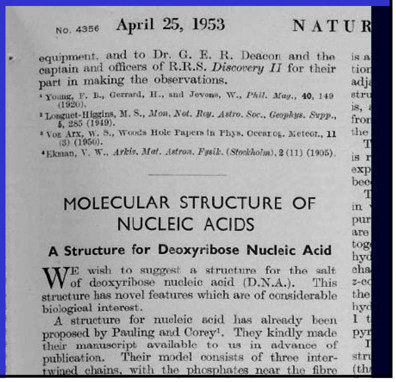
## Chromosomes and cell division in animals



## Genome and gene

- Genome: unit of information transmission by DNA replication
- Gene : unit of information expression by transcription to RNA or translation to protein

# Double helix model of DNA, Watson Crick 1953



No. 4356 April 25, 1953 NATUR

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

**A Structure for Deoxyribose Nucleic Acid**

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre

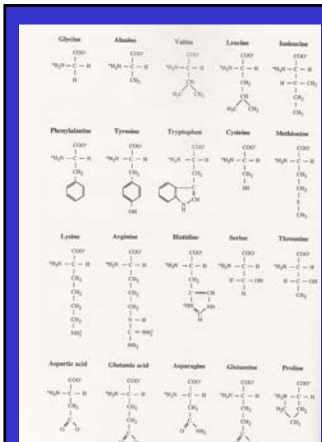
# Nucleotides

- A Adenine
- G guanine
- C Cytosine
- T Thymine

# DNA double helix

# Amino acids

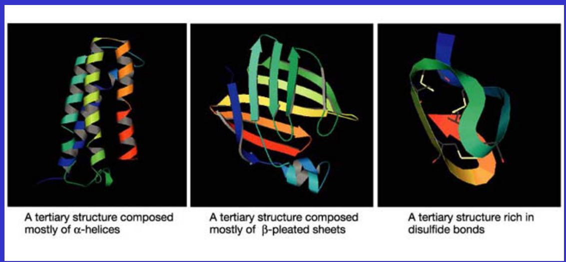
| | | |
|---|---|---|
| Ala | A | Alanine |
| Arg | R | Arginine |
| Asn | N | Asparagine |
| Asp | D | Aspartic acid |
| Cys | C | Cysteine |
| Gln | Q | Glutamine |
| Glu | E | Glutamic acid |
| Gly | G | Glycine |
| His | H | Histidine |
| Ile | I | Isoleucine |
| Leu | L | Leucine |
| Lys | K | Lysine |
| Met | M | Methionine |
| Phe | F | Phenylalanine |
| Pro | P | Proline |
| Ser | S | Serine |
| Thr | T | Threonine |
| Trp | W | Tryptophan |
| Tyr | Y | Tyrosine |
| Val | V | Valine |
| Asx | B | Asn or Asp |
| Glx | Z | Gln or Glu |
| Sec | U | Selenocysteine |
| Unk | X | Unknown |



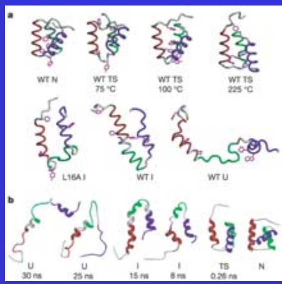
# Amino acids

# 3D protein structure protein: example myoglobin

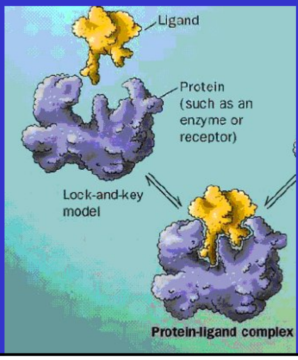
## 3D protein structure protein: alpha and beta helices



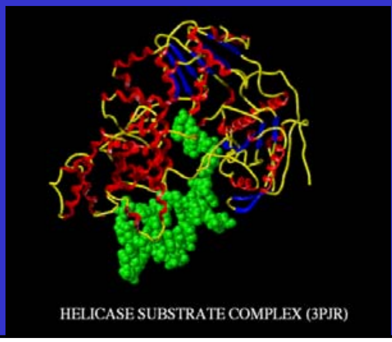

## Unfolding and folding of the engrailed homeodomain protein (Molecular Dynamics simulations, Nature 421:863-7, 2003)



## Enzyme substrate




## Enzyme substrate recognition



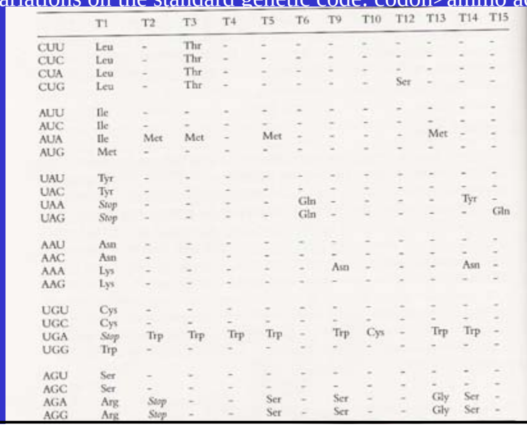

## Standard genetic code: codon->amino acid

| | U | | C | | A | | G | | |
|---|---|---|---|---|---|---|---|---|---|
| U | UUU | Phe | UCU | Ser | UAU | Tyr | UGU | Cys | U |
| | UUC | Phe | UCC | Ser | UAC | Tyr | UGC | Cys | C |
| | UUA | Leu | UCA | Ser | UAA | *Stop* | UGA | *Stop* | A |
| | UUG | Leu | UCG | Ser | UAG | *Stop* | UGG | Trp | G |
| C | CUU | Leu | CCU | Pro | CAU | His | CGU | Arg | U |
| | CUC | Leu | CCC | Pro | CAC | His | CGC | Arg | C |
| | CUA | Leu | CCA | Pro | CAA | Gln | CGA | Arg | A |
| | CUG | Leu | CCG | Pro | CAG | Gln | CGG | Arg | G |
| A | AUU | Ile | ACU | Thr | AAU | Asn | AGU | Ser | U |
| | AUC | Ile | ACC | Thr | AAC | Asn | AGC | Ser | C |
| | AUA | Ile | ACA | Thr | AAA | Lys | AGA | Arg | A |
| | AUG | Met, *Start* | ACG | Thr | AAG | Lys | AGG | Arg | G |
| G | GUU | Val | GCU | Ala | GAU | Asp | GGU | Gly | U |
| | GUC | Val | GCC | Ala | GAC | Asp | GGC | Gly | C |
| | GUA | Val | GCA | Ala | GAA | Glu | GGA | Gly | A |
| | GUG | Val | GCG | Ala | GAG | Glu | GGG | Gly | G |

## Nucleotides:

| DNA | RNA |
|---|---|
| • A Adenine | A Adenine |
| • G guanine | G Guanine |
| • C Cytosine | C Cytosine |
| • T Thymine | U Uracil |

## Variations on the standard genetic code: codon>amino acid I

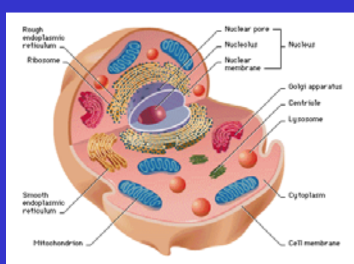


| | T1 | T2 | T3 | T4 | T5 | T6 | T9 | T10 | T12 | T13 | T14 | T15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CUU | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUC | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUA | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUG | Leu | - | Thr | - | - | - | - | - | Ser | - | - | - |
| AUU | Ile | - | - | - | - | - | - | - | - | - | - | - |
| AUC | Ile | - | - | - | - | - | - | - | - | - | - | - |
| AUA | Ile | Met | Met | - | Met | - | - | - | - | Met | - | - |
| AUG | Met | - | - | - | - | - | - | - | - | - | - | - |
| UAU | Tyr | - | - | - | - | - | - | - | - | - | - | - |
| UAC | Tyr | - | - | - | - | - | - | - | - | - | - | - |
| UAA | *Stop* | - | - | - | - | Gln | - | - | - | - | Tyr | - |
| UAG | *Stop* | - | - | - | - | Gln | - | - | - | - | - | Gln |
| AAU | Asn | - | - | - | - | - | - | - | - | - | - | - |
| AAC | Asn | - | - | - | - | - | - | - | - | - | - | - |
| AAA | Lys | - | - | - | - | - | Asn | - | - | - | Asn | - |
| AAG | Lys | - | - | - | - | - | - | - | - | - | - | - |
| UGU | Cys | - | - | - | - | - | - | - | - | - | - | - |
| UGC | Cys | - | - | - | - | - | - | - | - | - | - | - |
| UGA | *Stop* | Trp | Trp | Trp | Trp | - | Trp | Cys | - | Trp | Trp | - |
| UGG | Trp | - | - | - | - | - | - | - | - | - | - | - |
| AGU | Ser | - | - | - | - | - | - | - | - | - | - | - |
| AGC | Ser | - | - | - | - | - | - | - | - | - | - | - |
| AGA | Arg | *Stop* | - | - | Ser | - | Ser | - | - | Gly | Ser | - |
| AGG | Arg | *Stop* | - | - | Ser | - | Ser | - | - | Gly | Ser | - |

## Variations on the standard genetic code: codon>amino acid II

- T1 standard code
- T2 vertebrate mitochondrial code
- T3 yeast mitochondrial code
- T4 mould, protozoan and coelenterate mitochondrial code; mycoplasma and spiroplasma code
- T5 invertebrate mitochondrial code
- T6 ciliate, dasycladacean and hexamita nuclear code
- T9 echinoderm mitochondrial code
- T10 euplotid code
- T12 alternative yeast code
- T13 ascidian mitochondrial code
- T14 flatworm mitochondrial code
- T15 blepharisma nuclear code

## Animal cell with mitochondria



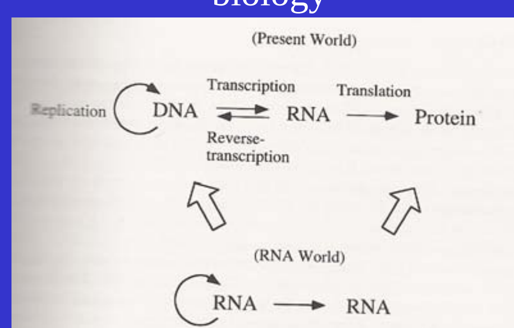

## 3 types of RNA

- Messenger RNA: mRNA
- Transfer RNA: tRNA
- Ribosomal RNA: rRNA

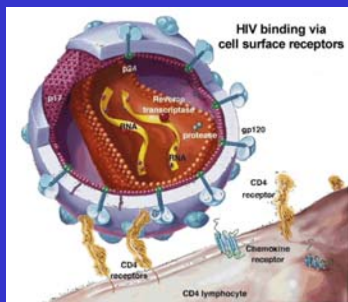
## Nucleotides: DNA (deoxribose backbone) RNA (ribose)

| | |
|---|---|
| • A Adenine | A Adenine |
| • G guanine | G Guanine |
| • C Cytosine | C Cytosine |
| • T Thymine | U Uracil |

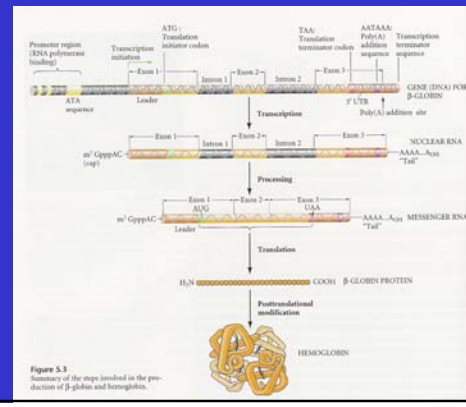
## Central dogma of molecular biology



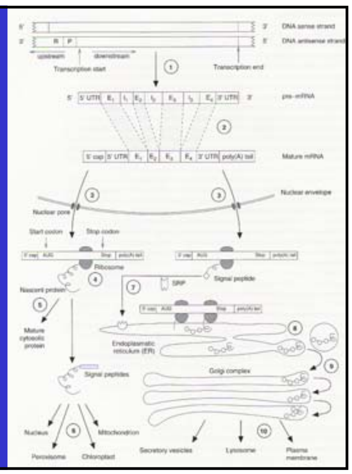
# RNA virus: HIV



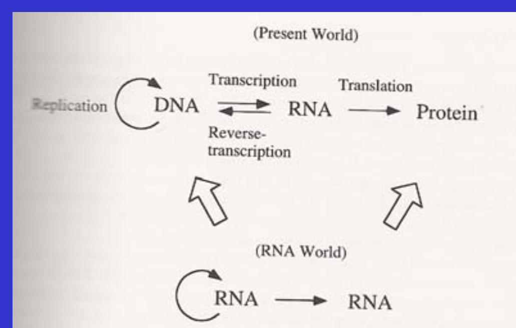

# mRNA splicing (in animals)




# Transcription and processing of mRNA (in eukaryotes, for example animals)



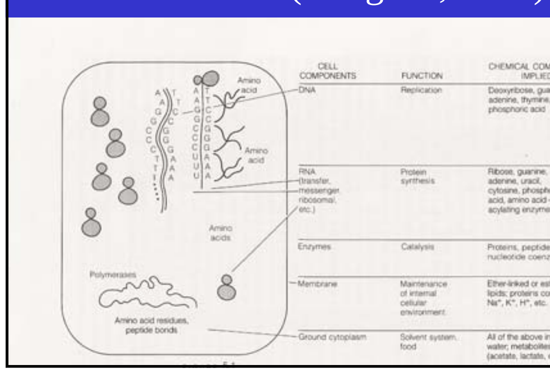

# RNA world




# Definition of life

- Self-maintaining metabolic system capable of reproduction

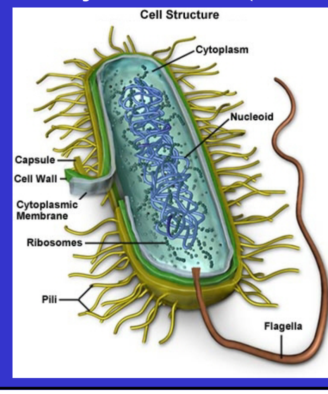
# Minimal cell (Margulis, 1993)



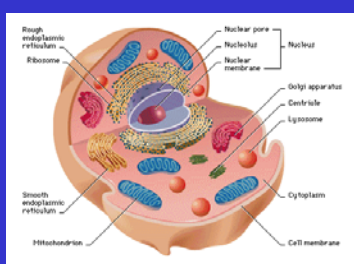
## Minimal cell II (Margulis, 1993)

- Cells, minimal autopoietic systems, are membrane entities, the smallest capable of self-reproduction. Even the tiniest cells today contain genes in the form of DNA molecules, single or in several copies, an protein-synthesizing machinery composed of several types of RNA, and many proteins. All cells contain ribosomes, bodies about 0.02$\mu$ m in diameter composed of at least three kinds of RNA and about 50 different proteins

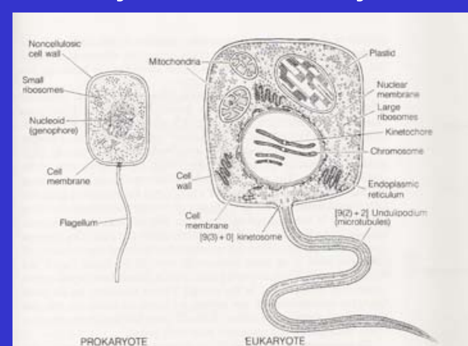
## Prokaryotic cell (E. coli)




## Eukaryotic cell (animal cell)



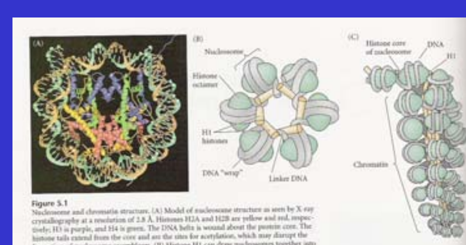

## Prokaryotic and Eukaryotic cell



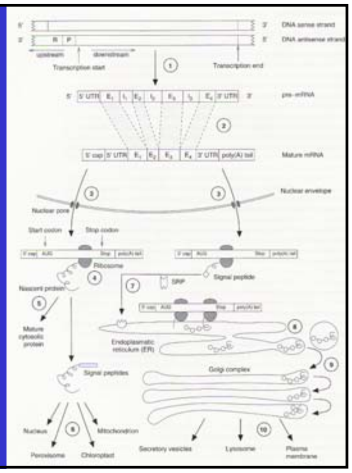

## Differences prokaryotes and eukaryotes

- No distinct compartments in prokaryotes: no nucleus, no organnelles (chloroplasts, mitochondria, endoplasmatic reticulum, Golgi complex, centriole, lysosome, peroxisome, etc.)
- Genome is not packed into a nucleus, no chromatin, no chromosomes, in prokaryotes
- Cell is smaller (in general) in prokaryotes (1-10 $\mu$ length; eukaryotes 10-100 $\mu$ length)
- Prokaryotes mostly single cellular, many eukaryotes are multi-cellular

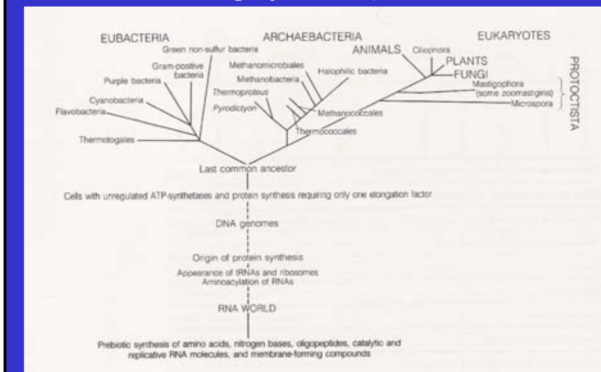
## Chromatin in eukaryotes



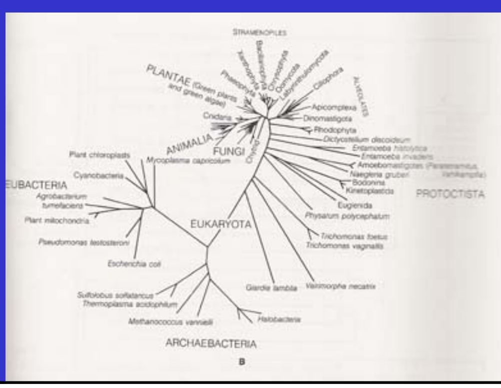
After transcription and processing of mRNA (in eukaryotes, for example animals)



Tree of life (Margulis, 1993; based on RNA polymerases)

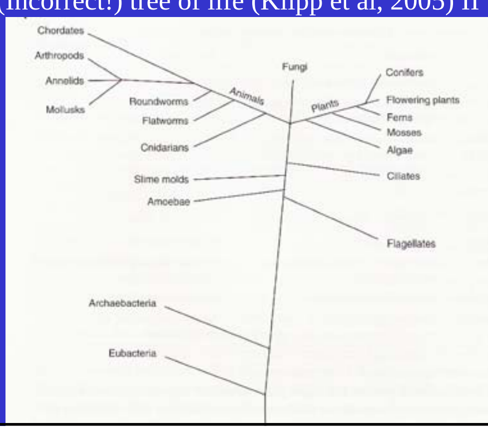



Tree of life (Margulis, 1993; based on ribosomal RNA sequences)

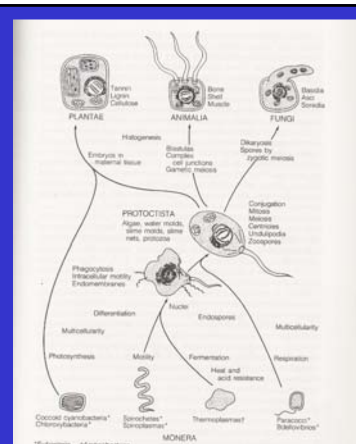



Example Rhodophyta (red seaweeds)

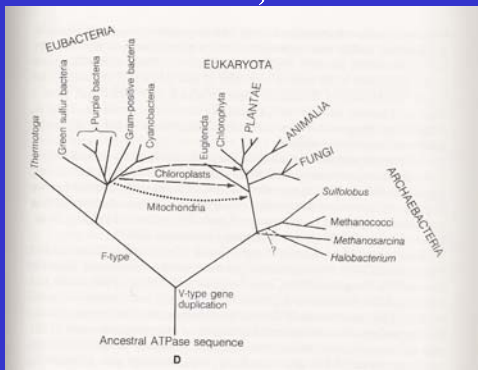
(Incorrect!) tree of life (Klipp et al, 2005) II



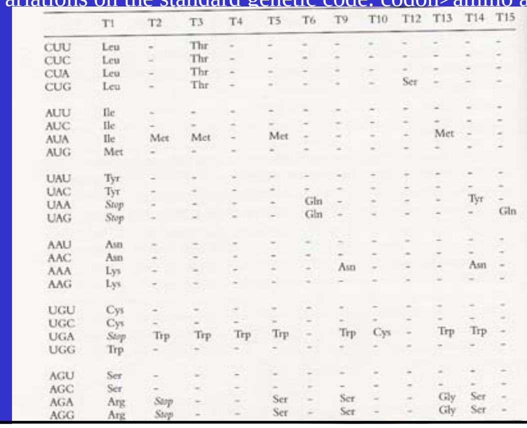

Tree of life symbiogenesis (Margulis, 1993)



## Diagram tree of life symbiogenesis (Margulis, 1993) II




## Variations on the standard genetic code: codon>amino acid I

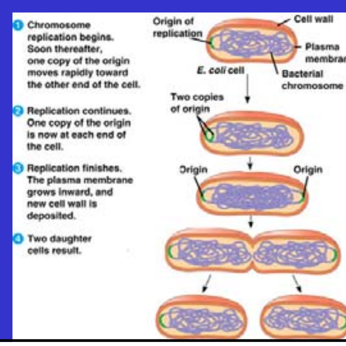


| | T1 | T2 | T3 | T4 | T5 | T6 | T9 | T10 | T12 | T13 | T14 | T15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CUU | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUC | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUA | Leu | - | Thr | - | - | - | - | - | - | - | - | - |
| CUG | Leu | - | Thr | - | - | - | - | - | Ser | - | - | - |
| AUU | Ile | - | - | - | - | - | - | - | - | - | - | - |
| AUC | Ile | - | - | - | - | - | - | - | - | - | - | - |
| AUA | Ile | Met | Met | - | Met | - | - | - | - | Met | - | - |
| AUG | Met | - | - | - | - | - | - | - | - | - | - | - |
| UAU | Tyr | - | - | - | - | - | - | - | - | - | - | - |
| UAC | Tyr | - | - | - | - | - | - | - | - | - | - | - |
| UAA | *Stop* | - | - | - | - | Gln | - | - | - | - | Tyr | - |
| UAG | *Stop* | - | - | - | - | Gln | - | - | - | - | - | Gln |
| AAU | Asn | - | - | - | - | - | - | - | - | - | - | - |
| AAC | Asn | - | - | - | - | - | - | - | - | - | - | - |
| AAA | Lys | - | - | - | - | - | Asn | - | - | - | Asn | - |
| AAG | Lys | - | - | - | - | - | - | - | - | - | - | - |
| UGU | Cys | - | - | - | - | - | - | - | - | - | - | - |
| UGC | Cys | - | - | - | - | - | - | - | - | - | - | - |
| UGA | *Stop* | Trp | Trp | Trp | Trp | - | Trp | Cys | - | Trp | Trp | - |
| UGG | Trp | - | - | - | - | - | - | - | - | - | - | - |
| AGU | Ser | - | - | - | - | - | - | - | - | - | - | - |
| AGC | Ser | - | - | - | - | - | - | - | - | - | - | - |
| AGA | Arg | *Stop* | - | - | Ser | - | Ser | - | - | Gly | Ser | - |
| AGG | Arg | *Stop* | - | - | Ser | - | Ser | - | - | Gly | Ser | - |

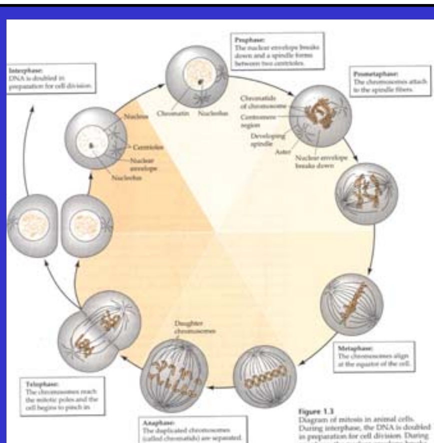
## Variations on the standard genetic code: codon>amino acid II

- T1 standard code
- T2 vertebrate mitochondrial code
- T3 yeast mitochondrial code
- T4 mould, protozoan and coelenterate mitochondrial code; mycoplasma and spiroplasma code
- T5 invertebrate mitochondrial code
- T6 ciliate, dasycladacean and hexamita nuclear code
- T9 echinoderm mitochondrial code
- T10 euplotid code
- T12 alternative yeast code
- T13 ascidian mitochondrial code
- T14 flatworm mitochondrial code
- T15 blepharisma nuclear code

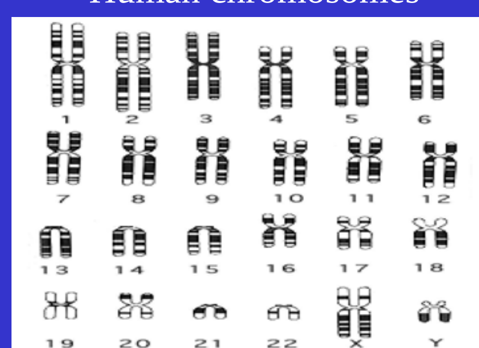
## Cell division in bacteria (note there are actually no chromosomes in bacteria!)



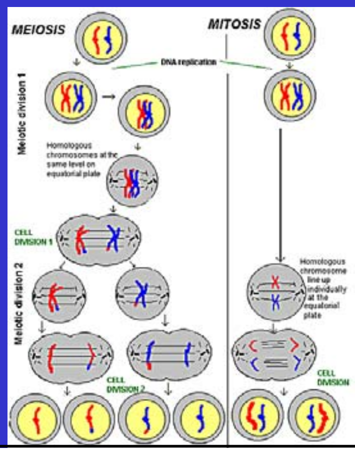

## Cell division in animals (eukaryotes) and cell cycle



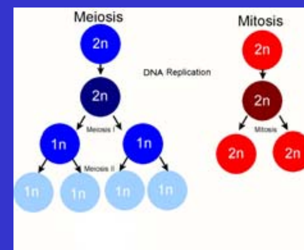

## Human chromosomes



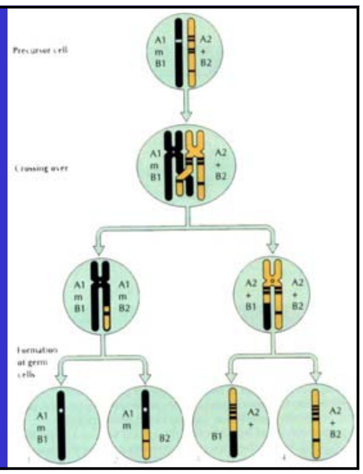
Mitosis and meiosis in animals (eukaryotes) I



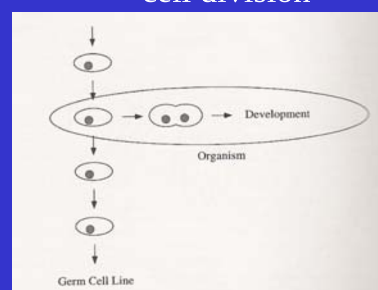

Mitosis and meiosis in animals (eukaryotes) II




Recombination of chromosomes (for example in animals)




Life is a complex system for information storage and processing + cell division




Computational Biology

End Part I blueprint of life

Jaap Kaandorp

12 ECTS course bioinformatics at Academic Medical Center Amsterdam

- Block 2b
- 12 ECTS
- See website programme bioinformatics MGC http://www.science.uva.nl/bioinformatics for details
- Registration via email contact Dr A.H. van Kampen (a.h.vankampen@amc.uva.nl)
- Number of students limited