Introduction to Molecular Biology

Part 1
Microscopic biology began in 1665

Robert Hooke (1635-1703) discovered organisms are made up of cells
Cells

• Fundamental working units of all living systems
• Every organism is composed of one or two different types of cells
  – Prokaryotic cells
  – Eukaryotic cells
• Prokaryotes and Eukaryotes are descended from the same primitive cell
• All extant cells are the result of 3.5 billion years of evolution
Cells
Cells

- A cell is the smallest structural unit of an organism that is capable of independent functioning.
- All cells have some common features.
Cell Cycles: Born, Eat, Replicate, Die
Cell: Contents...

• Chemical Composition (by weight)
  – 70% water
  – 7% small molecules
    • Salts, lipids, amino acids, nucleotides
  – 23% macromolecules
    • Proteins, polysaccharides, lipids

• Biochemical (Metabolic) Pathways

• Translation of mRNA into Proteins
Cells: Prokaryotes & Eukaryotes
Prokaryotes & Eukaryotes
Charles Darwin: Tree of Life

“I think case must be that one generation should have as many living as now. To do this and to have as many species in same genus (as is) requires extinction. Thus between A + B the immense gap of relation. C + B the finest gradation. B+D rather greater distinction. Thus genera would be formed. Bearing relation" (next page begins) "to ancient types with several extinct forms”

-: Charles Darwin, 1837
## Prokaryotes & Eukaryotes

<table>
<thead>
<tr>
<th>Prokaryotes</th>
<th>Eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>➢ Single Cell</td>
<td>➢ Single or multi cell</td>
</tr>
<tr>
<td>➢ No nucleus</td>
<td>➢ Nucleus</td>
</tr>
<tr>
<td>➢ No organelles</td>
<td>➢ Organelles</td>
</tr>
<tr>
<td>➢ One piece of Circular DNA</td>
<td>➢ Chromosomes</td>
</tr>
<tr>
<td>➢ No mRNA post-transcriptional modification</td>
<td>➢ Exons/Introns splicing</td>
</tr>
</tbody>
</table>
Prokaryotes & Eukaryotes
Structural Differences

Prokaryotes

- Eubacteria (blue green algae) and archaeobacteria
- Only one type of membrane
  Plasma membrane forms the boundary of the cell
- The smallest cells known are
  bacteria (E. Coli cell, 3 \times 10^6 protein molecules, 1000-2000 polypeptide species)

Eukaryotes

- Plants, animals, Protista, fungi
- Complex systems of internal
  membranes forms organelle and compartments
- Volume of cell is several hundred
  times larger (Hela cell, 5 \times 10^9 protein molecules, 5000-10000 polypeptide species)
## Prokaryotes & Eukaryotes

### Chromosomal Differences

<table>
<thead>
<tr>
<th>Prokaryotes</th>
<th>Eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The genome of E. Coli contains $4 \times 10^6$ base pairs</td>
<td>The genome of yeast contains $1.35 \times 10^7$ base pairs</td>
</tr>
<tr>
<td>&gt;90% of DNA encode protein</td>
<td>A small fraction of the DNA encodes protein (many repeats of non-coding sequences)</td>
</tr>
<tr>
<td>Lacks a membrane bound nucleus</td>
<td>All chromosomes are contained in a membrane bound nucleus (DNA is divided between one or more chromosomes)</td>
</tr>
<tr>
<td>Circular DNA and supercoiled domain</td>
<td>A set of five histones DNA packaging and gene expression regulation</td>
</tr>
<tr>
<td>Histones are unknown</td>
<td></td>
</tr>
</tbody>
</table>
Signaling Pathways Control Gene Activity

• Instead of having brains, cells make decisions through complex networks of chemical reactions, called pathways

  – Synthesize new materials

  – Break other materials down for spare parts

  – Signal to eat or die
Cells: Information & Machinery

• Cells store all information to replicate
  – Human genome is around $3 \times 10^9$ base pairs long
  – Almost every cell in a human body contains same set of genes
  – But not all genes are used/expressed by all cells

• Machinery
  – Collect and manufacture components
  – Carry out replication
  – Kick-start its new offspring

(A cell is like a car factory!)
The Human Genome Project

- **1986** Leroy Hood: Developed automated sequencing mechanism
- **1986** Human Genome Initiative announced
- **1990** The 15 year Human Genome project is launched by congress
- **1995** John Craig Venter: First bacterial genomes sequenced
- **1997** E. Coli sequenced
- **1996** First eukaryotic genome-yeast-sequenced
The Human Genome Project

• **2000** J. Craig Ventnor (Celera) and Francis Collins (IHGSC) announce sequencing the complete human genome

• **2001** International Human Genome Sequencing: First draft of the sequence of the human genome published

• False start...until **2003**
Organization of Life: Overview

- Nucleus = library
- Chromosomes = bookshelves
- Genes = books
- Almost every cell in an organism contains the same libraries and the same set of books
- Books represent all the information (DNA) that every cell in the body needs so it can grow and carry out its various functions
**Terminology**

- **Genome**: An organism’s genetic material
- **Gene**: Discrete unit of hereditary information located on the chromosomes and consisting of DNA
- **Genotype**: The genetic makeup of an organism
- **Phenotype**: The physical expressed traits of an organism
- **Nucleic Acid**: Biological molecules (DNA & RNA) that allow organisms to reproduce
Genotype & Phenotype

• Genes are inherited and are expressed
  – **genotype** (genetic makeup)
  – **phenotype** (physical expression)

• On the left, is the eye’s phenotypes of green and black eye genes.
Genotype & Phenotype

- Genes are like recipes (genotype)
- Think of a recipe for a cake...
- Only partly guarantee the end result (phenotype)
- Environment plays a crucial role
Terminology...

• The **genome** is an organism’s complete DNA. A bacteria contains about 600,000 DNA base pairs. Human and mouse genomes have 3 billion base pairs.

• Human genome has 24 distinct chromosomes. Each chromosome contains many **genes**.

• **Genes** are the basic functional units of heredity. Specific sequences of DNA bases that encode instructions on how to make **proteins**.

• **Proteins** make up the cellular structure. Large complex molecules made up of smaller subunits called **amino acids**.
Life: 3 Critical Molecules

• **DNA**
  Holds information on how cell works

• **RNA**
  Acts to transfer short pieces of information to different parts of a cell
  Provides templates to synthesize proteins

• **Proteins**
  Form enzymes that send signals to other cells and regulate gene activity
  Form body’s major components (e.g. hair, skin, etc.)
DNA: The Double Helix

• **1952-1953** James D. Watson and Francis H. C. Crick deduced the double helical structure of DNA
DNA: The Code of Life

- The structure and the four genomic letters code for all living organisms
- Adenine, Guanine, Thymine, and Cytosine which pair A-T and C-G on complimentary strands.
DNA: The Code of Life

• DNA has a double helix structure which composed of
  – sugar molecule
  – phosphate group
  – and a base (A,C,G,T)

• DNA always reads from 5’ end to 3’ end for transcription replication
  5’ ATTTAGGCC 3’
  3’ TAAATCCGG 5’
DNA & RNA: Flow of Information

Replication

DNA can replicate.

Transcription

RNA

Translation

Protein

Information coded in the sequence of base pairs in DNA is passed to molecules of RNA.

Information in RNA is passed to proteins. It never passes from proteins to nucleic acids.
DNA & RNA: Flow of Information

Replication

DNA can replicate.

Transcription

Information coded in the sequence of base pairs in DNA is passed to molecules of RNA.

Translation

Information in RNA is passed to proteins. It never passes from proteins to nucleic acids.

Protein

*a*ka “The Central Dogma”!!
Francis Crick

“The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein of nucleic acid.”

DNA to RNA to Protein

A gene is expressed in two steps

1. **Transcription:** RNA Synthesis
2. **Translation:** Protein Synthesis
The Code Book

- DNA, RNA, and Proteins are examples of strings written in either the four-letter nucleotide of DNA and RNA (A C G T/U)

- or the twenty-letter amino acid of proteins. Each amino acid is coded by 3 nucleotides called codons
Sheldon’s Favorite Amino Acid...
The Code Book
DNA & RNA

- DNA = Deoxyribonucleic acid
- RNA = Ribonucleic acid
- They are almost the same...
- There is no T base in RNA
- A similar base U takes its place
- An oxygen atom is added to the sugar component of RNA
References

• Adapted from slides posted at the web site of the above book.