The study of chemical and physical structure of macromolecule was first referred as Molecular Biology by William Astbury.

The discipline is enormously complex and crosses traditional boundaries between genetics, biochemistry, cell biology and biophysics.
**Introduction to Molecular Biology**

<table>
<thead>
<tr>
<th>Major events in the genetics century</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850</td>
</tr>
<tr>
<td>1865 Genes are particulate factors</td>
</tr>
<tr>
<td>1871 Discovery of nucleic acids</td>
</tr>
<tr>
<td>1903 Chromosomes are hereditary units</td>
</tr>
<tr>
<td>1910 Genes lie on chromosomes</td>
</tr>
<tr>
<td>1913 Chromosomes are linear arrays of genes</td>
</tr>
<tr>
<td>1927 Mutations are physical changes in genes</td>
</tr>
<tr>
<td>1931 Recombination occurs by crossing over</td>
</tr>
<tr>
<td>1944 DNA is the genetic material</td>
</tr>
<tr>
<td>1945 A gene codes for protein</td>
</tr>
<tr>
<td>1951 First protein sequence</td>
</tr>
<tr>
<td>1953 DNA is a double helix</td>
</tr>
<tr>
<td>1958 DNA replicates semiconservatively</td>
</tr>
<tr>
<td>1961 Genetic code is triplet</td>
</tr>
<tr>
<td>1977 Eukaryotic genes are interrupted</td>
</tr>
<tr>
<td>1977 DNA can be sequenced</td>
</tr>
<tr>
<td>1995 Bacterial genomes sequenced</td>
</tr>
<tr>
<td>2001 Human genome sequenced</td>
</tr>
</tbody>
</table>

**Mendel’s Pea experiment**

Molecular Biology of Gene Watson et al ©
Nucleic acid carry genetic specificity

Frederick Griffith
1928
Streptococcus pneumonia

Transformation of bacteria

<table>
<thead>
<tr>
<th>Pneumococcus types</th>
<th>Injection of cells</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capsule smooth (S)</td>
<td>Living S</td>
<td>Dies</td>
</tr>
<tr>
<td>appearance</td>
<td>Heat-killed S</td>
<td>Lives</td>
</tr>
<tr>
<td>No capsule</td>
<td>Living R</td>
<td>Lives</td>
</tr>
<tr>
<td>rough (R) appearance</td>
<td>Heat-killed S</td>
<td>Dies</td>
</tr>
<tr>
<td></td>
<td>Living R</td>
<td>Lives</td>
</tr>
</tbody>
</table>

The transforming principle is DNA

Mouse injected with heat-killed S and living R bacteria
Living S bacteria recovered from dead mouse
Extract DNA
R bacteria Transform S bacteria

Genes IX Lewin et al ©
Nucleic acid is inherited in next generation

Nucleic acid controls Amino acid sequence in proteins

The first evidence that DNA controls AA sequence came from the study of genetic disease sickle cell anemia

When the protein was analyzed in these patients. Most of them found to contain change in Amino acid sequence at 6th position Glu=>Val.

Genes IX Lewin et al ©
The DNA molecule

Francis Crick,
James Watson
Maurice Wilkins
Rosalind Franklin

BOX 2-1 TABLE 1 Data Leading to the Formulation of Chargaff's Rules

<table>
<thead>
<tr>
<th>Source</th>
<th>Adenine to Guanine</th>
<th>Thymine to Cytosine</th>
<th>Adenine to Thymine</th>
<th>Guanine to Cytosine</th>
<th>Purines to Pyrimidines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ox</td>
<td>1.29</td>
<td>1.43</td>
<td>1.04</td>
<td>1.00</td>
<td>1.1</td>
</tr>
<tr>
<td>Human</td>
<td>1.56</td>
<td>1.75</td>
<td>1.00</td>
<td>1.00</td>
<td>1.0</td>
</tr>
<tr>
<td>Hen</td>
<td>1.45</td>
<td>1.29</td>
<td>1.06</td>
<td>0.91</td>
<td>0.99</td>
</tr>
<tr>
<td>Salmon</td>
<td>1.43</td>
<td>1.43</td>
<td>1.02</td>
<td>1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>Wheat</td>
<td>1.22</td>
<td>1.18</td>
<td>1.00</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>Yeast</td>
<td>1.67</td>
<td>1.92</td>
<td>1.03</td>
<td>1.20</td>
<td>1.0</td>
</tr>
<tr>
<td>Hemophilus influenzae</td>
<td>1.74</td>
<td>1.54</td>
<td>1.07</td>
<td>0.91</td>
<td>1.0</td>
</tr>
<tr>
<td>Escherichia coli K2</td>
<td>1.05</td>
<td>0.95</td>
<td>1.09</td>
<td>0.99</td>
<td>1.0</td>
</tr>
<tr>
<td>Avian tuberculosis bacillus</td>
<td>0.4</td>
<td>0.4</td>
<td>1.09</td>
<td>1.08</td>
<td>1.1</td>
</tr>
<tr>
<td>Serratia marcescens</td>
<td>0.7</td>
<td>0.7</td>
<td>0.95</td>
<td>0.86</td>
<td>0.9</td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>0.7</td>
<td>0.7</td>
<td>1.12</td>
<td>0.69</td>
<td>1.0</td>
</tr>
</tbody>
</table>

DNA replication

Arthur Kornberg: discover first DNA polymerase enzyme (Pol I). It works only in presence of DNA template.
DNA replicate by semi-conservative mode

Meselson and Stahl @CalTech
Experiment demonstrating that RNA is synthesized in the nucleus and move to cytoplasm
RNA dominate the information decoding pathway
Central Dogma of Molecular Biology

DNA → RNA → PROTEIN

Biotechnology Clark et al ©

Encarta ©
**Virus**

- Viruses are the entities that border on living.
- Made up of outer protein coat (capsid) and a genetic material ss/dsDNA or ss/dsRNA.
- Devoid of cellular machinery and cannot make energy or duplicate it genome.
- Virus are obligatory parasites.
- Most of the understanding of genetic material and genes functions comes from the study of Bactriophages.
- T4, lambda, Pl and Mu were commonly used for the study.
- The genetic material of certain viruses (QB) are as small as 3500 bp consist of only 4 genes. Some viruses genome is complex and may contains around 200 genes.
- The genes are categorized as early or late depending upon their action in time of infection.
- The virus hijacks host translation machinery to synthesis its proteins and DNA for survival.
- This is good system to study DNA replication and RNA transcription.
- Virus inserts/ integrate it DNA into host genome and modify host genome.
- To alter genome of host organism (Lenti virus).
- For gene therapy on humans.

Biotechnology Clark et al ©
**Escherichia. coli**

- Most important organism of Modern biotechnology (work horse)
- Prokaryote, Gram –ve, rod-shaped and about 1x2.5 micron in size. with 10-12 flagella for locomotion and numerous pili.
- No nucleus and single circular chromosome. It has around 4000 genes.
- Usually found in the gut, but are generally harmless. Occasional strains are pathogenic that secrete enterotoxin causing diarrhea.
- They grow on simple liquid or solid nutrient. The doubling time is 20 min. It can be stored in a refrigerator for weeks and can be maintained frozen at -70°C for 20 or more years.
- The have extra chromosomal genetic material called colicin plasmid, which harbors genes to kill neighboring cells. This plasmid found pivotal role in modern biotechnology after removal of toxin genes.
Fungi: *Saccharomyces cerevisiae* (baker’s yeast)

- Unicellular eukaryote with diploid or haploid chromosome.
- Distinct nucleus, cellular components are compartmentalized.
- Cellular organelles includes endoplasmic reticulum, Golgi and mitochondria.
- *S. cerevisiae* has 16 linear chromosomes with telomeres and centromeres, the features absents in bacteria.
- The genome is 12 Mb of DNA with 6000 genes.
- Like bacteria, yeast is also easy to maintain in simple media. Their doubling time is 90 min (compare to 20 min in *E. coli*).
- Can be stored frozen at -70°C for years.
- Yeast also has extra-chromosomal plasmid called 2-micron circle. It is very useful in yeast manipulation.
- Yeast grow by budding to give one daughter and mother cell.
- It also shows sexual mating: with a and α type haploid cells mate to form genetically unique diploid cells.
- Under poor environment spore formation takes place called Ascus, subsequent meiosis give 4 haploid spores.
- The organism is very easy to manipulate genetically.
- Useful to express proteins from higher eukaryotes with post-translational modifications.

Biotechnology *Clark et al* ©
Caenorhabditis elegans

- C. elegans is a small round worm found in rotting vegetation and feeds on bacteria.
- They are self-fertilizing hermaphrodite (and separate male).
- There are exactly 959 somatic cells in C. elegans.
- Life cycle of this worm lasts for 3 days (Fast reproduction).
- The organism is transparent and hence very useful in studying real-time cellular physiology with fluorescent labeling.
- The advantage of using C. elegans is to study development, aging, sexual dimorphism, alcohol metabolism and other phenomena that apply to humans.

Easy to grow and make genetic clones of itself
**Drosophila melanogaster** (fruit fly)

- Multi cellular invertebrate commonly found around rotting fruit.
- Small insect to handle, simple food source, and easy to maintain.
- Life span is 2 weeks.
- Genome is sequenced (165 Mb DNA) divided into 3 autosome and x/y sex chromosome.
- 12,000 predicted genes.
- Very versatile model to study genetic.
- Visible marker for genetic study includes wings, legs, antenna, eyes, hair formation, etc.
- True sexual reproductions.
- Easy genetic crosses.

Polytene chromosomes: While Drosophilla’s larval development the number of cells remains constant while the size of cells increase dramatically. In order for the cells to work properly there need lot of mRNA and proteins. This is accomplished by duplication of genome hundreds of times to give giant polytene chromosome.
Simple vertebrate, easy to maintain and breed in home made aquarium
Great model system to study genetic.
Adult lays about 200 eggs, great numbers of offspring to study in subsequent generation.
Embryo develop out side the mother and transparent, good system to study the embryo development.
Cells inside the embryo can be manipulated to study the effect of position on cellular development.
75% of zebra fish DNA is homologous to humans thus the finding from zebra fish gene study can be extrapolated to human genes.
**Arabidopsis thaliana** (wild mustard seeds)

- Excellent model system for plant biotechnology.
- Easily grown and maintained in laboratory settings.
- Entire generation is finished in 6-10 weeks.
- It has small genome (125Mb) and 25,000 genes as compared to 40-50,000 genes in rice.
- This plant can be maintained in haploid state like yeast.

**Mus musculus** (Mouse)

- This model is most closely related to humans.
- 20 different chromosome, and less than 1% genes has no human counterpart.
- Very easy to manipulate genetically by creating knockout mouse or transgenic mouse (adding extra genes).
- Ideal to study growth, development and physiology.
Tissue culture

- Mammalian cell culture is another way to study in-vitro effects.
- No ethical issues to address.
- Different human, monkey, hamster cell lines are available (kidney, liver, heart, NIH3T3, Hella, HK213, etc).
- These are immortal cell lines and can be grown forever with proper maintenance.
- Direct genetic manipulation (deletion or mutations) is possible.
- Recombinant protein from human origin can purified for medical use.
- Every expensive to maintain. Need fetal bovin serum and CO$_2$ incubator.
- Insect cell lines are also available for heterologous protein expression.
- Advantage: Can be maintain on serum free media, grow on lower temperature and CO$_2$ free condition.