CHAPTER 1 DNA and Chromosome

1.1 DNA is genetic material

1.2 DNA Structure

1.3 Chromosome

1.4 Genome and Gene

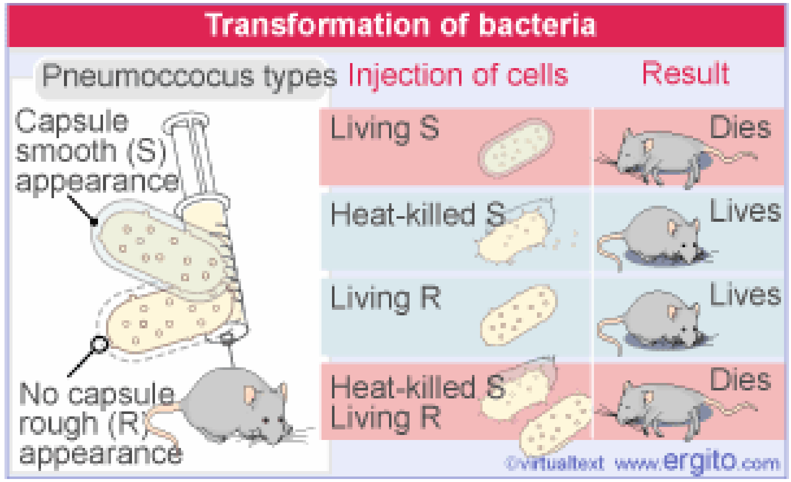
1.1 DNA is Genetic Material*

1.1.1 DNA is the genetic material of bacteria

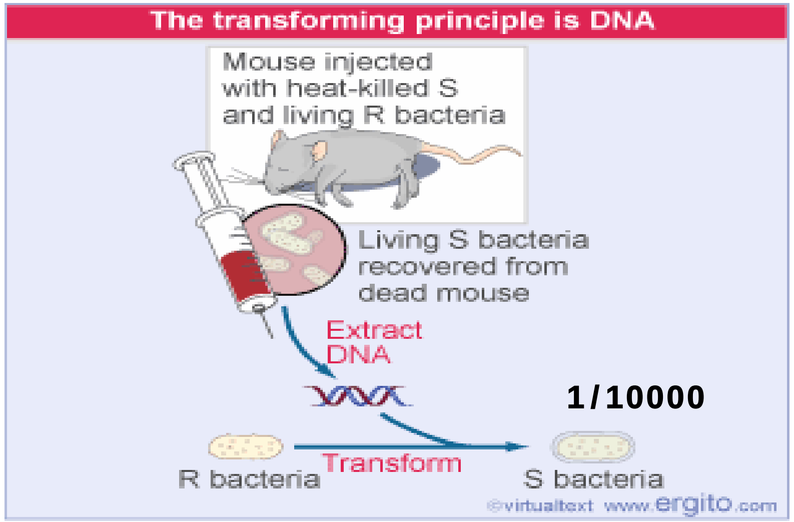
1928, Fred Griffith

~16 years

1944,
Oswald Avery
Colin MacLeod
Maclyn McCarty



- 转化（Transformation）

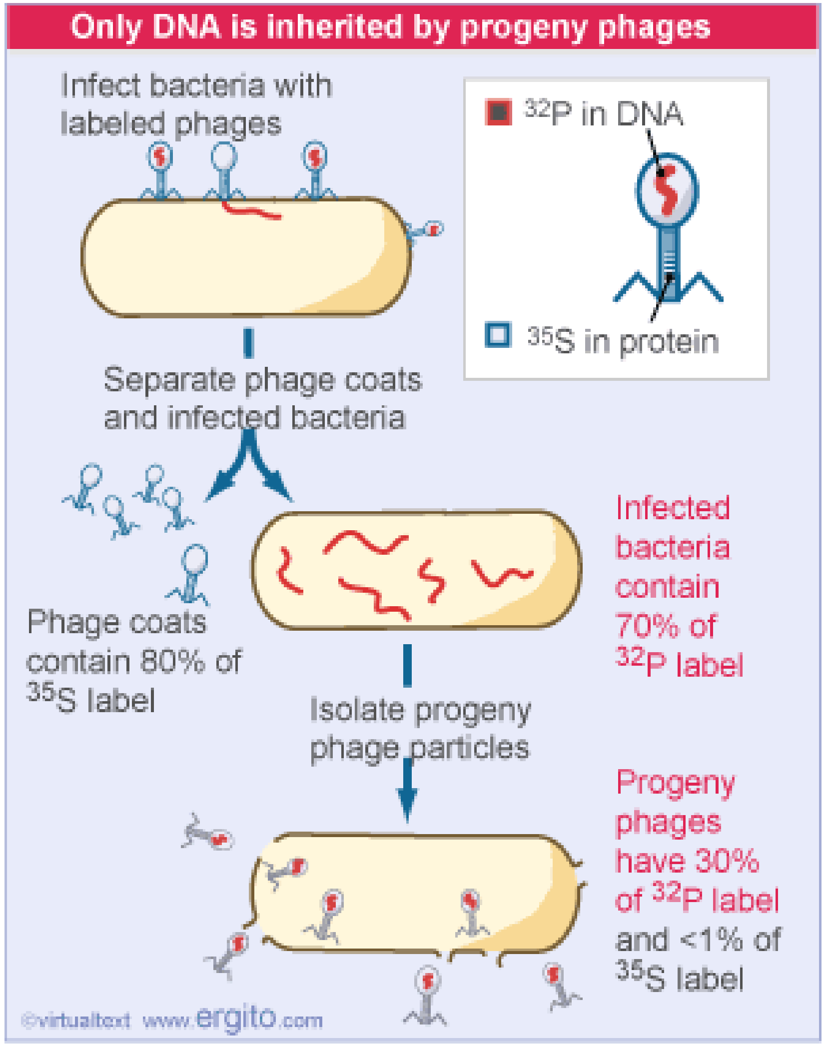


1/10000

- 转化因子（Transforming principle ）

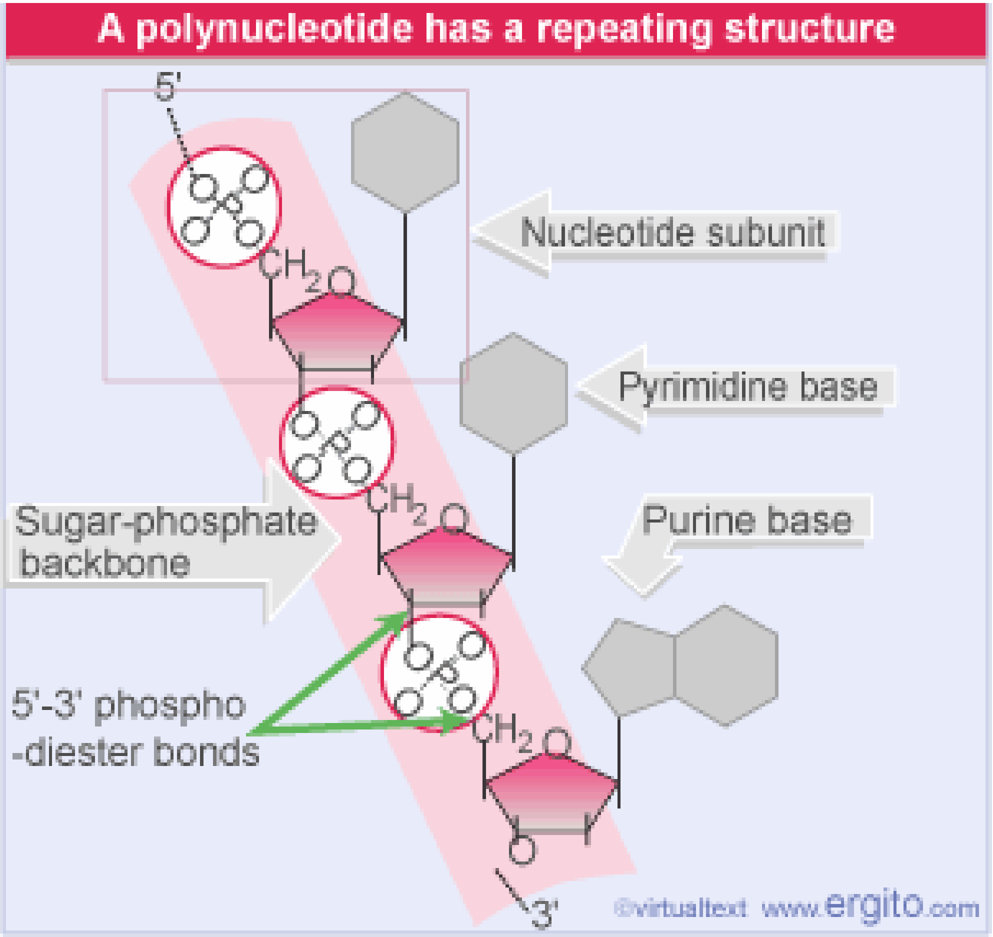
1.1.2 The genetic material of phage T2 is DNA

1952
Alfred Hershey
Martha Chase



1.2 DNA Structure

1.2.1 A nucleic acid (DNA/RNA) is a polynucleotide

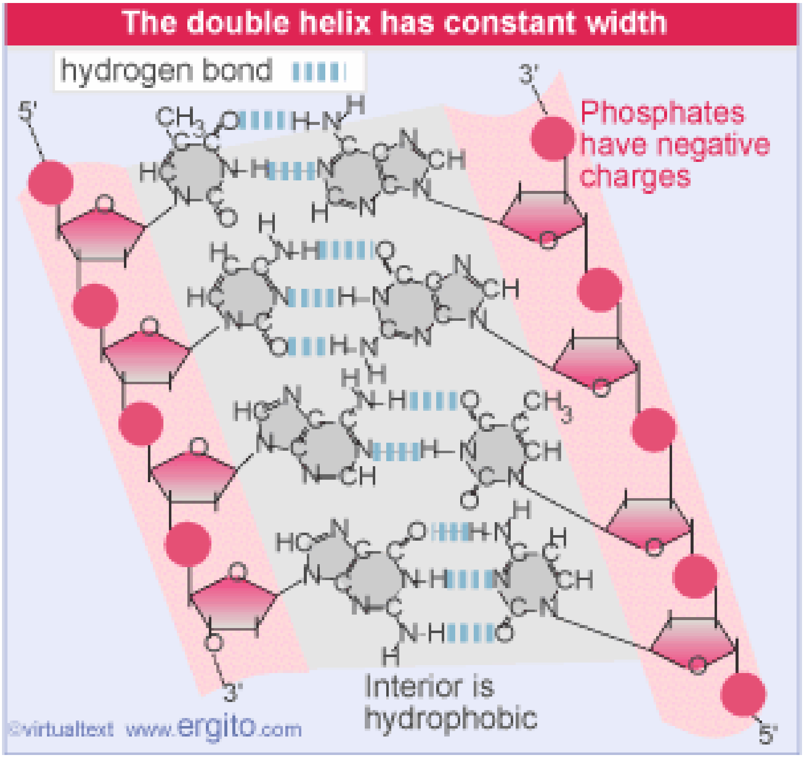


DNA: dA, dG, dC, dT
2′-H (deoxyribose sugar)

RNA: A, G, C, U
2′-OH (ribose sugar)

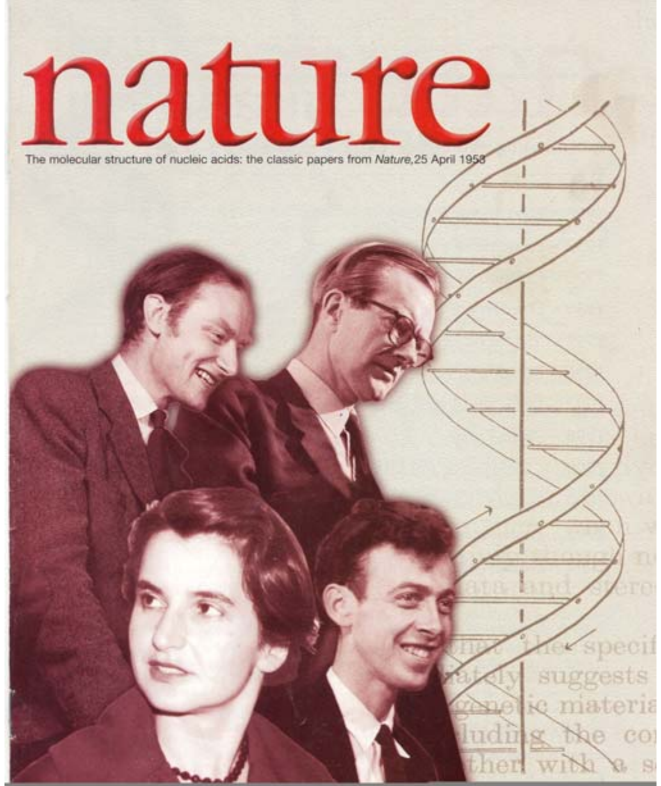
Direction: 5′-p → 3′-OH

1.2.2 DNA is a double helix



Key term: Complementary Base Pair A-T, C-G

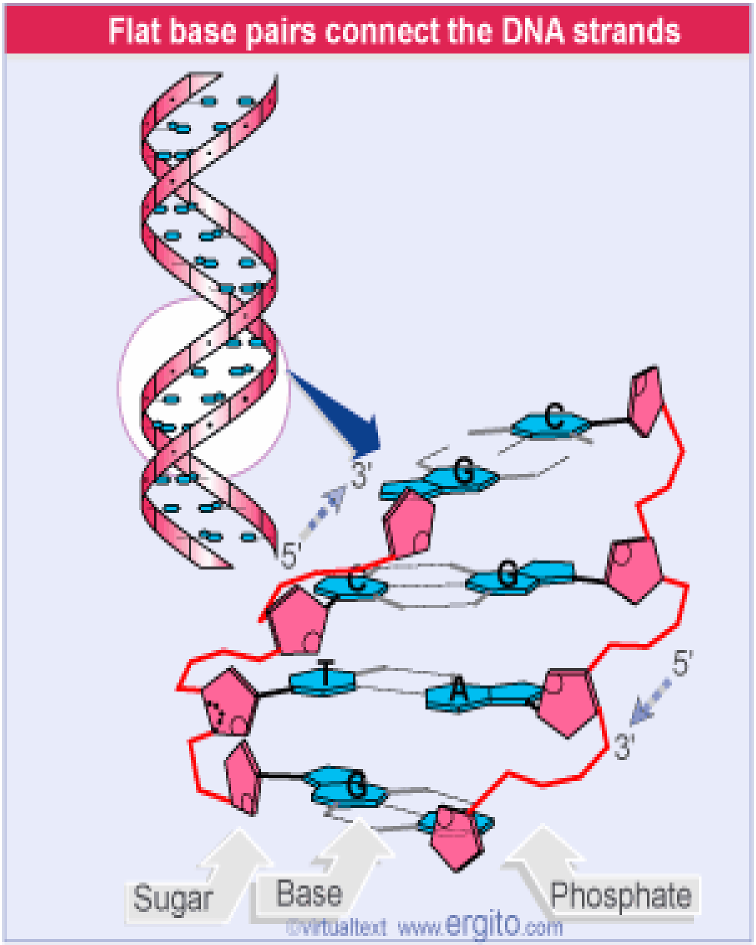
DNA double helix —— a research legacy

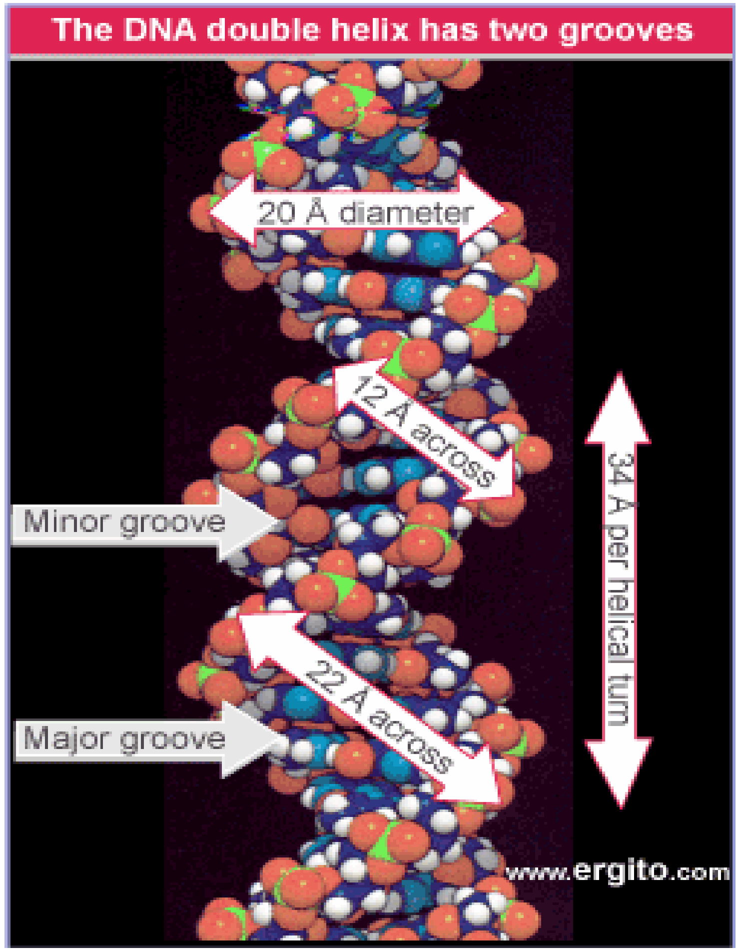


Yesterday's research –
today's legacy

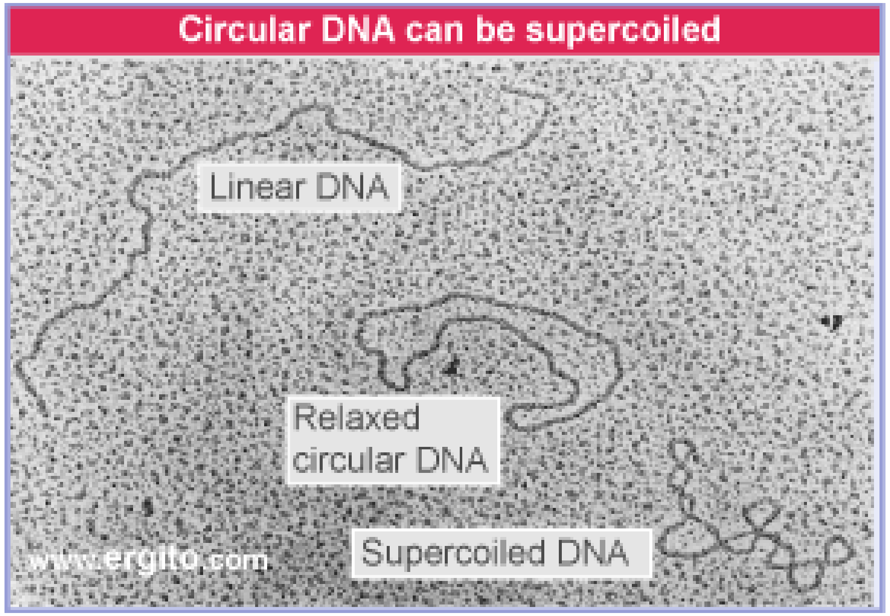
DNA double helix

B-form DNA


Flat base pairs connect the DNA strands


The DNA double helix has two grooves

1.2.3 Supercoiling structure of DNA



Positive supercoiling

The DNA is twisted around itself in the same sense as the two strands within the double helix,causing the DNA strands to wound around one another more tightly, so that there are more base pairs per turn.

Negative supercoiling

The DNA is twisted around itself in the opposite sense as the two strands within the double helix, causing the DNA strands to be twisted around one another less tightly, so there are fewer base pairs per turn.

Type I topoisomerases

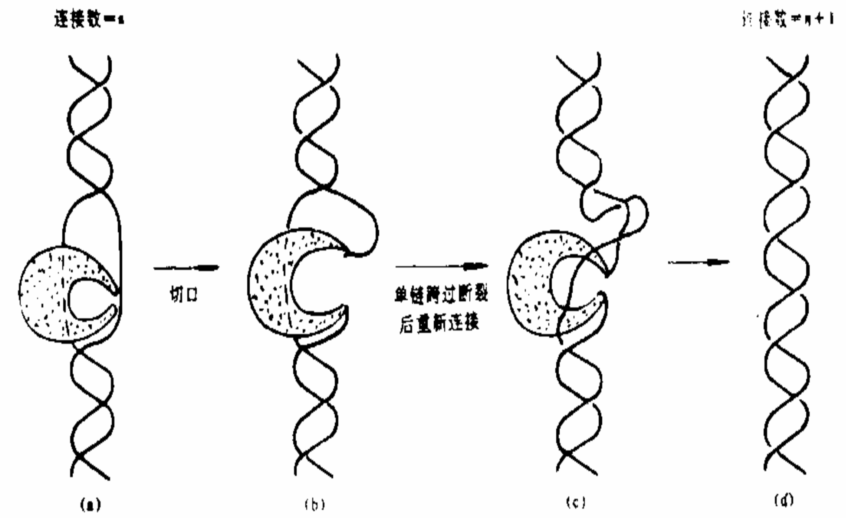
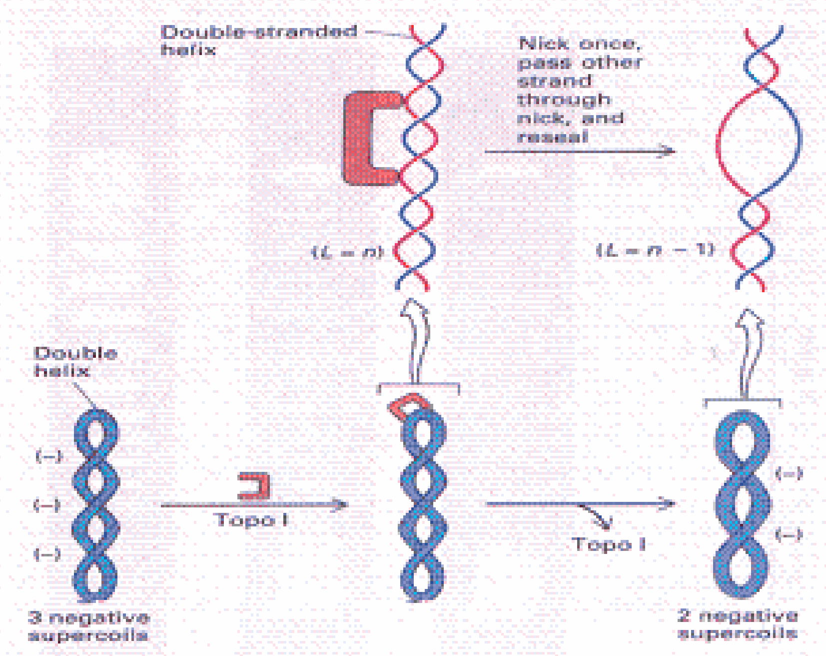
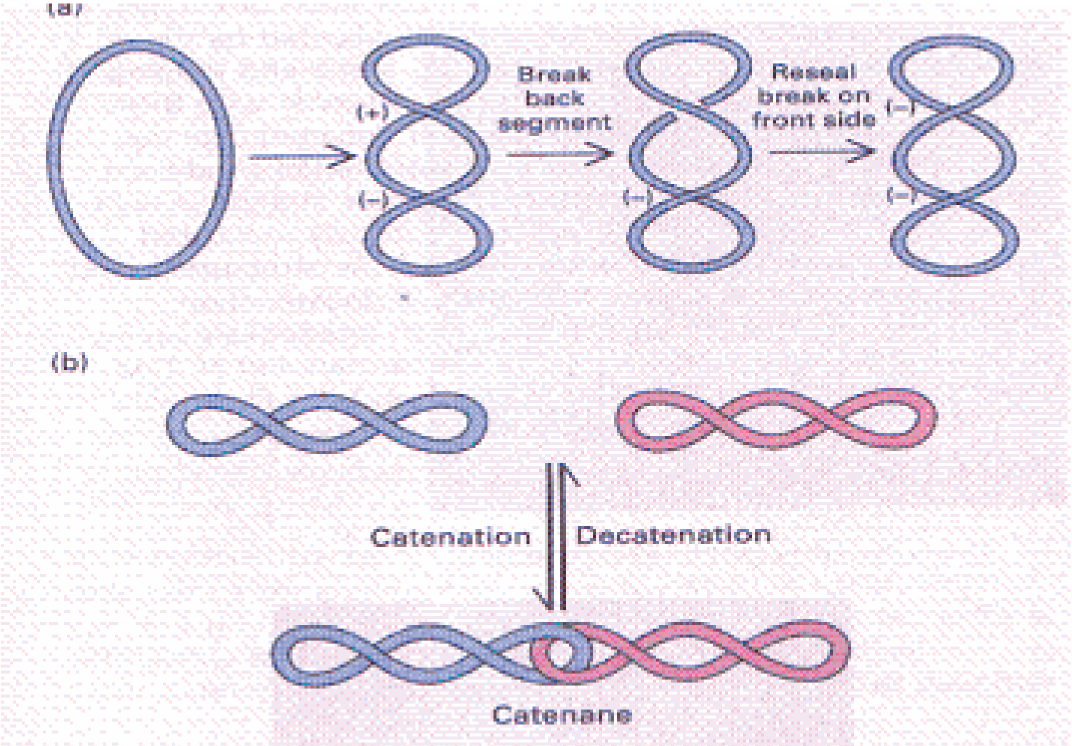


图 1-14 拓扑异构酶 I 的作用机理



Type II topoisomerases



链接反应

Summary of topoisomerases

A **type I topoisomerase** is an enzyme that changes the topology of DNA by nicking and resealing one strand of DNA without requiring any input of energy.

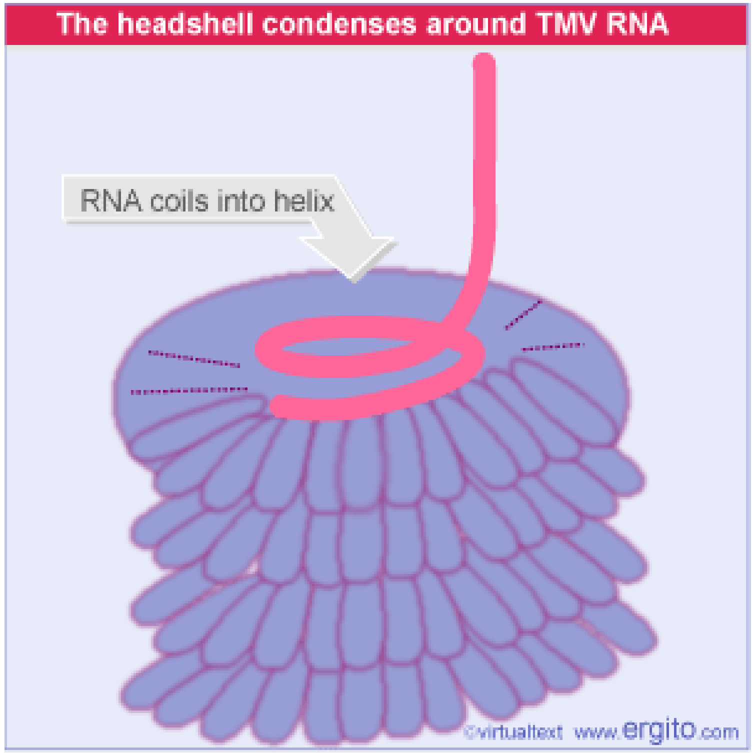
A **type II topoisomerase** is an enzyme that changes the topology of DNA by nicking and resealing both strands of DNA. The hydrolysis of ATP may be used to drive the reactions.

E. coli DNA **Gyrase** is a type II topoisomerase that can introduce negative supercoils.

1.3 Chromosome

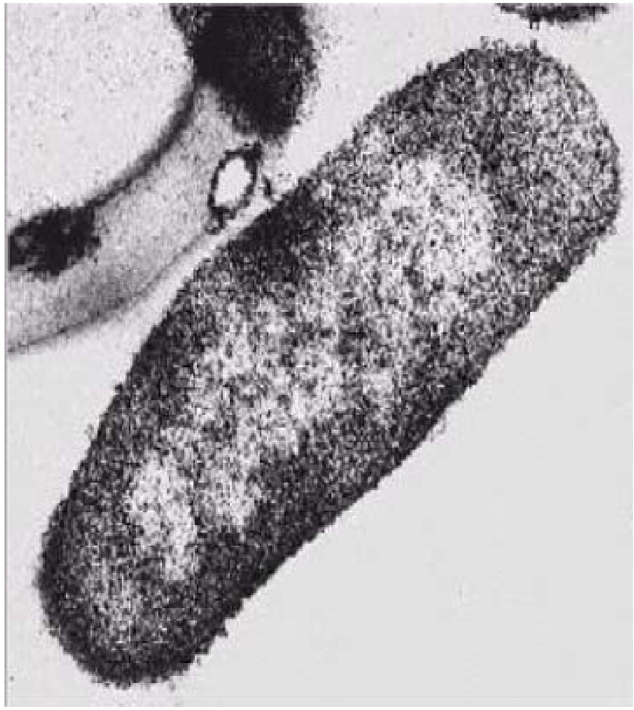
种类	形状	尺寸	核酸类型	长度
TMV	细丝	0.008 x 0.3 mm	1 单链 RNA	2 mm = 6.4 kb
噬菌体fd	细丝	0.006 x 0.85 mm	1 单链 DNA	2 mm = 6.0 kb
腺病毒	二十面体	0.07 mm	1 双链 DNA	11 mm = 35.0 kb
T4	二十面体	0.065 x 0.10 mm	1 双链 DNA	55 mm = 170.0 kb
E. coli	柱体	1.7 x 0.65 mm	1 双链 DNA	1.3 mm = 4.2×10^3 kb
线粒体 (人类)	扁平棒状	3.0 x 0.5 mm	~10 相同的 双链 DNAs	50 mm = 16.0 kb
细胞核 (人类)	棒状	6 mm	46 染色体双链 DNA	1.8 m = 6×10^6 kb

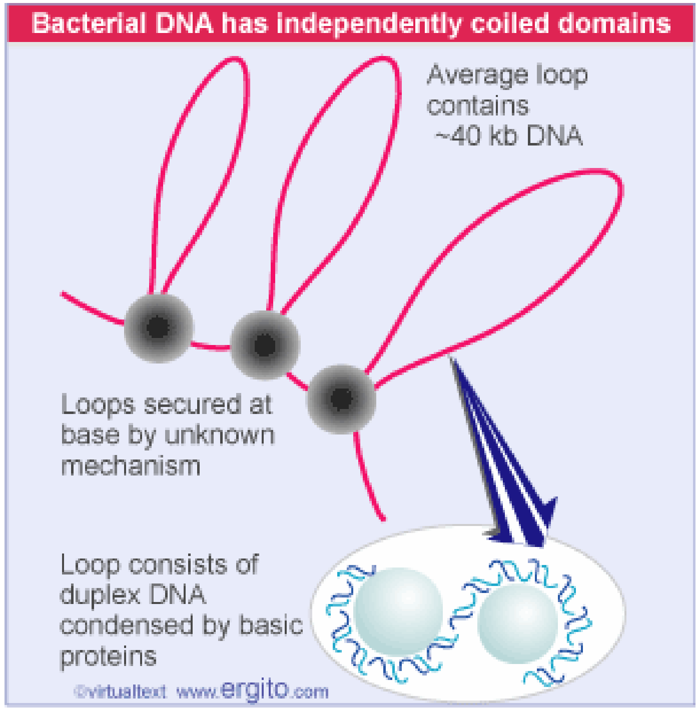
1.3.1 Viral genomes are packaged into their coats



Note: Viruses have genomes of double-stranded or single-stranded DNA or RNA.

1.3.2 Prokaryotic chromosome, a supercoiled nucleoid





The nucleoid is the region in a prokaryotic cell that contains the genome. The DNA is bound to proteins and is not enclosed by a membrane.

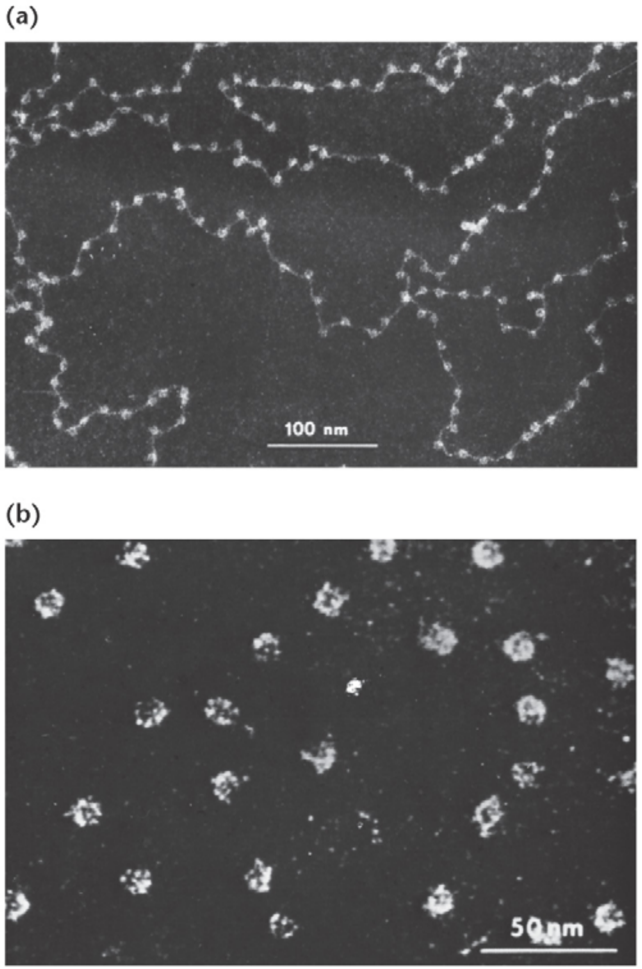
1.3.3 Eukaryotic chromatin and chromosomes

Chromatin is a mass of beads on a string

www.ergito.com

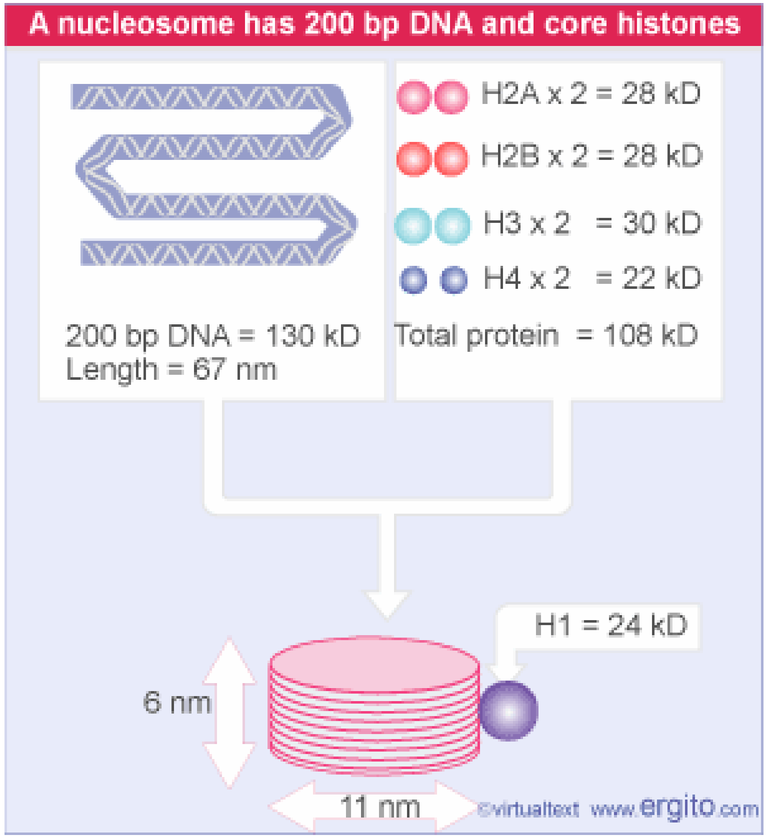
Chromatin describes the state of nuclear DNA and its associated proteins during the interphase (between mitoses) of the eukaryotic cell cycle.

Nucleosome is the subunit of chromatin



Individual nucleosomes are released by digestion of chromatin with micrococcal nuclease.

Nucleosome and Histones



Histones are extremely rich in the positive charged amino acids lysine and arginine.

Nucleosome:

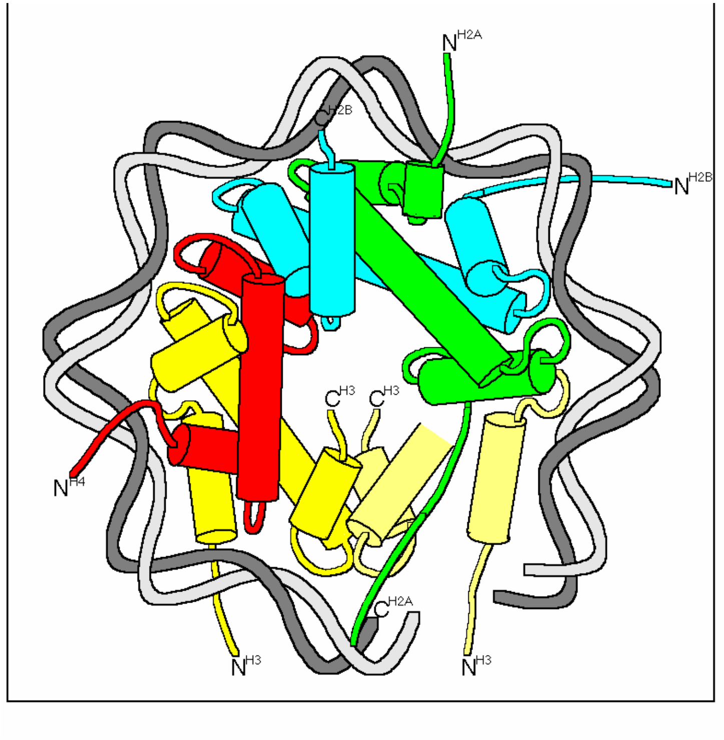
Histones: H1×1, H2A ×2,
H2B ×2, H3 ×2,
H4 ×2
DNA: ~200bp DNA

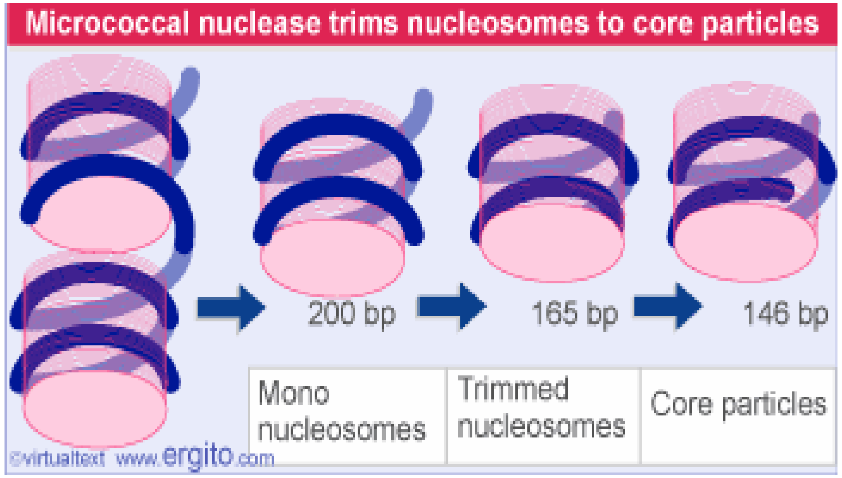
Table 12.2

Categories and Properties of Histone Proteins

Histone Type	Lysine-Arginine Content	Molecular Weight (Da)
H1	Lysine-rich	23,000
H2A	Slightly lysine-rich	14,000
H2B	Slightly lysine-rich	13,800
H3	Arginine-rich	15,300
H4	Arginine-rich	11,300

Core particle of nucleosome



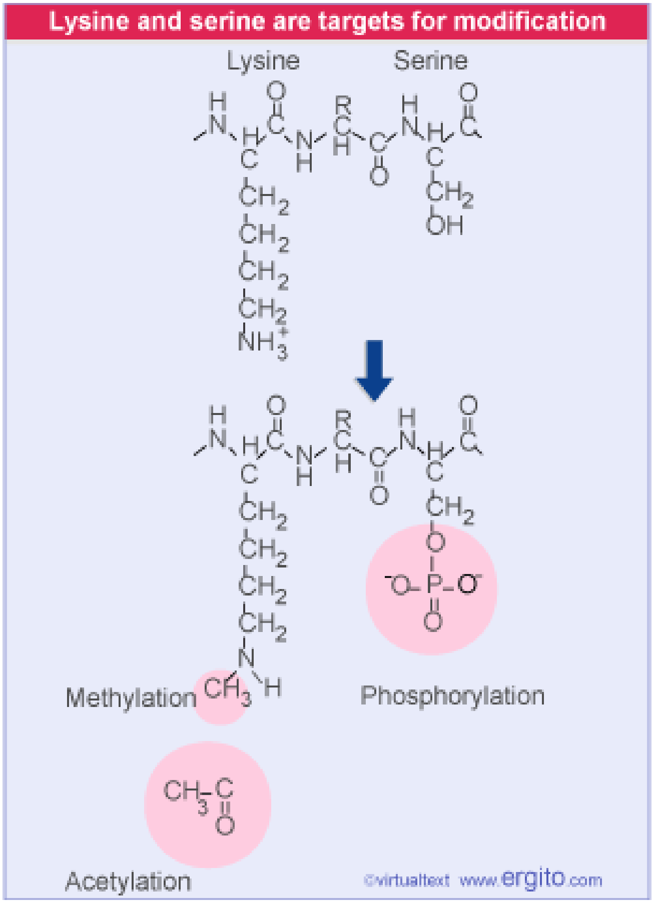


Core Particle:

Histones: H2A ×2, H2B ×2,H3 ×2, H4 ×2

DNA: 146bp DNA

The N-terminal tails of histones are modified



Histones are modified by methylation, acetylation, and phosphorylation.

The sites that are modified are concentrated in the N-terminal tails. These modifications have important effects on the structure of chromatin and in controlling gene expression.

Chromatin Remodeling

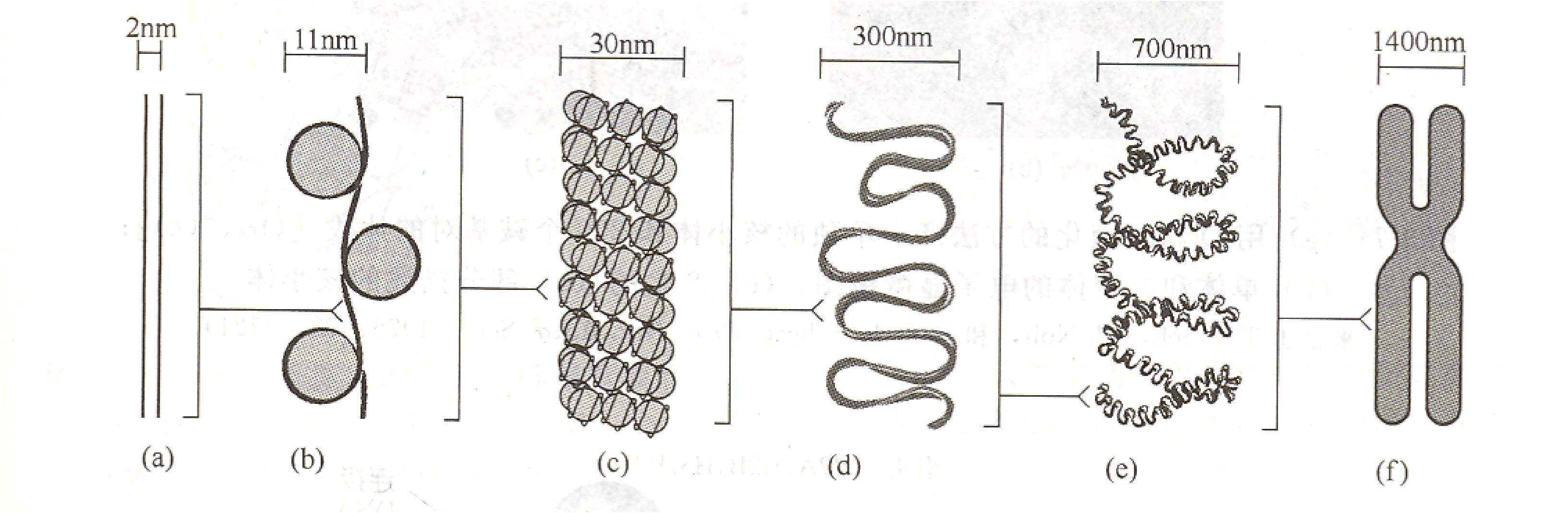
- Chromatin structure is dynamic
- Induced change in chromatin structure
    - Replication, gene expression
- Histone modification
    - Acetylation by histone acetyltransferase (HAT)
    - Methylation by methyl transferases
    - Phosphorylation by kinases
- DNA modifications
    - Methylation of cytosine (5-methyl C) in CpG islands

Heterochromatin

- 1928, staining differences in nuclei lead to terms euchromatin and heterochromatin
- Heterochromatin
  - Dark staining
  - Genetically inactive
    - Few genes, those present repressed
    - Replicates late in S phase
    - Centromeres and telomeres are heterochromatic
    - Portion of Y and inactivated X chromosomes
    - Position effect when genes are translocated to location adjacent to heterochromatin

DNA Condensation

The path of nucleosomes in the chromatin fiber



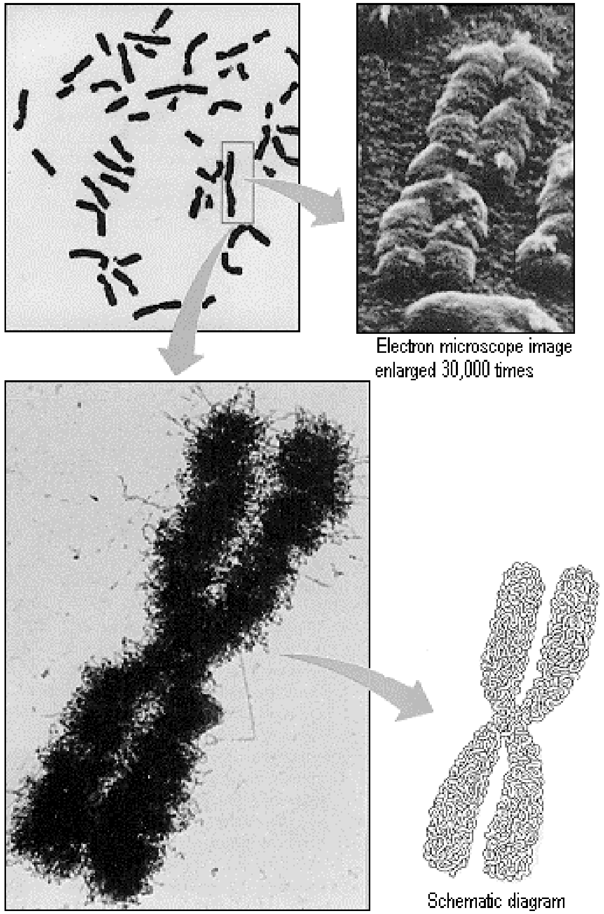
Chromatin proteins are subdivided into histones and nonhistones.

Nonhistone proteins:

e.g. high mobility group proteins (HMG protein);
and other DNA binding proteins.

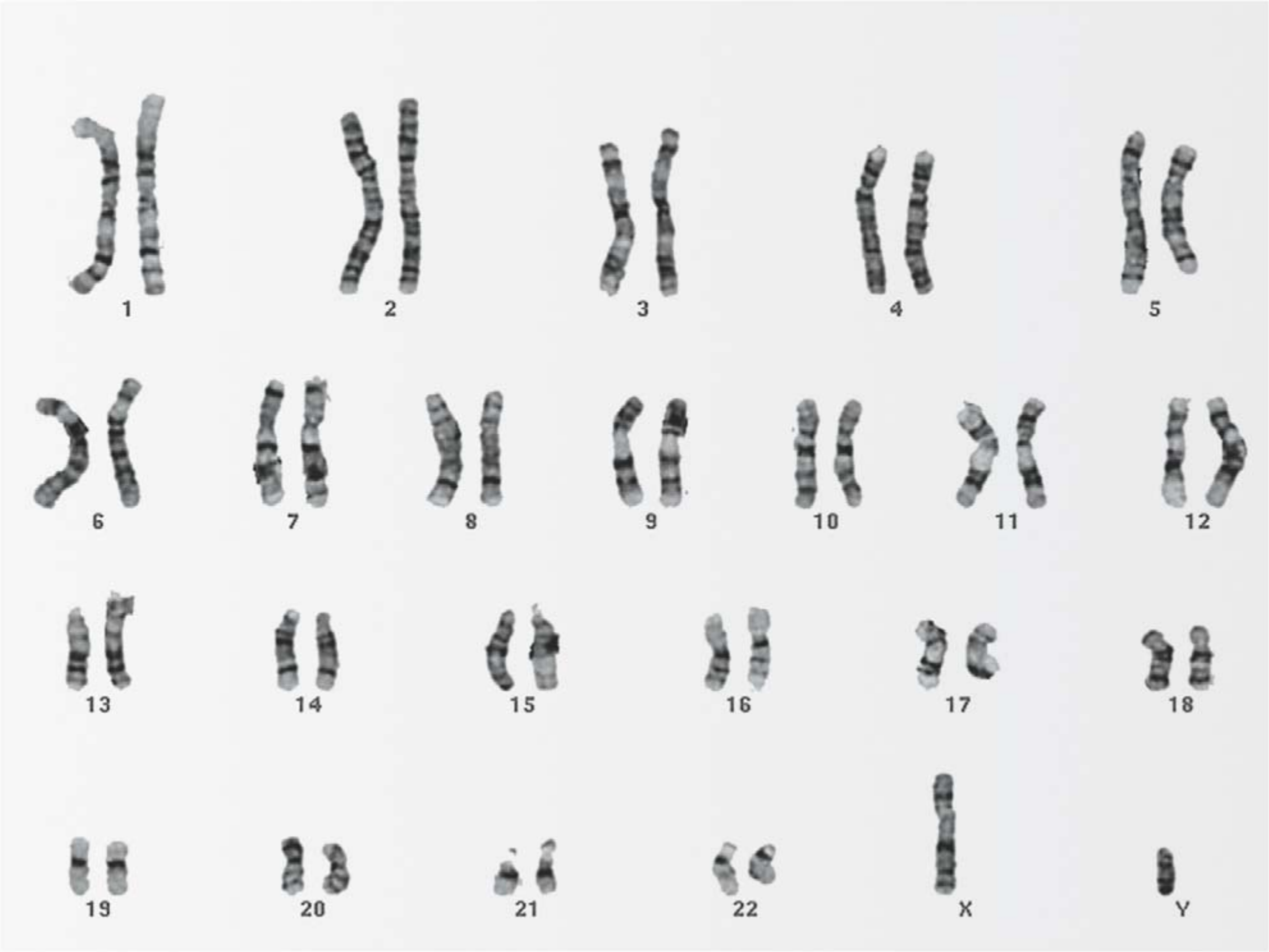
Chromosome

Human Chromosomes



A **chromosome** is a discrete unit of the genome carrying many genes. Each chromosome consists of a very long molecule of duplex DNA and an approximately equal mass of proteins. It is visible as a morphological entity only during cell division.

Human Chromosomes

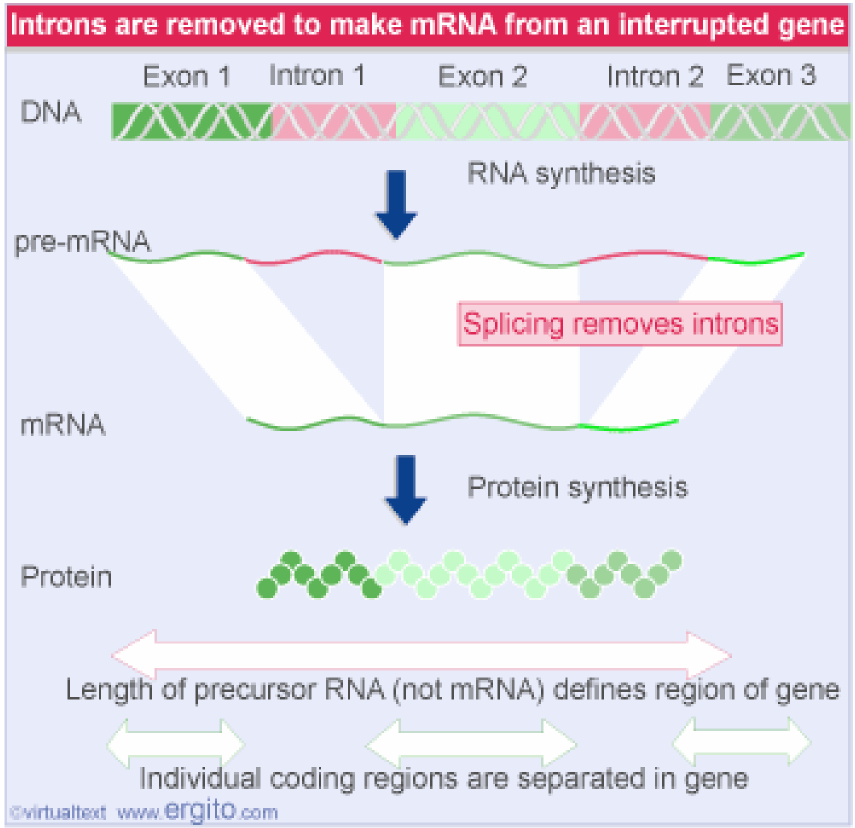


1.4 Genome and Gene

The **genome** is the complete set of sequences in the genetic material of an organism. It includes the sequence of each chromosome plus any DNA in organelles.

One gene : one enzyme hypothesis is the basis of modern genetics.
One gene : one polypeptide chain : a gene is a stretch of DNA coding for a single polypeptide chain.

1.4.1 Gene

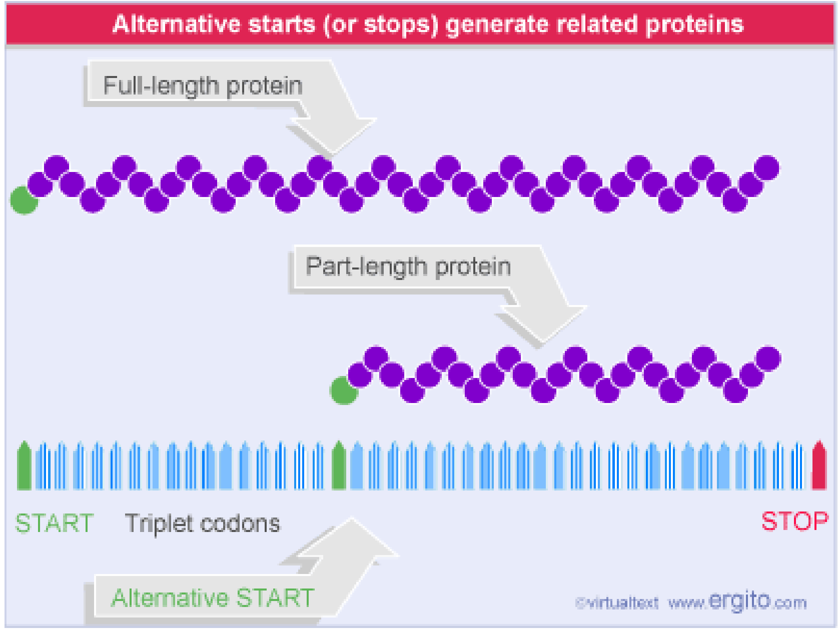


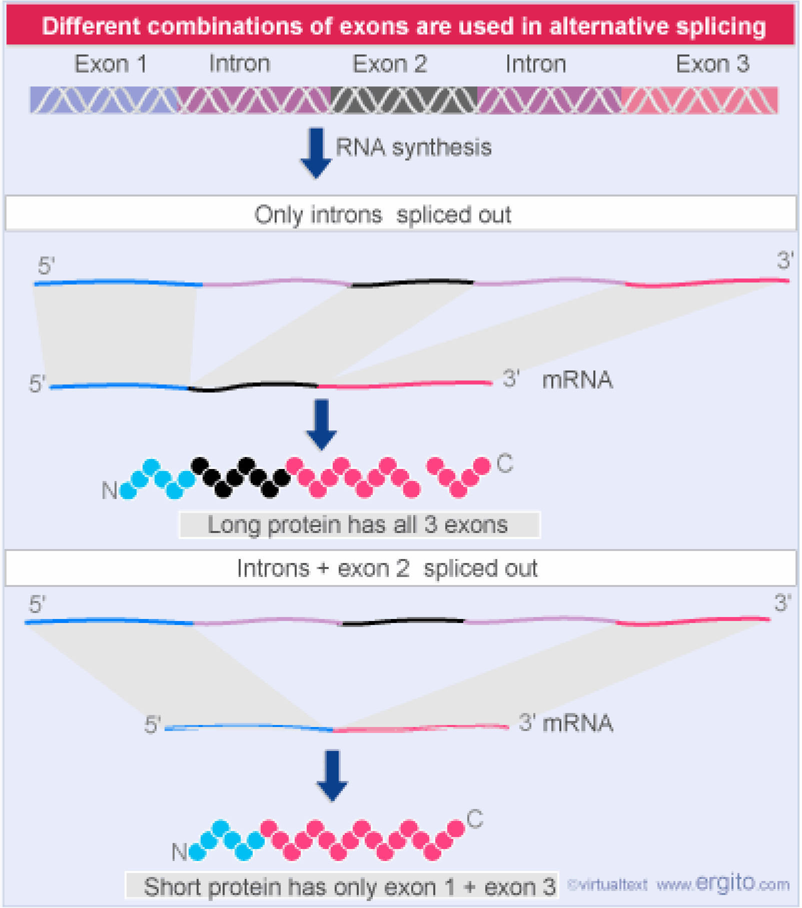
A **gene** is a stretch of DNA that functions as a unit to give rise to an RNA or protein product; it may include **exons** and **introns**.

An **exon** is any segment of a gene that is represented in the mature RNA product.

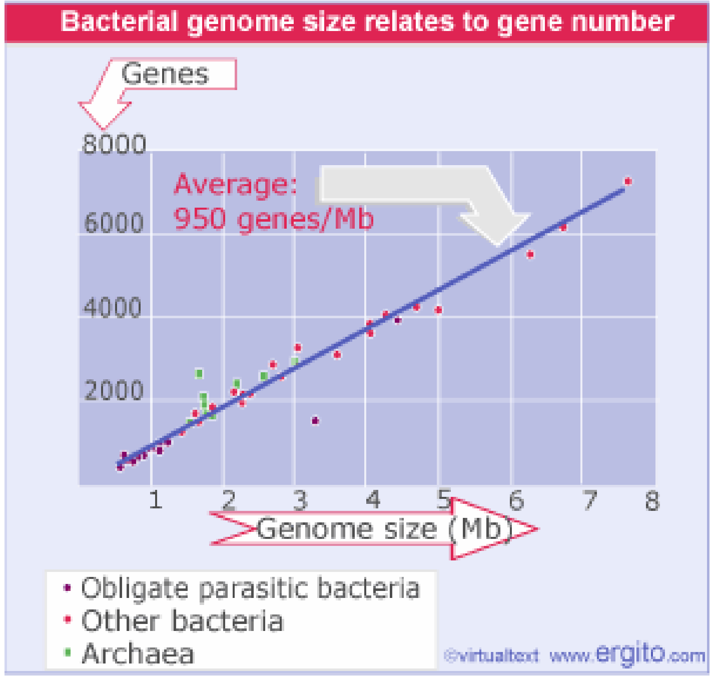
An **intron** is any segment of a gene that is transcribed, but removed from the transcript in RNA splicing.

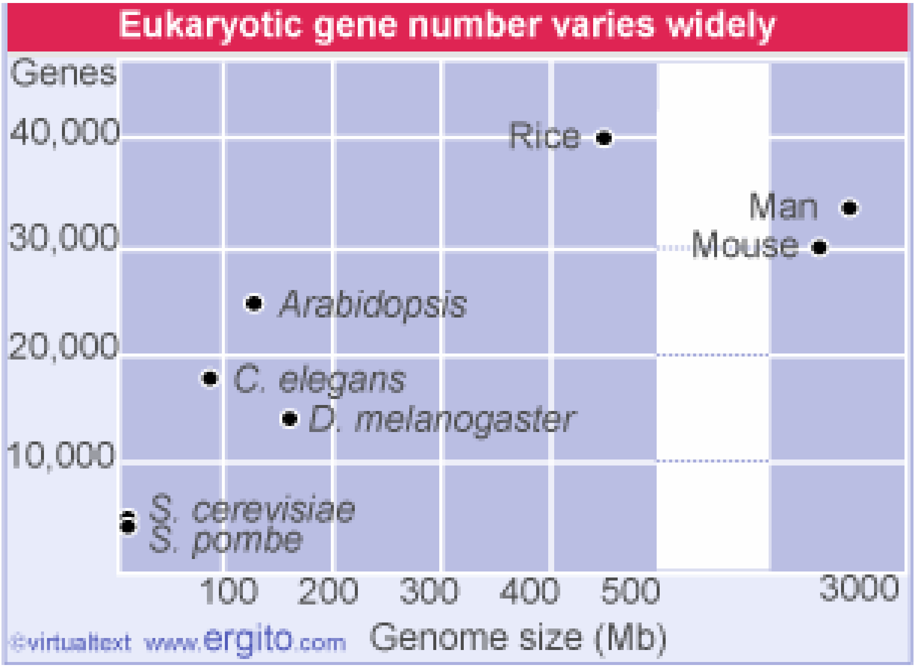
Some DNA sequences code for more than one protein





1.4.2 Genome Size and Gene Number


Bacterial genome size relates to gene number


Eukaryotic gene number varies widely

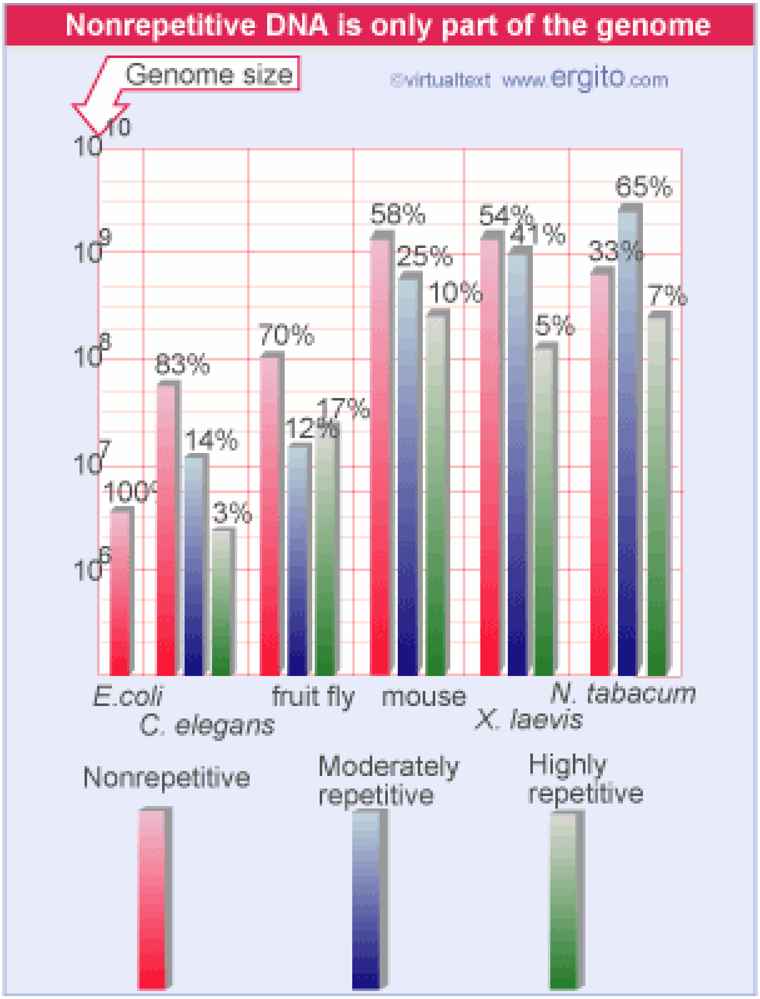
Repetitive DNA

Unique (single copy):
Most of the genes, a small percentage of the DNA

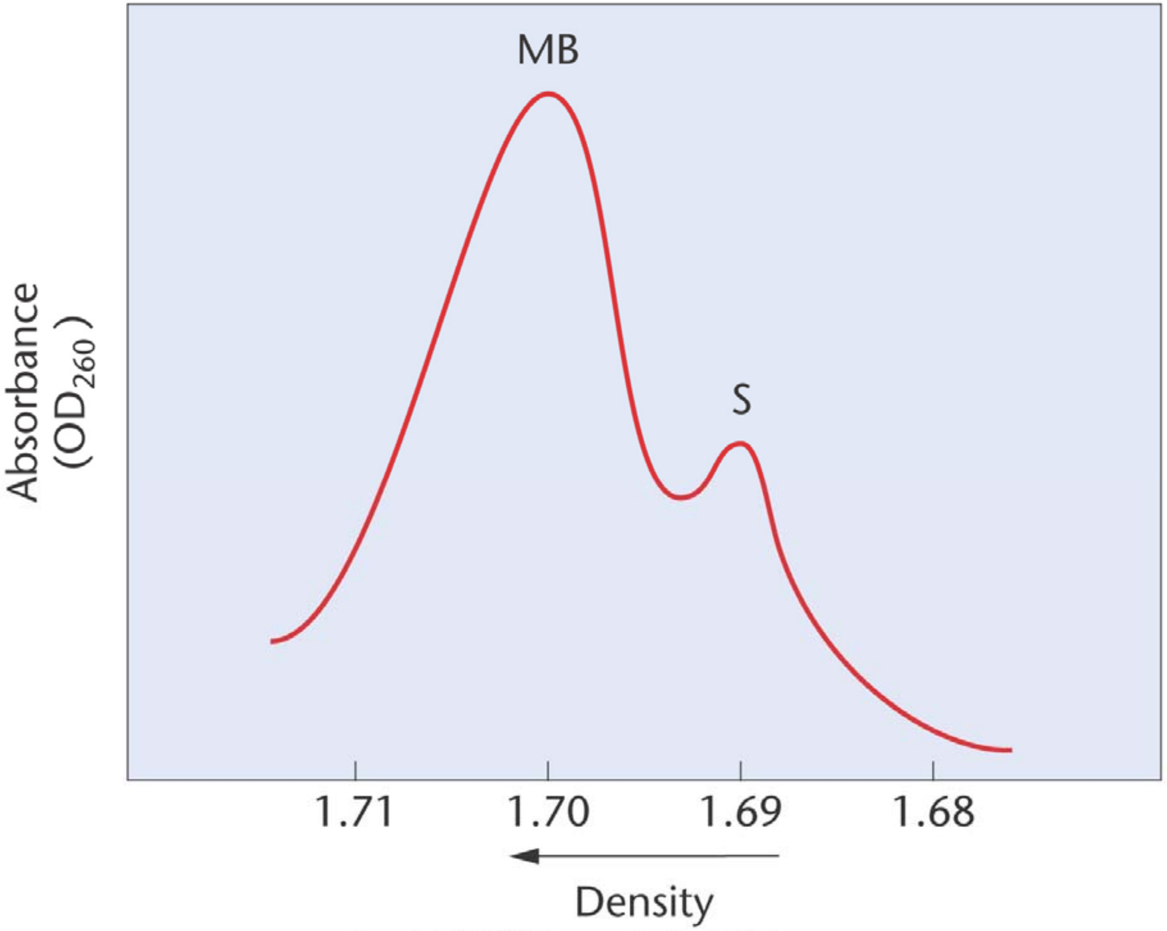
Slightly repetitive (1-10 copies):
Some genes

Middle repetitive (10 to several hundred copies):
e.g. rRNA genes

Highly repetitive (several hundred to several million copies):
e.g. satellite DNA



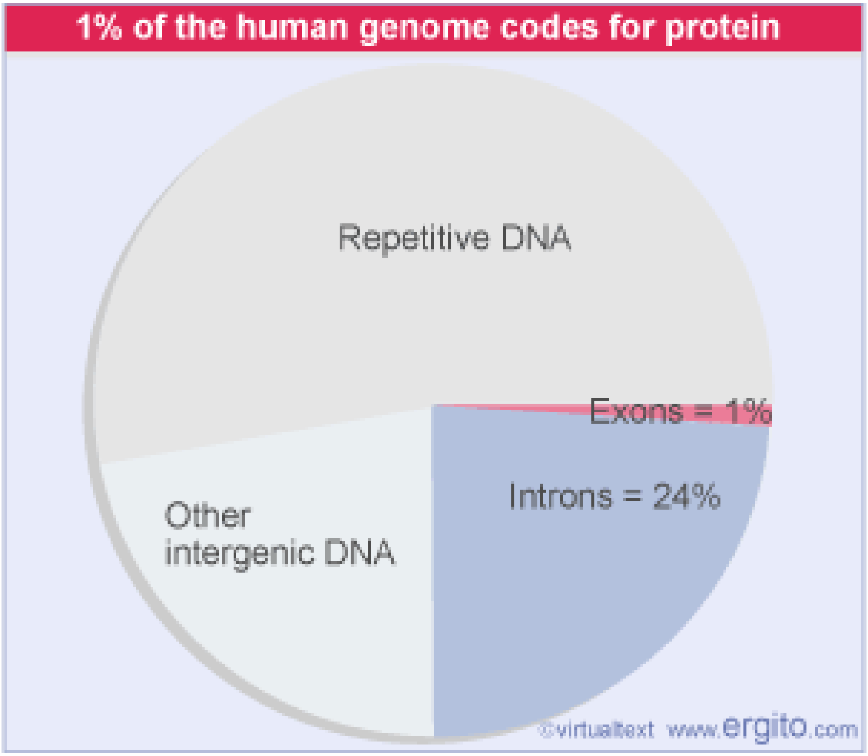
Buoyant Density Analysis



Satellite DNAs

- Some DNAs, when subjected to buoyant density analysis, give a density profile that is not a single peak as expected for random variations in the AT/CG content from one region of DNA to another
  - Some densities are over represented
    - Satellite DNAs
- Short sequence tandem repeats
- Heterochromatic regions, mostly associated with centromeres

The human genome has fewer genes than expected



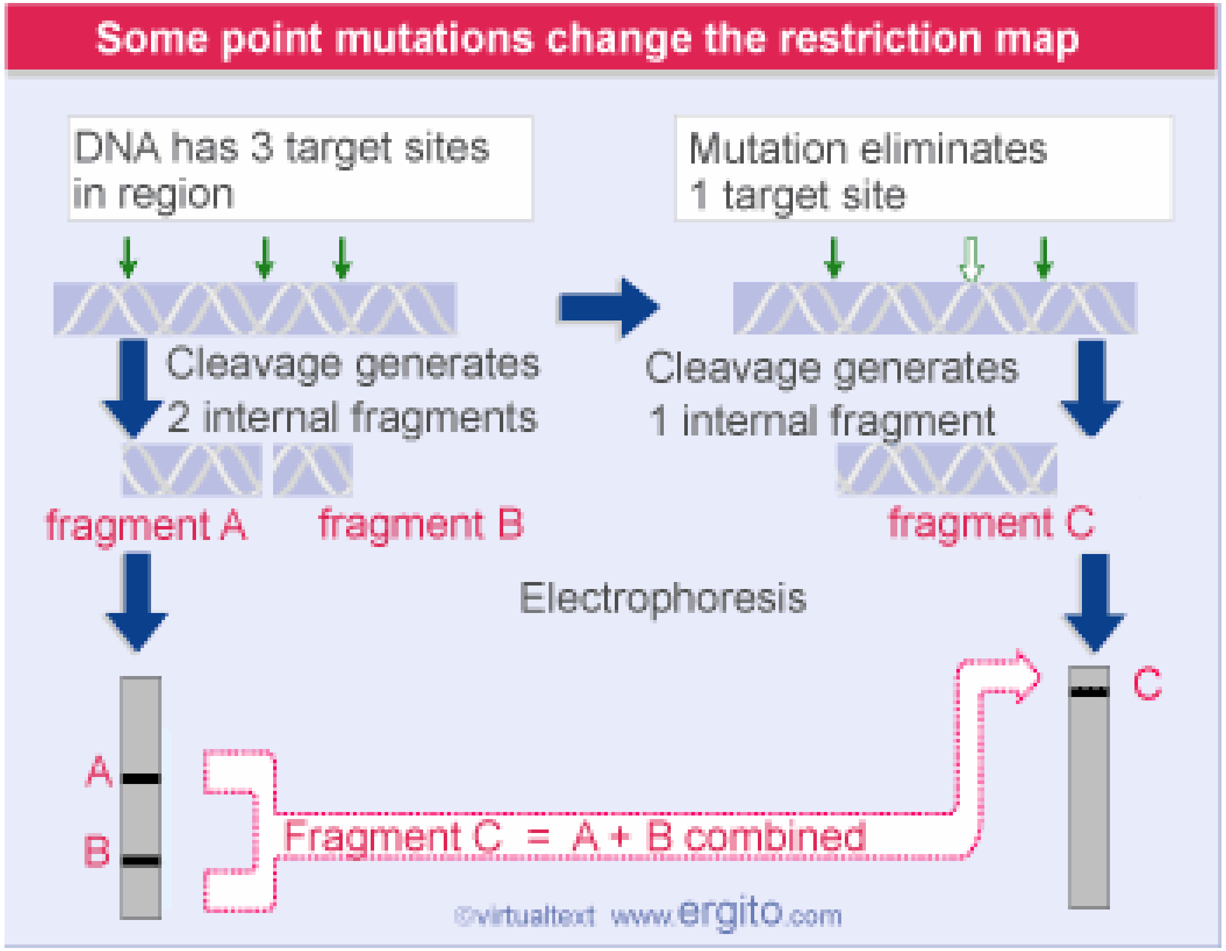
- Only 1% of the human genome consists of coding regions.
- The exons comprise ~5% of each gene, so genes (exons plus introns) comprise ~25% of the genome.
- The human genome has 30,000-40,000 genes.
- ~60% of human genes are alternatively spliced.
- Up to 80% of the alternative splices change protein sequence, so the proteome has ~50,000-60,000 members.

1.4.3 Individual genomes show extensive variation

- **Single nucleotide polymorphism (SNP)** describes a polymorphism (variation in sequence between individuals) caused by a change in a single nucleotide. This is responsible for most of the genetic variation between individuals.

- **Restriction fragment length polymorphism (RFLP)** refers to inherited differences in sites for restriction enzymes (for example, caused by base changes in the target site) that result in differences in the lengths of the fragments produced by cleavage with the relevant restriction enzyme.

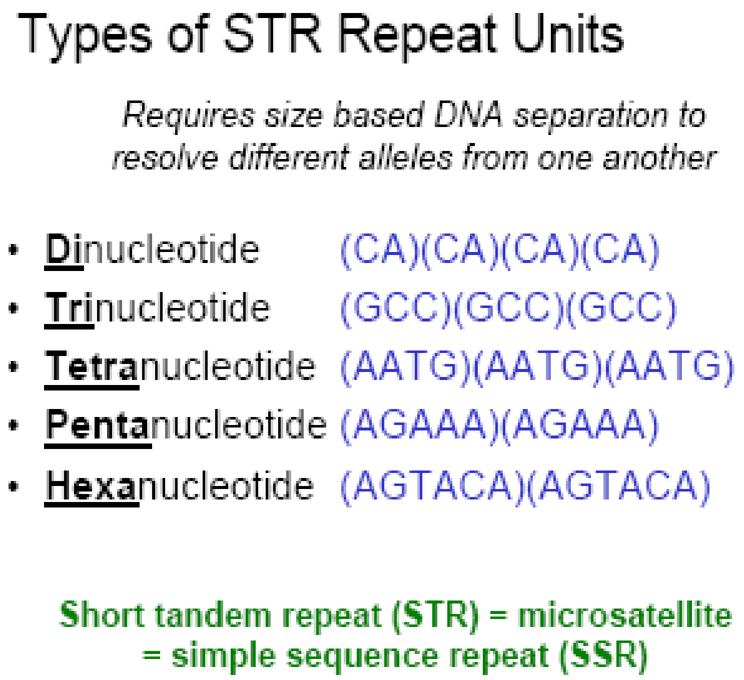
Restriction fragment length polymorphism (RFLP)



Some point mutations change the restriction map

©virtualtext www.ergito.com

Short Tandem Repeat (STR) Markers



Types of STR Repeat Units

Requires size based DNA separation to resolve different alleles from one another

Dinucleotide	(CA)(CA)(CA)(CA)
Trinucleotide	(GCC)(GCC)(GCC)
Tetranucleotide	(AATG)(AATG)(AATG)
Pentanucleotide	(AGAAA)(AGAAA)
Hexanucleotide	(AGTACA)(AGTACA)

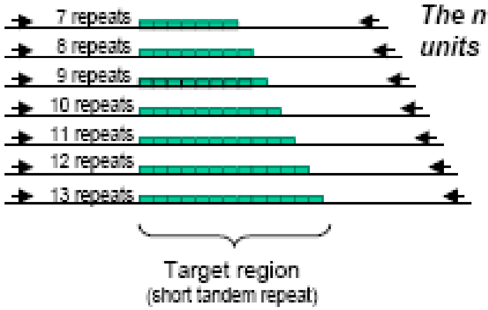
**Short tandem repeat (STR) = microsatellite
= simple sequence repeat (SSR)**

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

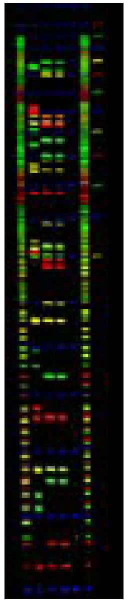
TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACA
GGTGGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGA
TAGATAGATATCATTGAAAGACAAAACAGAGATGGATGATAGAT
ACATGCTTACAGATGCACAC

= 12 GATA repeats ("12" is all that is reported)



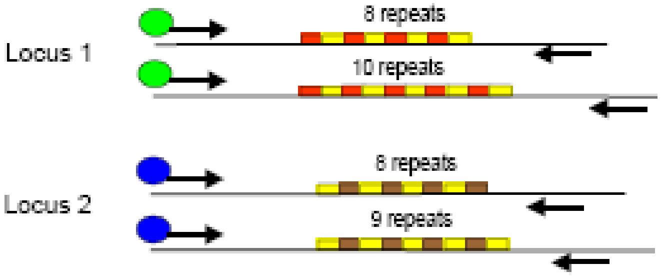
The number of consecutive repeat units can vary between people

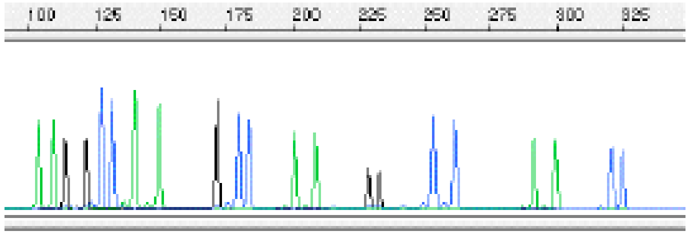
The FBI has selected 13 core **STR loci** that must be run in all DNA tests in order to provide a common currency with DNA profiles



Scanned Gel Image

The polymerase chain reaction (PCR) is used to amplify STR regions and label the amplicons with fluorescent dyes using locus-specific primers



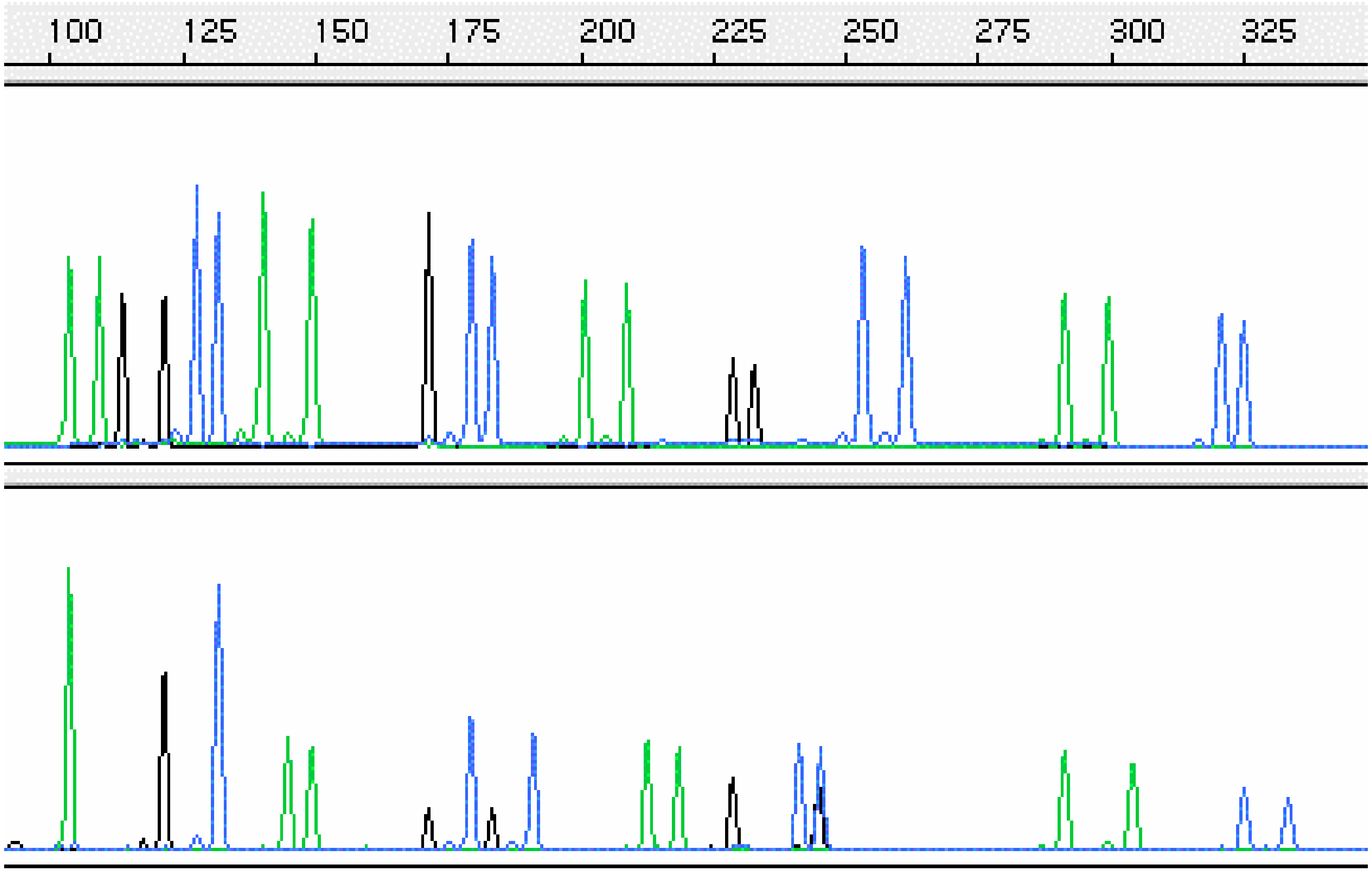


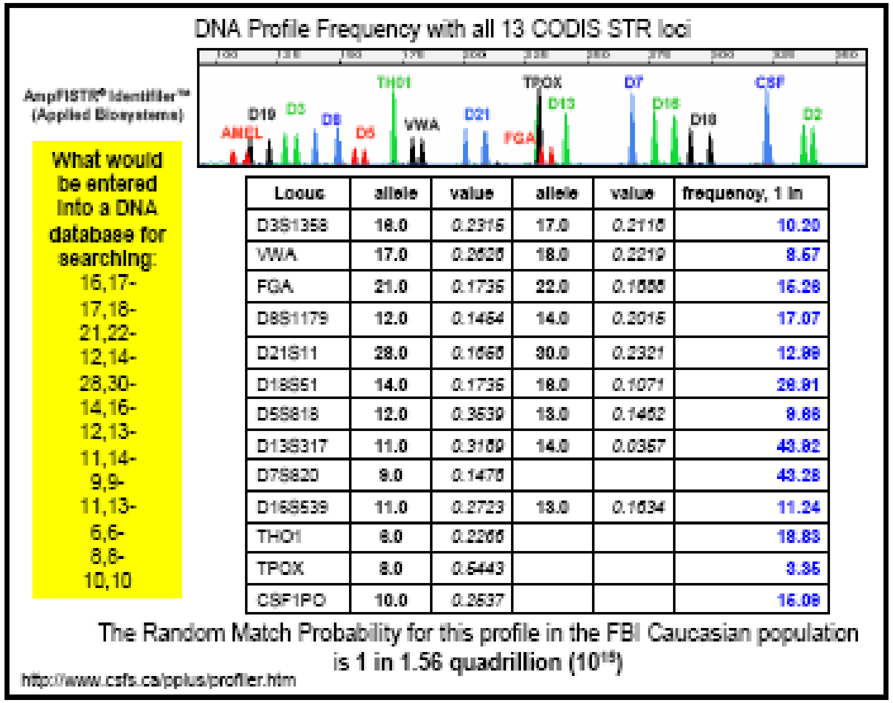
Capillary Electropherogram

Multiplex STR Analysis

AmpFlSTR® SGM Plus™ kit

Two different individuals





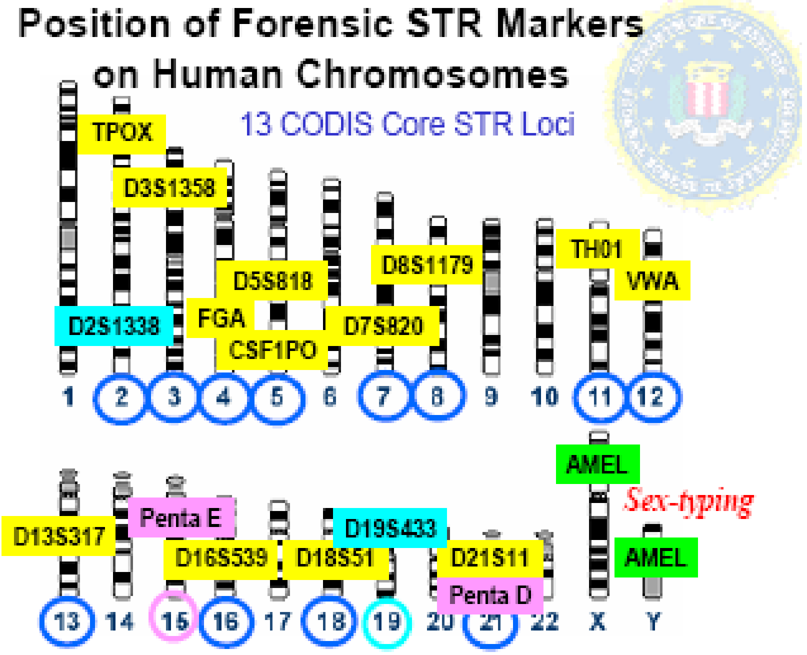
DNA Profile Frequency with all 13 CODIS STR loci

What would be entered into a DNA database for searching:
16,17-
17,18-
21,22-
12,14-
28,30-
14,16-
12,13-
11,14-
9,9-
11,13-
6,6-
8,8-
10,10

Locus	allele	value	allele	value	frequency, 1 in
D3S1358	16.0	0.2316	17.0	0.2116	10.20
VWA	17.0	0.2626	18.0	0.2219	8.57
FGA	21.0	0.1736	22.0	0.1656	16.26
D8S1179	12.0	0.1464	14.0	0.2016	17.07
D21S11	28.0	0.1666	30.0	0.2321	12.89
D18S51	14.0	0.1736	16.0	0.1071	26.91
D5S818	12.0	0.3539	13.0	0.1462	9.66
D13S317	11.0	0.3169	14.0	0.0357	43.82
D7S820	9.0	0.1476			43.28
D16S539	11.0	0.2723	13.0	0.1634	11.24
TH01	6.0	0.2266			18.83
TPOX	8.0	0.5443			3.35
CSF1PO	10.0	0.2537			16.09

The Random Match Probability for this profile in the FBI Caucasian population is 1 in 1.56 quadrillion (10^{15})

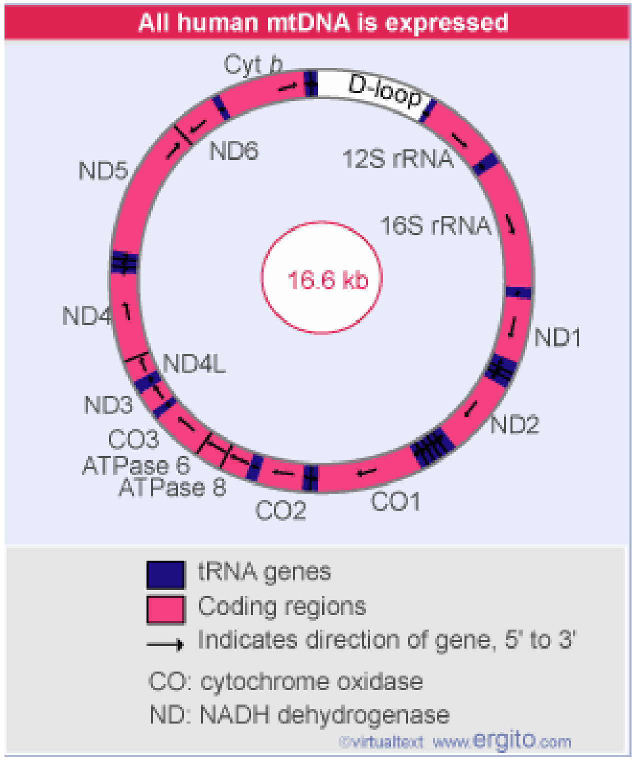
http://www.csfs.ca/pplus/profiler.htm

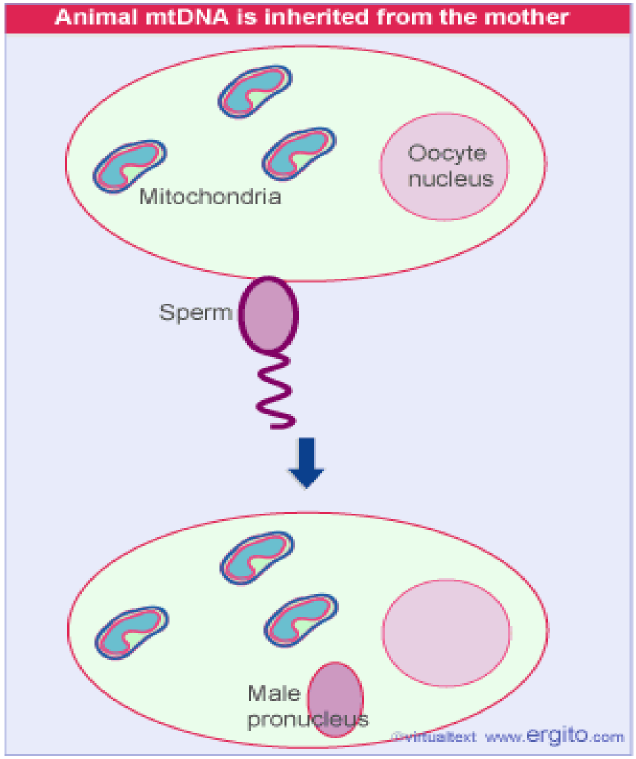

Position of Forensic STR Markers on Human Chromosomes

13 CODIS Core STR Loci

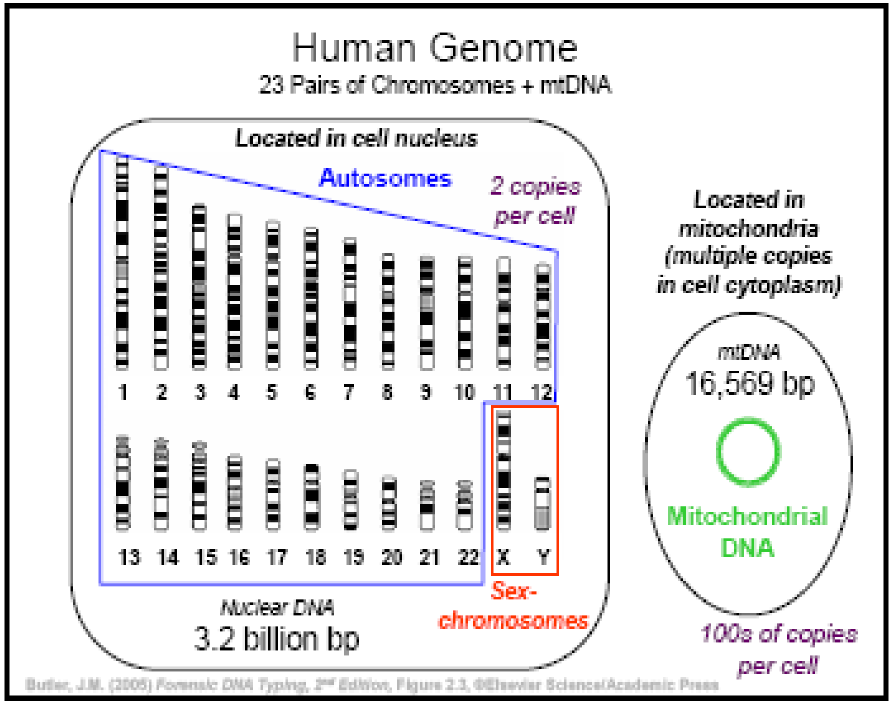
1.4.4 Extranuclear genes

Extranuclear genes reside outside the nucleus in organelles such as mitochondria and chloroplasts.





Human Genome



Butler, J.M. (2005) *Forensic DNA Typing*, 2nd Edition, Figure 2.3, ©Elsevier Science/Academic Press

Characteristics of Prokaryotic Genome

- The prokaryotic chromosome exits as a nucleoid without the membrane around.
- The structure of prokaryotic chromosome is much more simple than eukaryotic chromosome.
- Many prokaryotic gene are polycistronic (多顺反子），and some genes are overlapping (重叠）.
- Prokaryotic chromosomal DNA do not contain highly repeated sequences.
- Prokaryotic genes do not contain the noncoding introns.

Characteristics of Eukaryotic Genome

- The eukaryotic cell contains a membranous nucleus that includes the chromosome as carriers of genetic information.
- The chromosomal DNA of eukaryotes is packaged into a highly regular nucleoprotein complex called chromatin.
- A typical eukaryotic gene is monocistronic (单顺反子）.
- Many DNA segments contain highly repeated sequences.
- Eukaryotic genes are split with the noncoding introns.

断裂基因（interrupted gene）

内含子(intron)、　外显子(exon)

Summary

1. 中心法则(Central dogma)；
2. 复制(Replication)，转录(Transcription)，翻译(Translation)，逆转录(Reverse transcription)；
3. 拟核(Nucleoid)，染色体(Chromosome)，染色质(Chromatin)；
4. 核小体(Nucleosome)及其核心颗粒，组蛋白(Histones)及其修饰；
5. 基因(Gene)，外显子(Exon)，内含子(Intron)；
6. 单核苷酸多样性(SNP)，限制性片段长度多样性(RFLP)；STR

DNA与RNA的区别；DNA是遗传物质的证据； DNA作为遗传物质的特征； Watson&Crick双螺旋模型；拓扑异构酶I和II的作用机制；原核和真核基因组的主要差异