DNA as the Hereditary Molecule

"Nuclein" (later called nucleic acid) discovered by Fredrick Miescher 1869

Where did he find it?

Salmon sperm and Pus cells.

Why did he look there?

Beats me!!

Essentially ignored until the 1920s as Proteins which are more complex were thought to be more important.

Key Topics in DNA Structure Chapter 10

10.3 Evidence that DNA was the hereditary molecule

Three experiments

10.5 RNA in Viruses (we will cover this later when discussing viruses).

10.6 DNA Chemistry

Know the names of the 5 bases (Adenine, Guanine, Cytosine, Thymine (in DNA) and Uracil (in RNA). Which are purines and pyrimidines?

Know the names of dNTPs building blocks for DNA; dATP (deoxy adenosine triphosphate), dGTP, dCTP, dTTP

Be able to recognize these (as dAMP etc within DNA)

Know the difference between dNTPs and NTPs (C2 OH); mostly ATP (Adenosine triphosphate).

10.7 Structure of DNA --- Know this! Polarity of the molecules

10.8 Alternative Forms of DNA: …. Who cares!

10.9 RNA structure: We will come back to this when we discuss RNA in transcription (Ch 14)

10.10 Analytical Techniques: We will discuss some of these briefly.

Genome Organization:

How much DNA do you have and what is it doing?

Additional Pieces from around the book

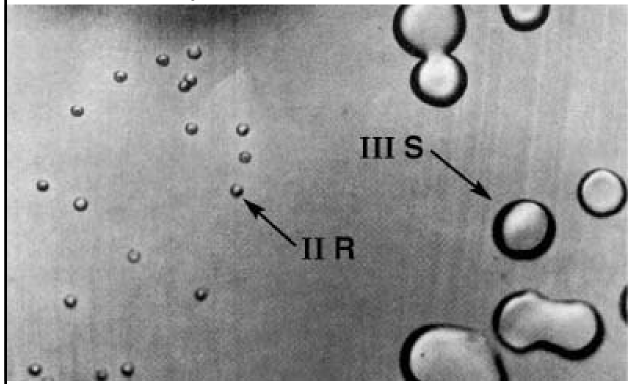
Repetitive DNA: 10.10

Repetitive DNA, Telomeres, Centromeres, SINES, LINES…:12.6, 12.7

Fragile X and microsatellites or simple sequence repeats (SSRs): 8.10 .

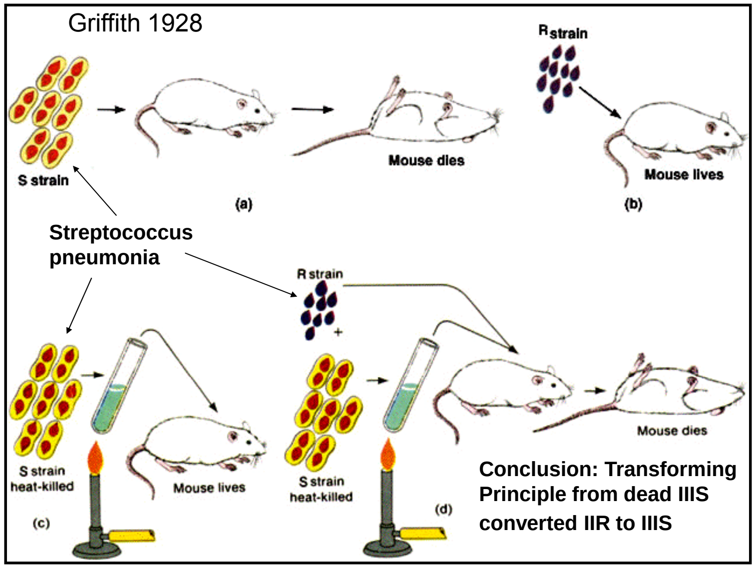
Transposons: 22.1

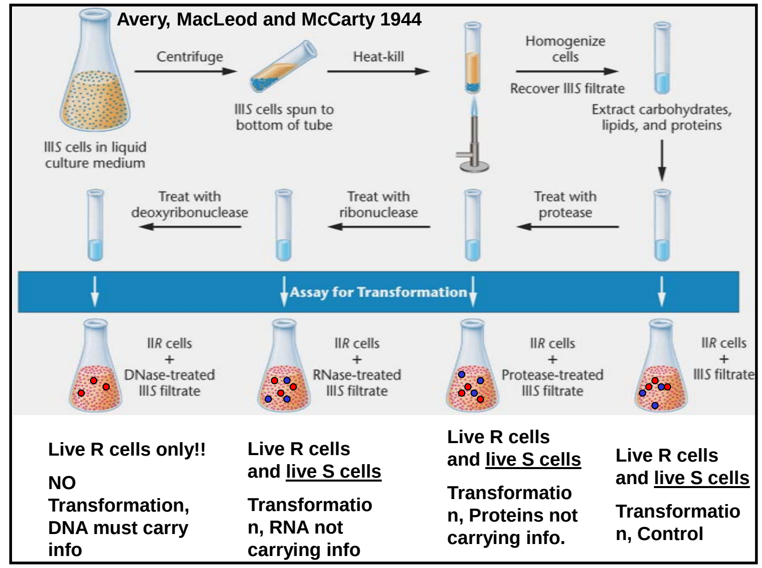
The Discovery of DNA as the Heritable Material

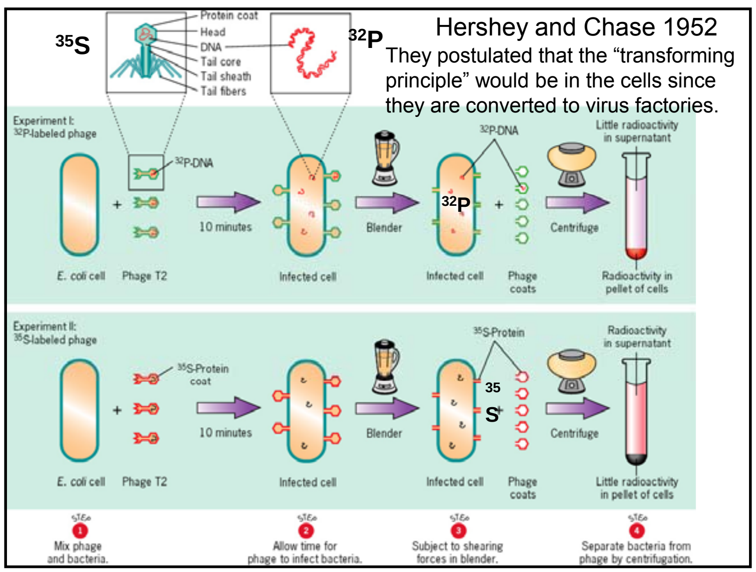


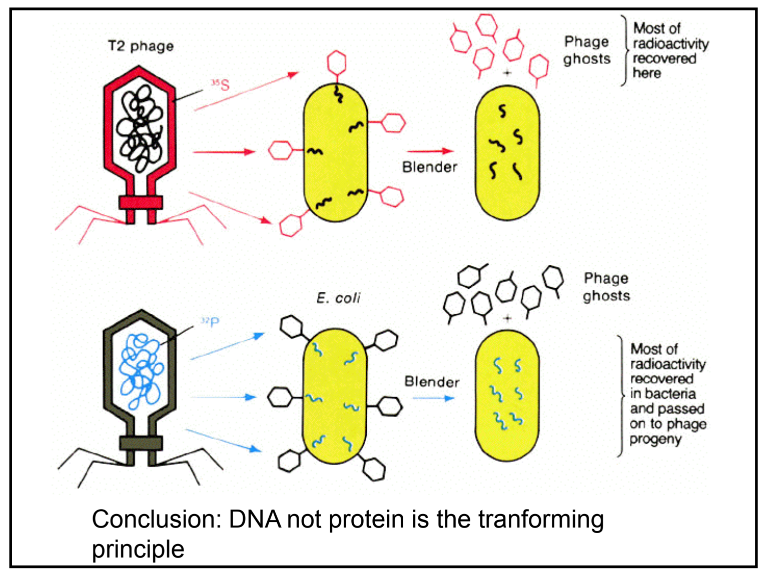
Streptococcus pneumoniae **studied by Griffith in 1928.**

IIR	rough	No polysaccharide coat	Avirulent
IIIS	smooth	Polysaccharide coat	Virulent


Griffith 1928
R strain
S strain
Mouse dies
Mouse lives
(a)
(b)
Streptococcus pneumonia
R strain
S strain heat-killed
Mouse lives
(c)
S strain heat-killed
(d)
Mouse dies
Conclusion: Transforming Principle from dead IIIS converted IIR to IIIS


Avery, MacLeod and McCarty 1944
IIIS cells in liquid culture medium
Centrifuge
IIIS cells spun to bottom of tube
Heat-kill
Homogenize cells
Recover IIIS filtrate
Extract carbohydrates, lipids, and proteins
Treat with deoxyribonuclease
Treat with ribonuclease
Treat with protease
Assay for Transformation
IIR cells + DNase-treated IIIS filtrate
IIR cells + RNase-treated IIIS filtrate
IIR cells + Protease-treated IIIS filtrate
IIR cells + IIIS filtrate
Live R cells only!!
NO Transformation, DNA must carry info
Live R cells and live S cells
Transformation, RNA not carrying info
Live R cells and live S cells
Transformation, Proteins not carrying info.
Live R cells and live S cells
Transformation, Control


35S
32P
Protein coat
Head
DNA
Tail core
Tail sheath
Tail fibers
Hershey and Chase 1952
They postulated that the "transforming principle" would be in the cells since they are converted to virus factories.
Experiment I: 32P-labeled phage
32P-DNA
10 minutes
Blender
Centrifuge
Little radioactivity in supernatant
Radioactivity in pellet of cells
E. coli cell
Phage T2
Infected cell
Phage coats
Experiment II: 35S-labeled phage
35S-Protein coat
35S-Protein
Radioactivity in supernatant
Little radioactivity in pellet of cells
Mix phage and bacteria.
Allow time for phage to infect bacteria.
Subject to shearing forces in blender.
Separate bacteria from phage by centrifugation.


T2 phage
35S
Blender
Phage ghosts
Most of radioactivity recovered here
E. coli
32P
Blender
Phage ghosts
Most of radioactivity recovered in bacteria and passed on to phage progeny
Conclusion: DNA not protein is the tranforming principle

Structure of DNA

J Watson and F Crick 1953

The Data:

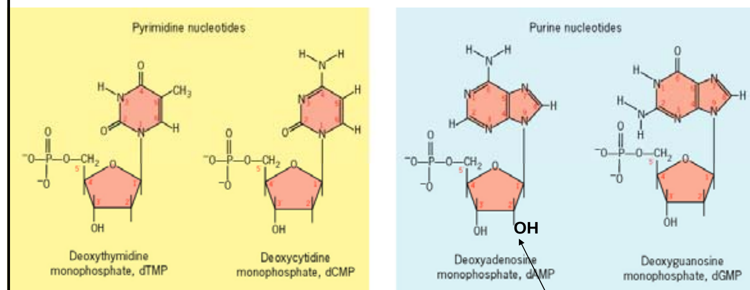
DNA composition

- sugar (deoxyribose)
- phosphate
- nitrogenous bases (adenine, thymine, cytosine, guanine)

Proportions of the bases (E. Chargoff's 1950 data)

Physical structure of DNA (R. Franklin's 1953 data)

This one page paper earned them the Nobel Prize in 1962



The building blocks of DNA are the dNTP (deoxyribose nucleotide triphosphate) forms (dATP, dCTP, dGTP, dTTP) not the dNMP shown.

The "energy" molecule is the ribose form ATP with OH at Carbon 2.

Fig 10.7 Structures of the four common deoxyribonucleotides present in DNA.

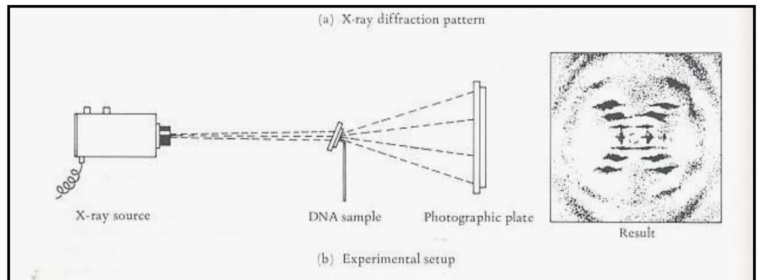
© 2003 John Wiley and Sons Publishers

Table 11-1 Molar Properties of Bases* in DNAs from Various Sources

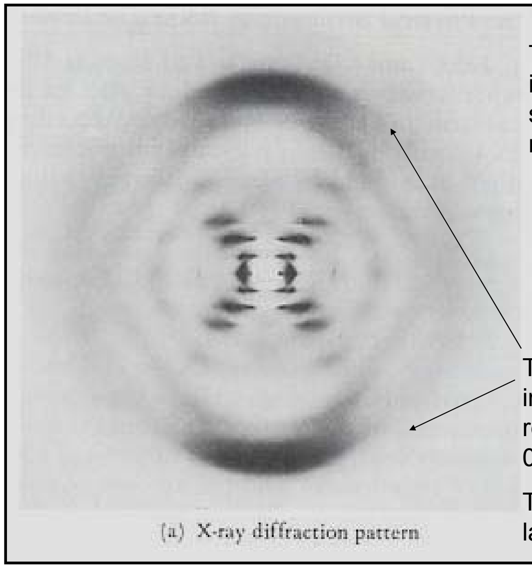
Chargaff's data

Organism	Tissue	Adenine	Thymine	Guanine	Cytosine
Escherichia coli (K12)	—	26.0	23.9	24.9	25.2
Diplococcus pneumoniae	—	29.8	31.6	20.5	18.0
Mycobacterium tuberculosis	—	15.1	14.6	34.9	35.4
Yeast	—	31.3	32.9	18.7	17.1
Paracentrotus lividus (sea urchin)	Sperm	32.8	32.1	17.7	18.4
Herring	Sperm	27.8	27.5	22.2	22.6
Rat	Bone marrow	28.6	28.4	21.4	21.5
Human	Thymus	30.9	29.4	19.9	19.8
Human	Liver	30.3	30.3	19.5	19.9
Human	Sperm	30.7	31.2	19.3	18.8

*Defined as moles of nitrogenous constituents per 100 g-atoms phosphate in hydrolysate.
Source: E. Chargaff and J. Davidson, eds., *The Nucleic Acids.* Academic Press, 1955.



Rosalind Franklin's photos of unpublished work were available to Watson and Crick. These were the best at the time and allowed for clear interpretations. Her work was also published in Nature in 1953.



The “X” pattern indicates that the structure is a regular helix

This distance indicates that the repeat pattern is 0.34 nm(3.4A)

The steps of the ladder!

How do you get the “Chargaff” pairing of bases and an even, regular helix?

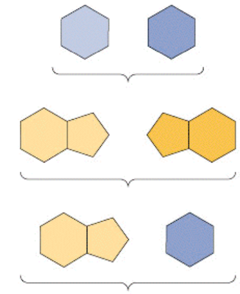
Pyrimidine + pyrimidine: DNA too thin

T and C

Purine + purine: DNA too thick

A and G

Purine + pyrimidine: thickness compatible with X-ray data



Purine + pyrimidine would give a regular structure to the DNA, and if you remember, from Chargaff, A = T and G ≡ C.

A = T

G ≡ C

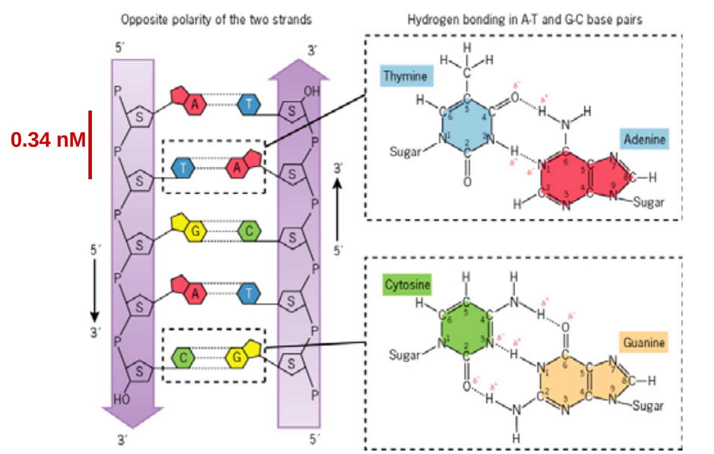
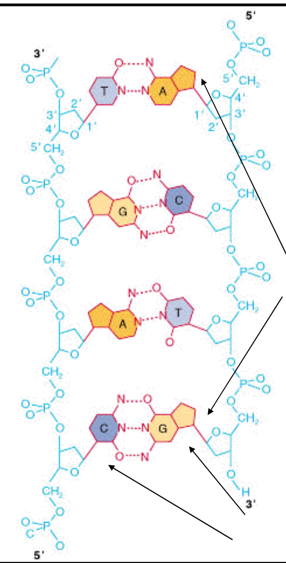


Fig. 10.12 Diagram of a DNA double helix.

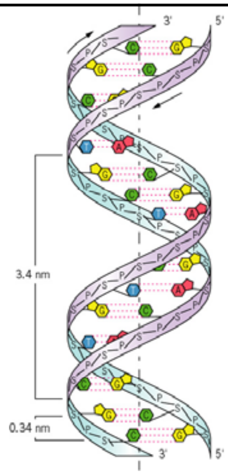
© 2003 John Wiley and Sons Publishers



Two features of the structure of DNA were obvious to Watson and Crick

No specific order of bases along the length allows for unlimited variability

Absolute specificity of bases across the molecule provides a mechanism for fidelity of replication



Watson and Crick's Model

Basic Points to Remember:

1) The two sugar-phosphate backbone chains are anti-parallel.

2) The pairing of the N bases suggests that if you know one strand you can reproduce the other (Complementarity)

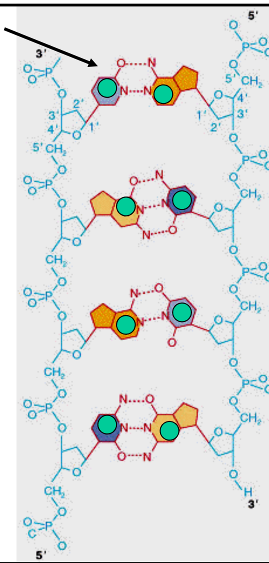
3) Pairing of fat purines and thin pyrimidines maintains constant width of specified dimensions (fits X-Ray data)

4)The order of the bases in a strand is not predictable, suggesting a means of creating variability among species

Fig 10.11 Diagram of the double-helix structure of DNA.

© 2003 John Wiley and Sons Publishers

Reading this side



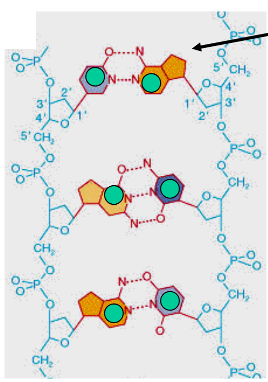
A) A G T C

B) A C T G

C) T C A G

D) G T C A

E) T G A C



Reading this side

A) A G T

B) T G A

C) G T C

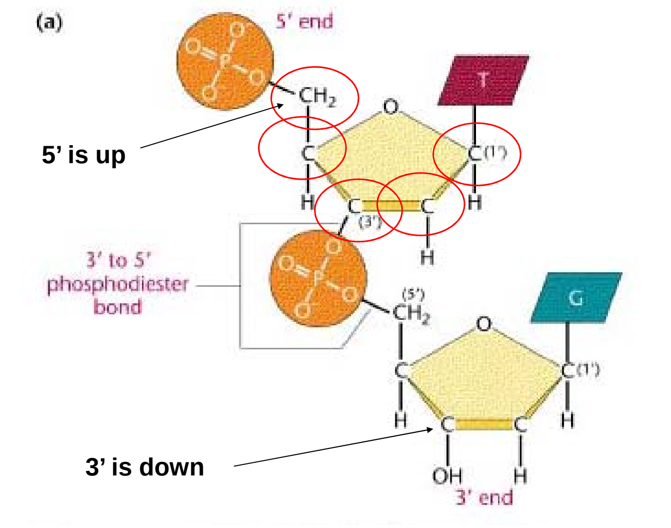
D) T C A

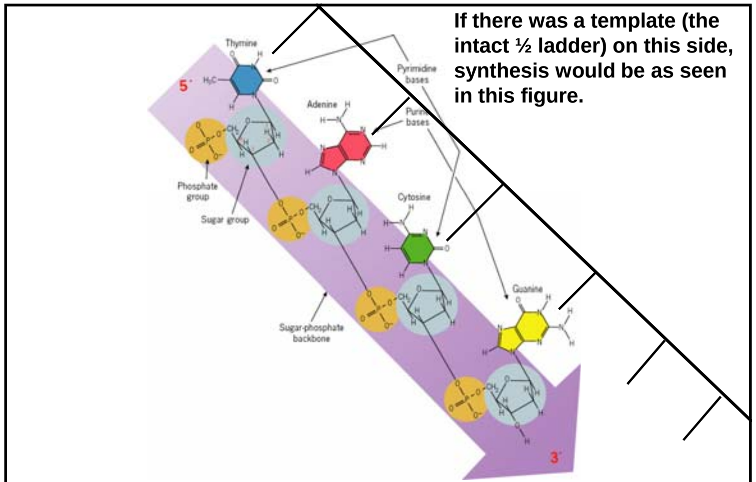
E) A C T

What do you need to know?

Identify the components of DNA

Determine the polarity of the different strands

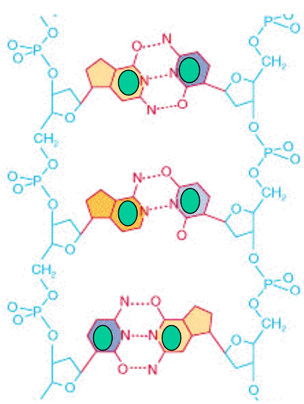




If there was a template (the intact ½ ladder) on this side, synthesis would be as seen in this figure.

Fig 10.8 Formation of a polynucleotide chain by joining nucleotides with phospodiester linkages.

© 2003 John Wiley and Sons Publishers



Which is true about the polarity in the molecule?

A) 5' is up, 3' is down

B) 5' is left, 3' is right

C) 3' is up, 5' is down

D) Left side is 5' up, 3' down

E) Right side is 5' up, 3' down

How much DNA do you have?

How much DNA, how many genes do we have?
Our genome (haploid) is about $3\text{x}10^9$ bp and we have about 25-35,000 genes.

How long is that DNA have per cell?
(0.34 nm/rung = $0.34 \text{ x } 10^{-9}$ m) x $3 \text{ x } 10^9$ bp/haploid human genome = 1 meter. Diploid cells have 2 meters.

How many cells in your body?
10 billion (10^{10}) in your brain, more than 1 trillion, 10^{12} in your body.

How much total DNA if aligned end to end?
If we go with the 1 trillion number, you have enough DNA to go to the moon (384,000,000 m away) 5200 times. Stretched out DNA is 2 nm wide so the "ribbon" of DNA would be about 10 um wide.

What is it all that DNA doing anyway?

If you have 30 - 35,000 genes and if genes (the coding region that gets translated into proteins) are about 1000 bp long, how much of your DNA is coding?

35,000 x 1000 bp = 35,000,000 bp

What proportion of your total DNA is coding?

35,000,000/ 3,000,000,000 $\cong$ 1 -2%

What is the rest of that "junk"?

Let's look at the characteristics of DNA among different groups of species to see if some interesting patterns emerge.

Relative amounts of AT vs. GC

% GC

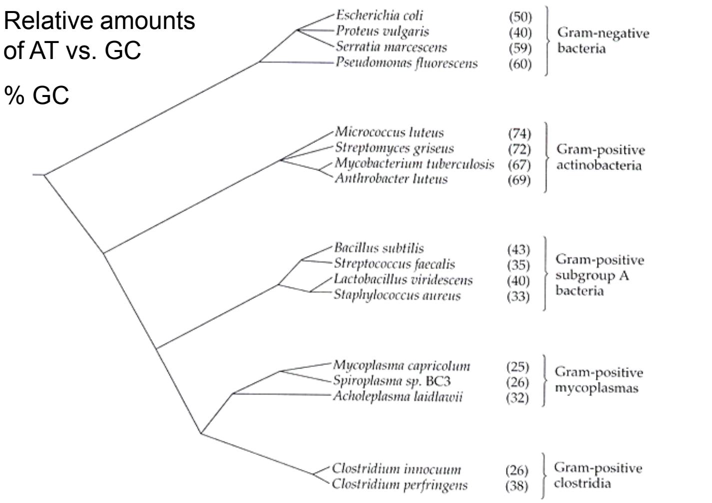
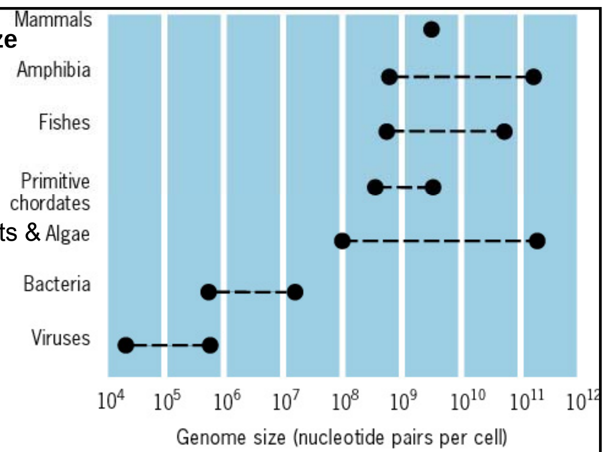


Figure 13.6 Phylogenetic tree of several eubacteria based on 5S rRNA sequences. Genomic GC contents are shown in parentheses. From Muto et al. (1986).

Genome Size Differences (haploid)



Does the size of a genome reflect the complexity of an organism?

Why do mammals have such similar genome

Different Species have characteristic DNA content

You are here

Amoebas are 200x bigger…and better?

Species	*C value (kb)*
Navicola pelliculosa (diatom)	35,000
Drosophila melanogaster (fruitfly)	180,000
Paramecium aurelia (ciliate)	190,000
Gallus domesticus (chicken)	1,200,000
Erysiphe cichoracearum (fungus)	1,500,000
Cyprinus carpio (carp)	1,700,000
Lampreta planeri (lamprey)	1,900,000
Boa constrictor (snake)	2,100,000
Parascaris equorum (roundworm)	2,500,000
Carcarias obscurus (shark)	2,700,000
Rattus norvegicus (rat)	2,900,000
Xenopus laevis (toad)	3,100,000
***Homo sapiens* (human)**	**3,400,000**
Nicotiana tabaccum (tobacco)	3,800,000
Paramecium caudatum (ciliate)	8,600,000
Schistocerca gregaria (locust)	9,300,000
Allium cepa (onion)	18,000,000
Coscinodiscus asteromphalus (diatom)	25,000,000
Lilium formosanum (lily)	36,000,000
Amphiuma means (newt)	84,000,000
Pinus resinosa (pine)	68,000,000
Protopterus aethiopicus (lungfish)	140,000,000
Ophioglossum petiolatum (fern)	160,000,000
Amoeba proteus (amoeba)	290,000,000
Amoeba dubia (amoeba)	670,000,000

Any ideas about why we see such a range in DNA content among flowering plants and fish, but not mammals or birds?

Polyploidy of course!!!

Let's look at the second major factor that has altered genome sizes.

Invasions !!

	Genome size range (kb)	Ratio (highest/lowest)
Protists	23,500–686,000,000	29,191
Euglenozoa	98,000–2,350,000	24
Ciliophora	23,500–8,620,000	367
Sarcodina	35,300–686,000,000	19,433
Fungi	8,800–1,470,000	167
Animals	49,000–139,000,000	2,837
Sponges	49,000–53,900	1
Annelids	882,000–5,190,000	6
Molluscs	421,000–5,290,000	13
Crustaceans	686,000–22,100,000	32
Insects	98,000–7,350,000	75
Echinoderms	529,000–3,230,000	6
Agnathes	637,000–2,790,000	4
Sharks and rays	1,470,000–15,800,000	11
Bony fishes	382,000–139,000,000	364
Amphibians	931,000–84,300,000	91
Reptiles	1,230,000–5,340,000	4
Birds	1,670,000–2,250,000	1
Mammals	1,420,000–5,680,000	4
Plants	50,000–307,000,000	6,140
Algae	80,000–30,000,000	375
Pteridophytes	98,000–307,000,000	3,133
Gymnosperms	4,120,000–76,900,000	17
Angiosperms	50,000–125,000,000	2,500

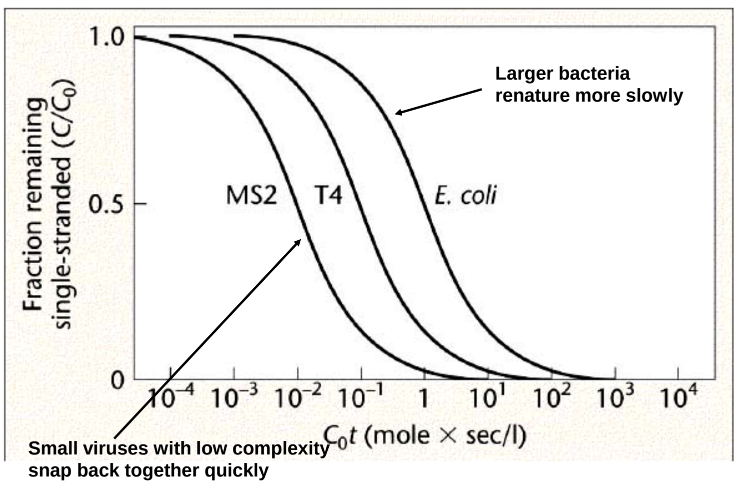
DNA complexity and Renaturation Experiments

1) Extract DNA: pieces of double stranded DNA
2) Heat DNA to denature in to single stranded pieces
3) Cool slowly and measure the rate that DNA reforms into double stranded pieces
4) Plot this process in a "Cot" (concentration-time) curve.

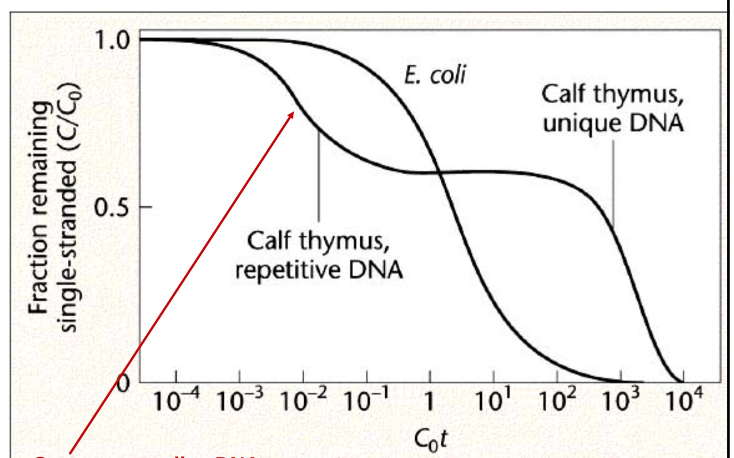
Expectations:

Smaller genomes will snap back together quickly

Genomes with lots of different genes will re-anneal slowly



What would your DNA do?



Some mammalian DNA seems to reassociate faster than E. coli DNA!!!

Why the rapid renaturation?

A major feature of Eukaryotic genomes is presence of repetitive DNA!! **10.8 and pieces of 12.6, 12.7, 22.1**

Is there a function for this repetitive DNA?

Sometimes Yes.

Some repetitive sequences are functional (rRNA genes)

Some form structural features of the chromosome

Telomeres and centromeres

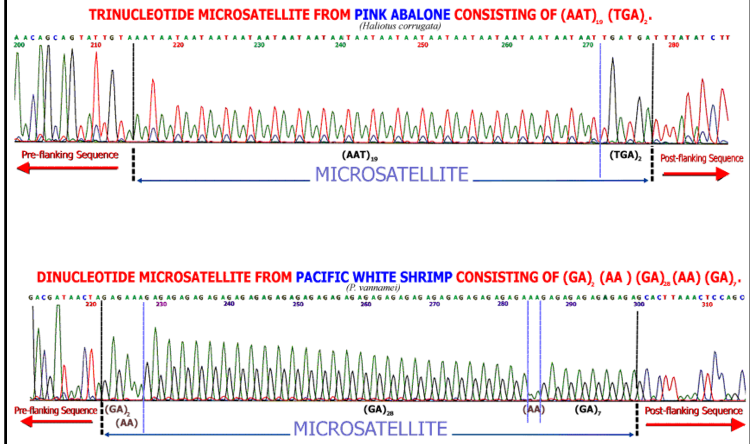
Most cases have no apparent function!

Some are caused by errors in replication

Most are likely non-functional duplicated regions (pseudogenes) or…

“DNA parasites”!!!!!

Errors during replication (DNA fingerprinting markers)



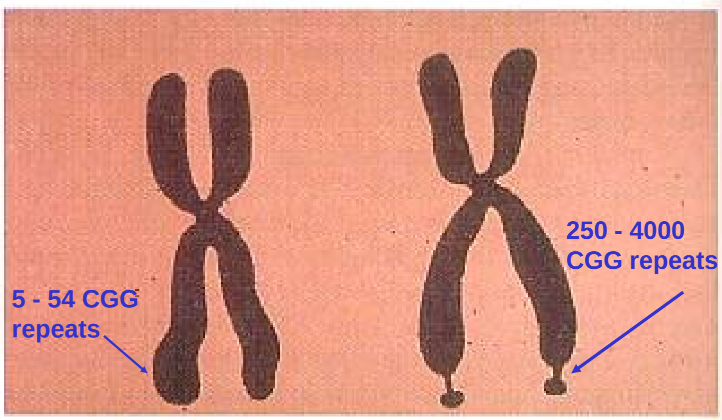
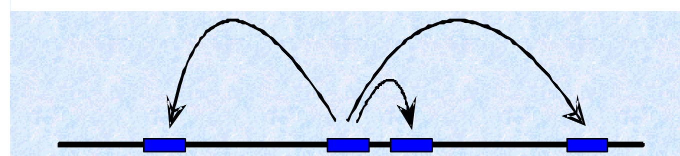


FIGURE 10.26 A normal human X chromosome (left) contrasted with a fragile X chromosome (right). The “gap” region (near the bottom of the chromosome) is associated with the fragile X syndrome.

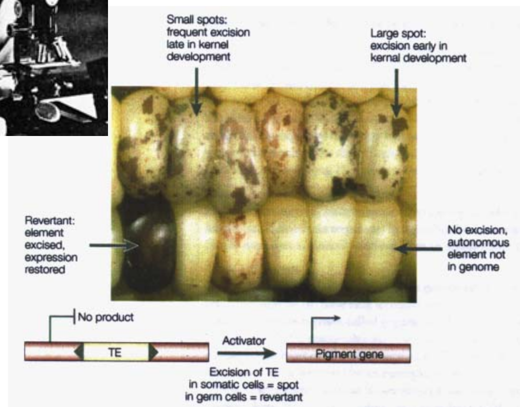
Transposable elements

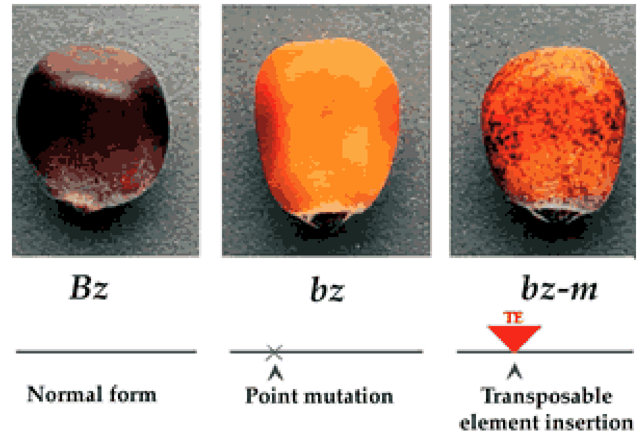
“Jumping Genes” are essentially DNA parasites



Barbara McClintock discovered the *Activator* (*Ac*) and related Transposable Elements in 1940s.

She won the Nobel Prize in 1983!!

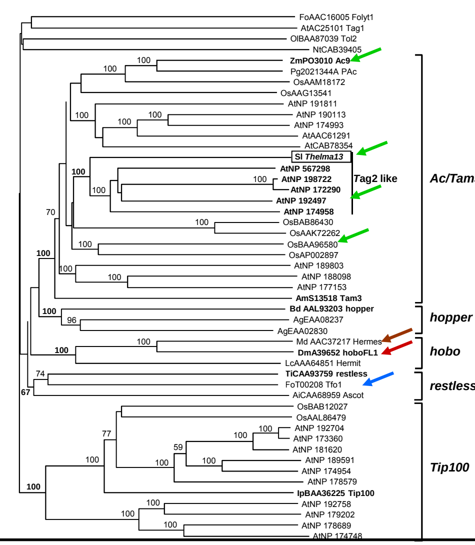




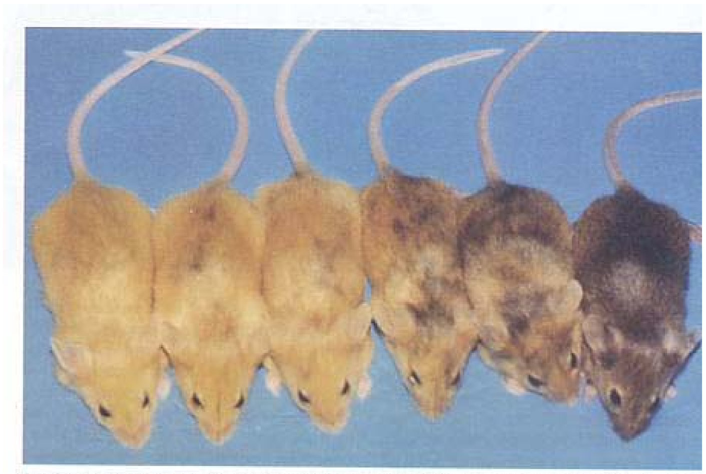
trendyscience.blogspot.com/2007/08/transposon.

An *Ac*-like Transposable Element Family With Transcriptionally Active Y-Linked Copies in the White Campion, *Silene latifolia*

Pritham, Zhang, Feschotte and Kesseli 2003

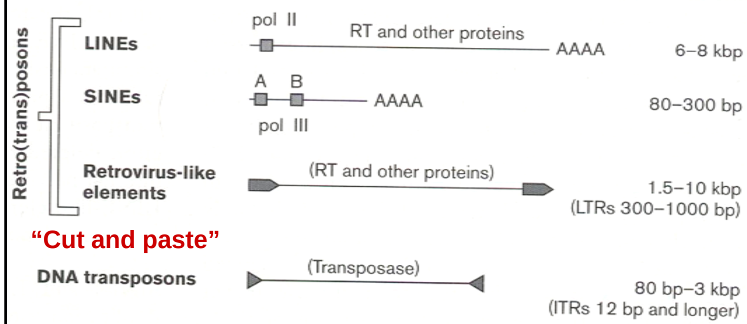


Maize/Corn
Silene
Arabidopsis
Rice
Drosophila
Mouse
Fungi



Say cheese: just one piece of 'junk' DNA can produce several coat colours in genetically identical mic[e]

Two classes of transposable elements



Selfish, parasitic or junk DNA

Interspersed repeat composition of the human genome.

CG level and total size database entries	SINEs *Alu*	SINEs MIR	LINEs LINE1	LINEs LINE2	Elements with LTRs HERVs	Elements with LTRs MaIRs	Elements with LTRs others	DNA transposons *mariner*	DNA transposons others	Unclassified elements	Total
Extrapolation to a 3 billion bp genome											
Copy number (in thousands)	1188	402	593	271	50	167	34	8	192	60	2969
Fraction of total genome	10.0%	1.7%	14.6%	2.1%	1.3%	2.6%	0.7%	0.1%	1.5%	0.8%	35.5%

Copy and paste — Cut and paste

Repetitive DNA

Functional Coding

Multigene families ex. Golbin genes

ribosomal RNA genes

Functional Structural

Centromere repeats ex. In mammals ≈ 171bp repeat (not completely conserved)

Telomere repeats ex. In mammals 5'**TTAGGG**TTAGGGTTAGGG…3'

Non-functional

Replication errors such as simple sequence repeats (aka microsatellites) ex. …TGTGTGTGTG…
…ACACACACAC…

Non-functional "parasitic" DNA

DNA Transposons ex. Activator (Ac) in Maize and other species

Retro Transposons ex. SINEs such as Alu element in humans