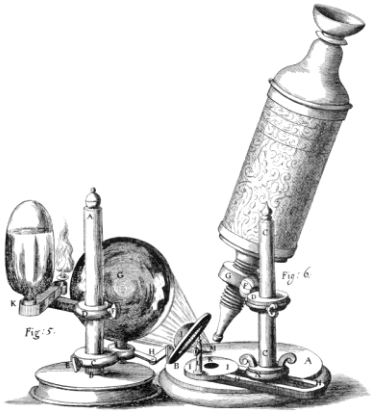
Introduction to Molecular Biology

Part 1

Discovery of Cells

- Microscopic biology began in 1665
- Robert Hooke (1635-1703) discovered organisms are made up of cells



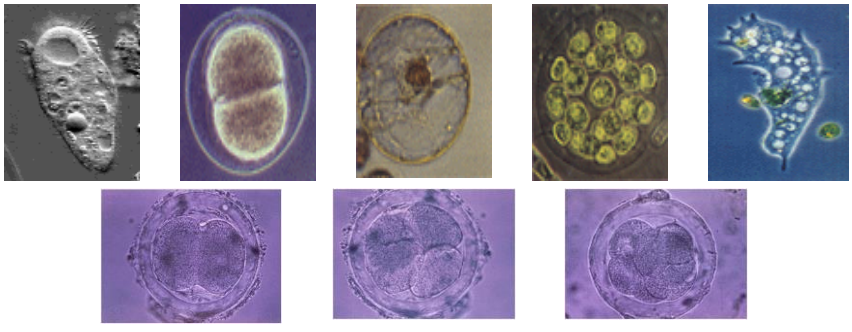
2

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 living cells are the result of 3.5 billion years of evolution

3

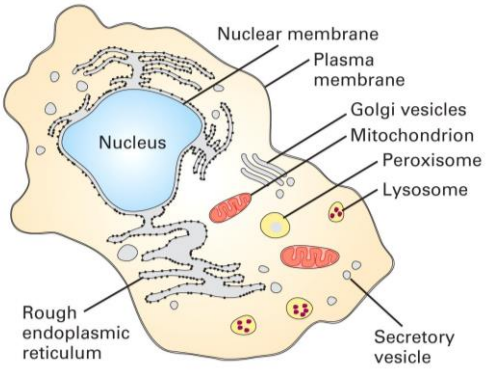
Cells



4

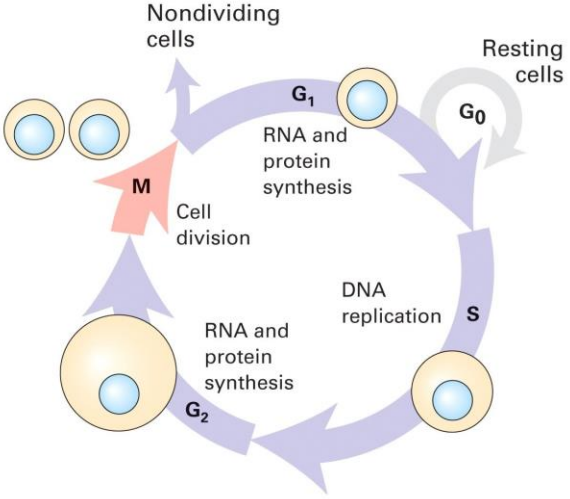
Cells

- A cell is a smallest structural unit of an organism that is capable of independent functioning
- All cells have some common features



5

Cell Cycles: Born, Eat, Replicate, Die



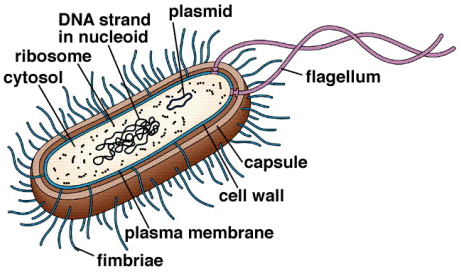
6

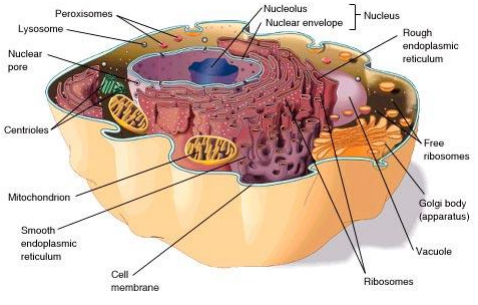
Cell Contents

- 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

7

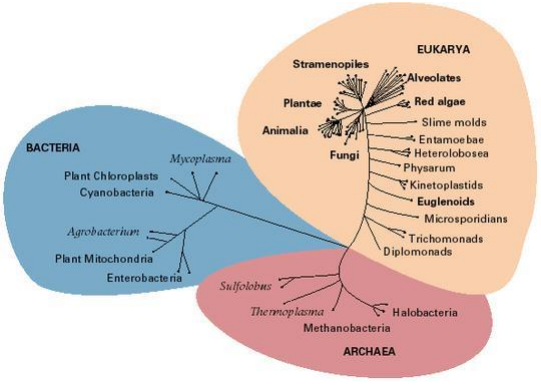
Cells: Prokaryotes & Eukaryotes





8

Prokaryotes & Eukaryotes

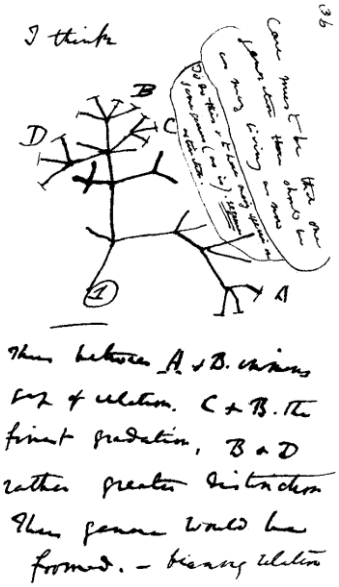


9

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



10

Prokaryotes & Eukaryotes

Prokaryotes	Eukaryotes
➢ Single Cell	➢ Single or multi cell
➢ No nucleus	➢ Nucleus
➢ No organelles	➢ Organelles
➢ One piece of Circular DNA	➢ Chromosomes
➢ No mRNA post-transcriptional modification	➢ Exons/Introns splicing

11

Prokaryotes & Eukaryotes

Prokaryotes	Eukaryotes
➢ Eubacteria (blue green algae) and archaebacteria	➢ Plants, animals, Protista, fungi
➢ Only one type of membrane Plasma membrane forms the boundary of the cell	➢ Complex systems of internal membranes forms organelle and compartments
➢ The smallest cells known are bacteria e.g. E. Coli cell, $3 * 10^6$ protein molecules, 1000-2000 polypeptide species	➢ Volume of cell is several hundred times larger e.g. Hela cell, $5 * 10^9$ protein molecules, 5000-10000 polypeptide species

12

Prokaryotes & Eukaryotes

Prokaryotes	Eukaryotes
➢ The genome of E. Coli contains $4 * 10^6$ base pairs	➢ The genome of yeast contains $1.35 * 10^7$ base pairs
➢ >90% of DNA encode protein	➢ A small fraction of the DNA encodes protein (many repeats of non-coding sequences)
➢ Lacks a membrane bound nucleus Circular DNA and supercoiled domain	➢ All chromosomes are contained in a membrane bound nucleus (DNA is divided between one or more chromosomes)
➢ Histones are unknown	➢ A set of five histones DNA packaging and gene expression regulation

13

Signalling 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

14

Cells: Information & Machinery

- Cells store all information to replicate
  - Human genome is around $3 * 10^9$ 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 that could mine for ore, build cars, replicate itself!

15

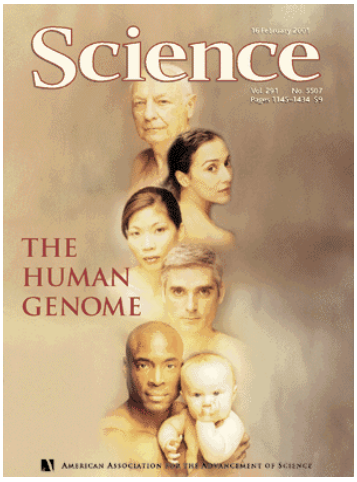
The Human Genome Project

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

16

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**



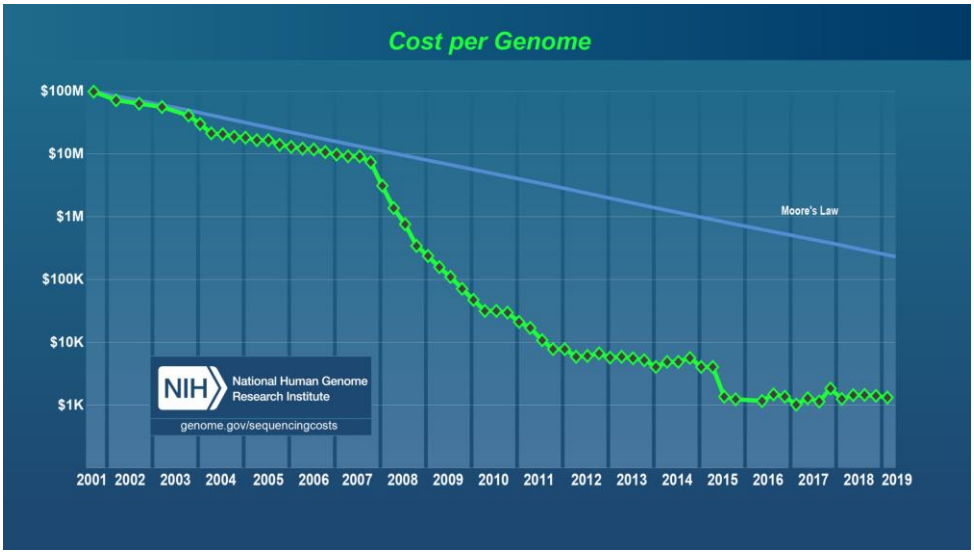
17

The Human Genome Project

- False start…until **2003**
  [Over 92% sampling with 99.99% accuracy]
- **2009** More accurate version by Genome Reference Consortium
  Still had some 300 gaps
- **2019** Still has 89 gaps
- **2025** Estimated that 1 billion individual genomes sequenced

18

Cost of Sequencing



19

But, what about 23andme??

20

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

21

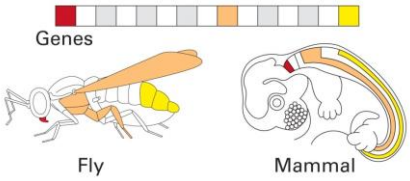
Terminology

- **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

22

Genotype & Phenotype

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



- On the right, is the eye's phenotypes of green and black eye genes.

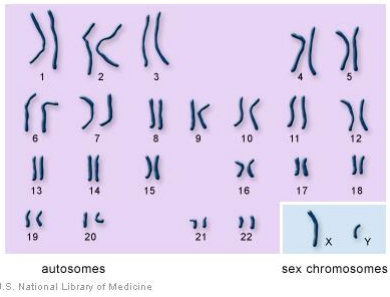
23

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

24

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 46 chromosomes (23 pairs)
  Each chromosome contains many **genes.**
- **Genes** are the basic functional units of heredity.
  Specific sequences of DNA bases encode instructions on how to make **proteins.**
- **Proteins** make up the cellular structure.
  Large complex molecules made up of smaller subunits called **amino acids.**

25

Life: Three 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 perform biochemical reactions
  Send signals to other cells and regulate gene activity
  Form body's major components (e.g. hair, skin, etc.)

26

Life: Three Critical Molecules

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

LIBRARY

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

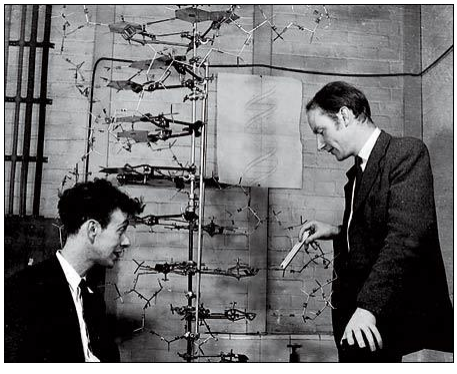
- **Proteins**
  Form enzymes that perform biochemical reactions
  Send signals to other cells and regulate gene activity
  Form body's major components (e.g. hair, skin, etc.)

PERFORM THE ACTUAL WORK OF THE CELL

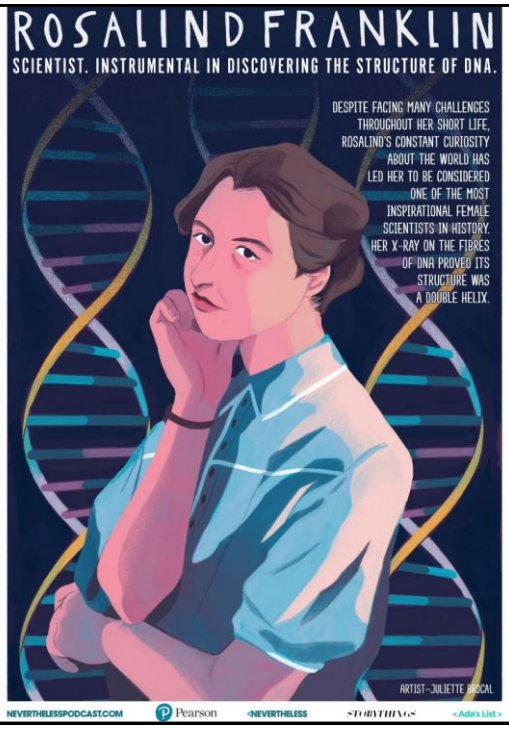
27

DNA: The Double Helix

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



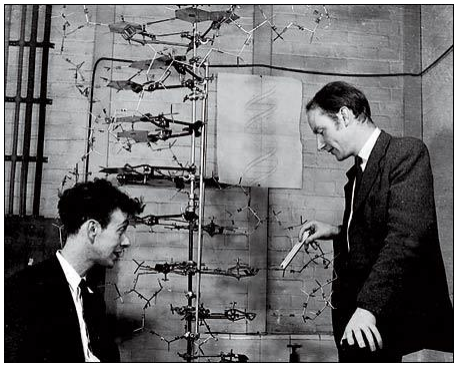
28



29

DNA: The Double Helix

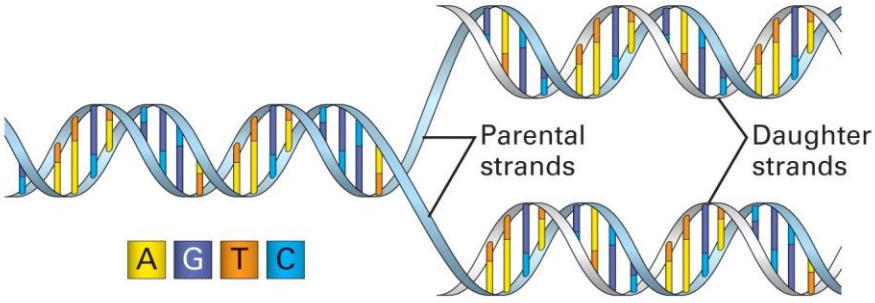
- **1952-1953** James D. Watson, Rosalind Franklin, and Francis H. C. Crick deduced the double helical structure of DNA



30

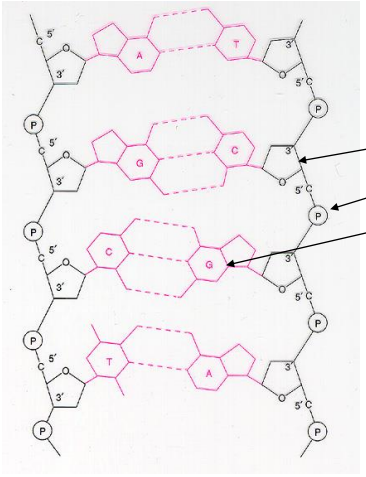
DNA: The Code of Life

- The structure and the four genomic letters code for all living organisms
- **A**denine, **G**uanine, **T**hymine, and **C**ytosine which pair A-T and C-G on complimentary strands.



31

DNA: The Code of Life



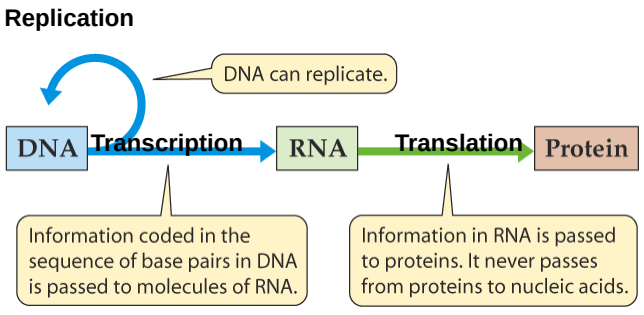
- DNA has a double helix structure which composed of
  - sugar molecule
  - phosphate group
  - and a base (A,C,G,T)
- DNA always reads from 5' end to 3' end for transcription replication

  5' ATTTAGGCC 3'

  3' TAAATCCGG 5'

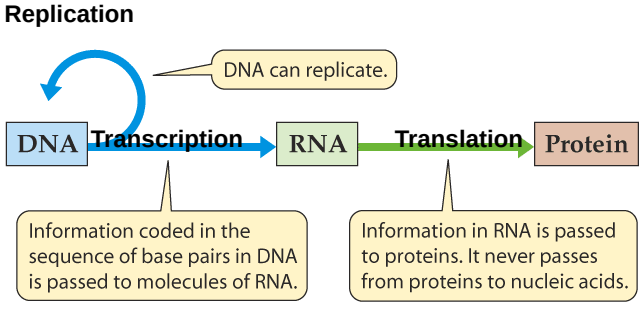
32

DNA &RNA: Flow of Information



33

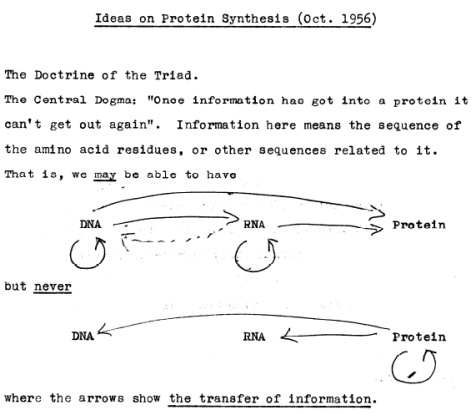
DNA & RNA: Flow of Information



aka "The Central Dogma"!!

34

Francis Crick



Ideas on Protein Synthesis (Oct. 1956)

The Doctrine of the Triad.

The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it.

That is, we may be able to have

but never

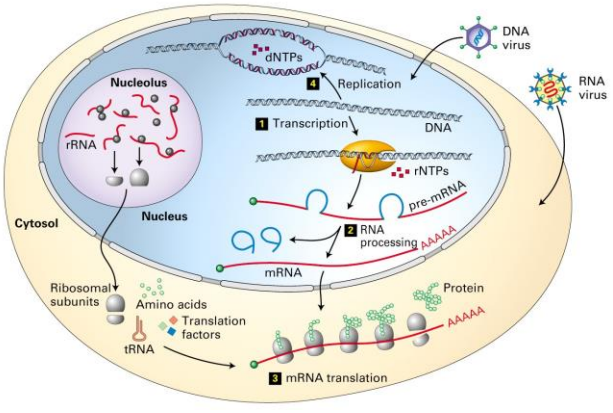
where the arrows show the transfer of information.

"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."

-: Francis Crick, *Central Dogma of Molecular Biology*, Nature, Volume 227, August 1970.

35

DNA to RNA to Protein



A gene is expressed in two steps

1. **Transcription:** RNA Synthesis
2. **Translation:** Protein Synthesis

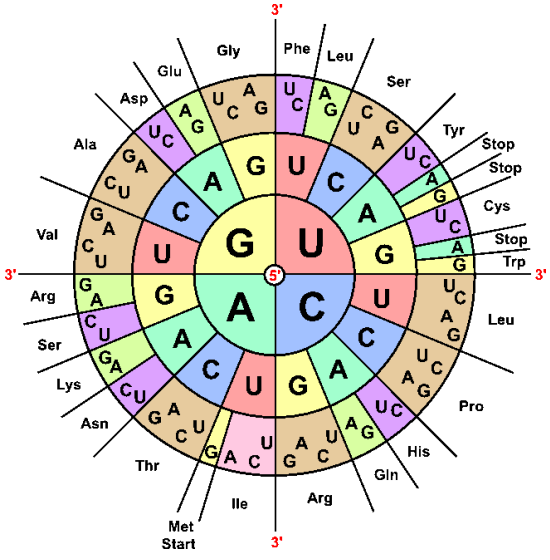
36

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 amino acids that make up proteins. Each amino acid is coded by 3 nucleotides called **codons**

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU, UUC: Phenyl-alanine	UCU, UCC, UCA, UCG: Serine	UAU, UAC: Tyrosine	UGU, UGC: Cysteine	U C A G
	UUA, UUG: Leucine		UAA, UAG: Stop codon Stop codon	UGA: Stop codon; UGG: Tryptophan	
C	CUU, CUC, CUA, CUG: Leucine	CCU, CCC, CCA, CCG: Proline	CAU, CAC: Histidine	CGU, CGC, CGA, CGG: Arginine	U C A G
			CAA, CAG: Glutamine		
A	AUU, AUC, AUA: Isoleucine	ACU, ACC, ACA, ACG: Threonine	AAU, AAC: Asparagine	AGU, AGC: Serine	U C A G
	AUG: Methionine; start codon		AAA, AAG: Lysine	AGA, AGG: Arginine	
G	GUU, GUC, GUA, GUG: Valine	GCU, GCC, GCA, GCG: Alanine	GAU, GAC: Aspartic acid	GGU, GGC, GGA, GGG: Glycine	U C A G
			GAA, GAG: Glutamic acid		

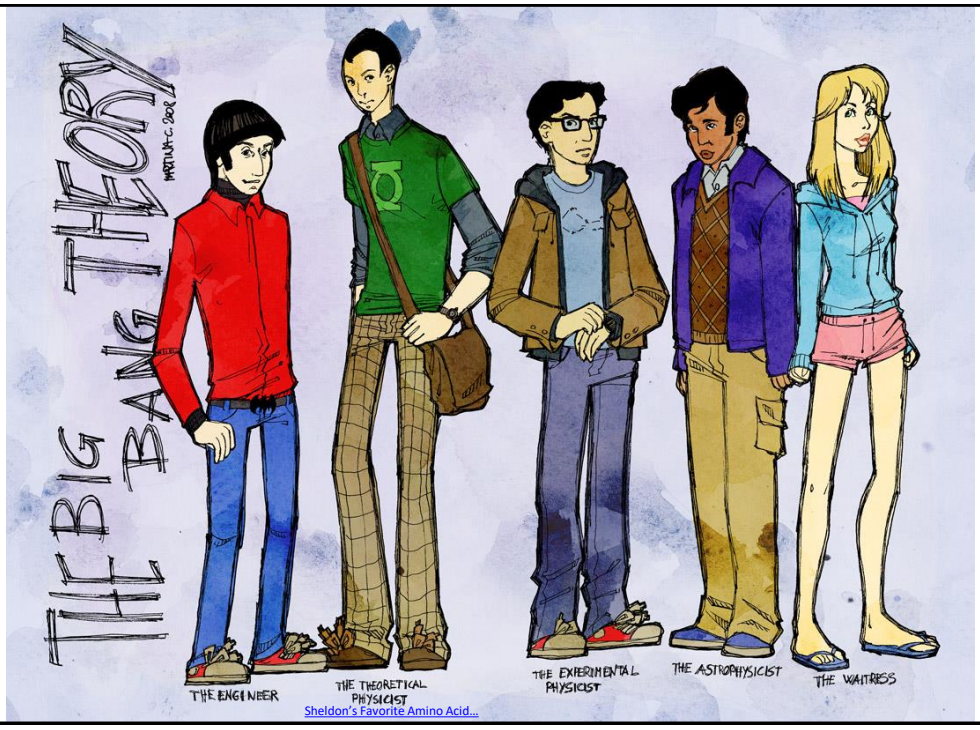
The Code Book



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

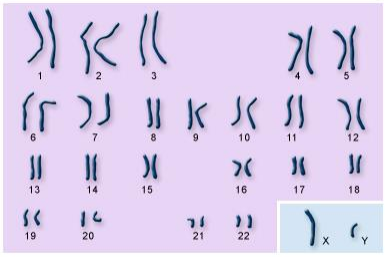
39



Sheldon's Favorite Amino Acid...

40

Sidebar…a little Bryn Mawr History



41

Edmund Wilson

- Professor of Biology at Bryn Mawr from 1885 to 1891.

  Wrote the first definitive book on the cell: Wilson E.B. 1896; 1900; 1925. *The cell in development and inheritance*. Macmillan. The third edition ran to 1232 pages, and was still in use after World War II.

  He also discovered the chromosomal XY sex-determination system in 1905—that males have XY and females XX sex chromosomes.

42

Thomas Morgan

- Professor of Biology at Bryn Mawr from 1890 to 1904 (Joined Edmund Wilson at Columbia)

43

Lilian Sampson

- Lilian Sampson started at Bryn Mawr as an undergraduate student in 1887. She majored in biology and was advised by M. Carey Thomas.

  After her graduation with honors in 1891, she spent the summer at the Marine Biological Laboratory in Woods Hole, Massachusetts, where Edmund Wilson, one of her previous zoology professors, introduced her to her future graduate advisor and husband Thomas Morgan (whom she married in 1904).

  Graduated with a Masters in Biology in 1894.

  Made seminal contributions to the genetics of fruit fly.

44

Nettie Stevens

- Earned her B.A. in Biology in 1899, and M.A. in 1900 from Stanford University.

  In 1900, Stevens enrolled at Bryn Mawr to do a PhD with advisor Thomas Morgan. Awarded PhD in Biology in 1903.

  Discovered the role of chromosomes in sex determination.

45

Nettie Stevens

- Although Stevens and Wilson both worked on chromosomal sex determination, many authors have credited Wilson alone for the discovery.

  Wilson did not realize how significant the small (Y) chromosomes are for sex determination until Stevens had completed and published her research.

  Before he read her papers, he had believed that environmental factors played a role in sex determination.

  Additionally, Thomas Hunt Morgan has been credited with the discovery of sex chromosomes although at the time of these discoveries, he argued against Wilson's and Stevens' interpretations.

  Stevens was not even recognized immediately after her discovery. For example, Morgan and Wilson were invited to speak at a conference to present their theories on sex determination in 1906 but Stevens was not invited to speak.

46

Nettie Stevens: Google Doodle, July 2016

What are the genetics behind gender?

This simple question drove the work of American geneticist Nettie Stevens. Building on research by Edmund Beecher Wilson and Thomas Hunt Morgan at Bryn Mawr, Stevens discovered the connection between chromosomes and physicality. Her breakthrough evolved into the XY sex-determination system, now taught in classrooms around the world.

Doodle by Lydia Nichols

47

References

- N. C. Jones & P. A. Pevzner, *An Introduction to Bioinformatics Algorithms*, MIT Press 2004.
- Some material adapted from slides posted at the web site of the above book.
- Francis Crick, *Central Dogma of Molecular Biology*, Nature, Volume 227, August 1970.
- Luciano Floridi, *Information: A Very Short Introduction*, Oxford 2010.
- Matthew Cobb. Sexism in science: did Watson and Crick really steal Rosalind Franklin's data? The Guardian, 2015. From: https://www.theguardian.com/science/2015/jun/23/sexism-in-science-did-watson-and-crick-really-steal-rosalind-franklins-data
- Wikipedia: Nettie Stevens. From: https://en.wikipedia.org/wiki/Nettie_Stevens
- Wikipedia: Lilian Vaughn Morgan. https://en.wikipedia.org/wiki/Lilian_Vaughan_Morgan
- Wikipedia: Thomas Hunt Morgan. https://en.wikipedia.org/wiki/Thomas_Hunt_Morgan
- Wikipedia: Edmund Beecher Wilson. https://en.wikipedia.org/wiki/Edmund_Beecher_Wilson

48