Bioinformatics
CS300
Crash course:
Structure and Replication of DNA

Fall 2019
Oliver BONHAM-CARTER

What is Bioinformatics?

- Clinical informatics
  - Systems used to deal with patient health
  - Clinical trial management systems, electronic health records, etc.
- Laboratory information
  - Systems to deal with scientific instrumentation and data management
  - Connecting instruments together, managing laboratory flow, etc.
- Bioinformatics
  - Systems to deal with basic research data
  - DNA, proteins, *molecular* things

Why Need Bioinformatics?

- Able to fight disease with medicine
- Avoid types of disorders by helping people make better health decisions before it’s too late
- To process the massive amounts of health data that already exists.
  - What can we learn from our experiments?
  - Can we incorporate data into computer models to be run instead of using animal models?
  - Other themes...

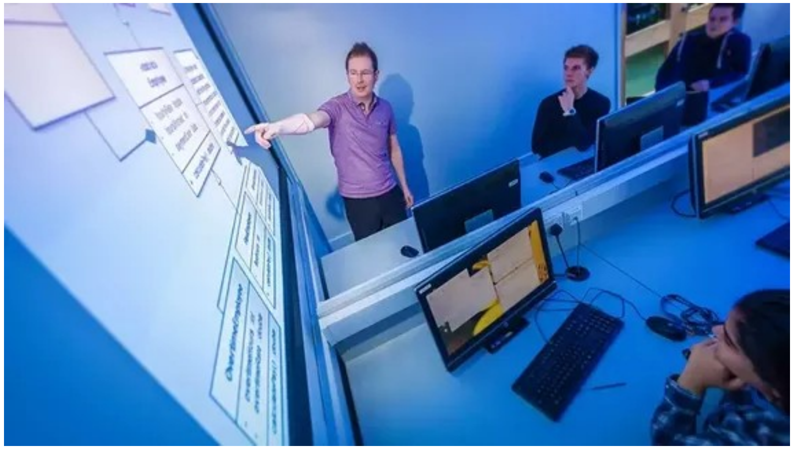
Study for Skills in Careers

- Biologists:
  - Computational skills
  - Mathematical /statistical
  - Programming for Automation

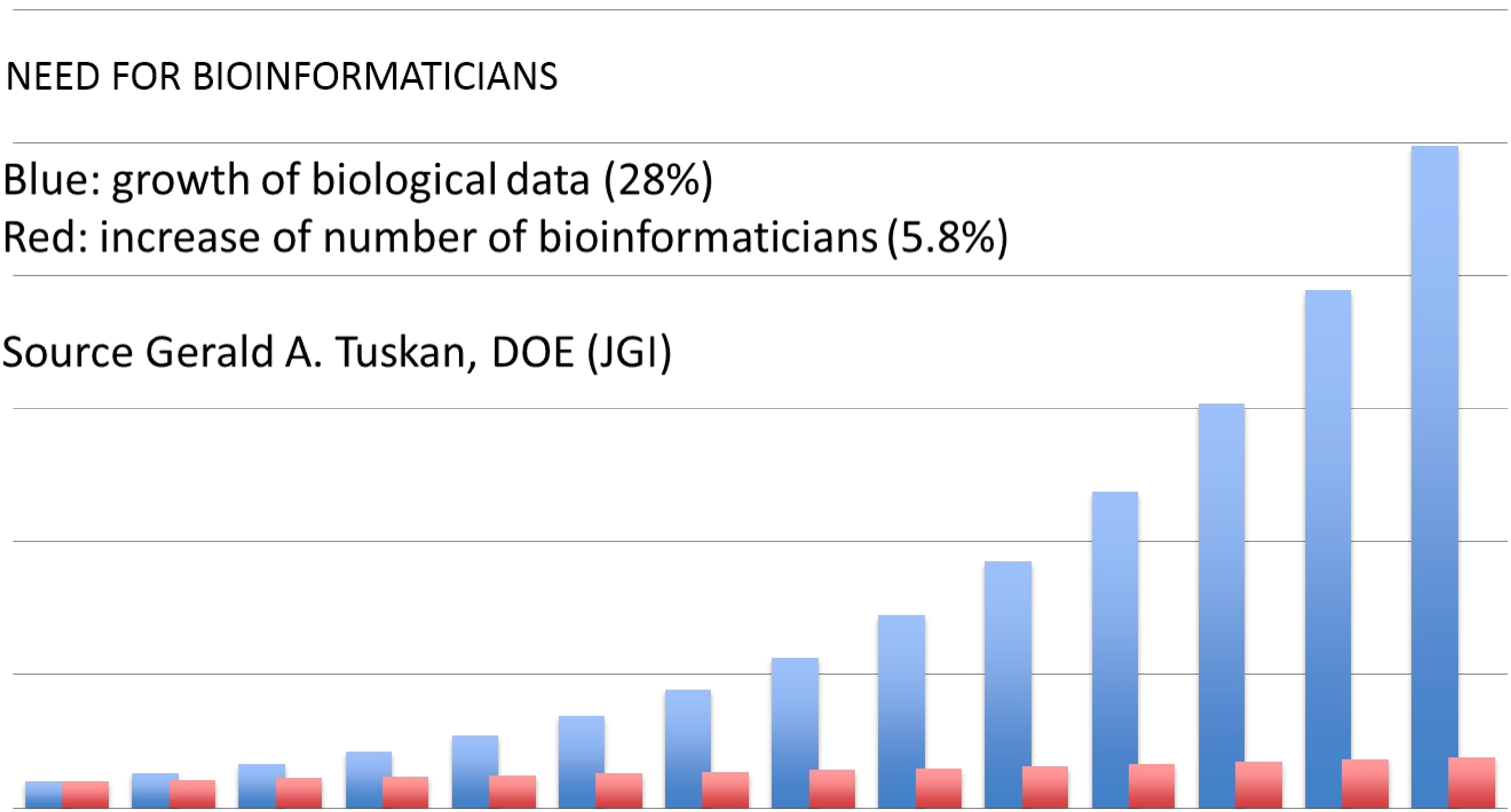
- Computer scientists
  - BioMedical skills
  - Understanding of biological systems and mechanisms
  - Early detection of disease by data
  - Modeling of therapeutic remedies
  - Others

Career Ideas

- Software (bioinformatics) engineer
- Research scientist in biotechnology
- Data scientist
- Project manager (pharmaceuticals, medical, etc)
- Computational immunologist
- Medical doctor (in clinical and research applications)



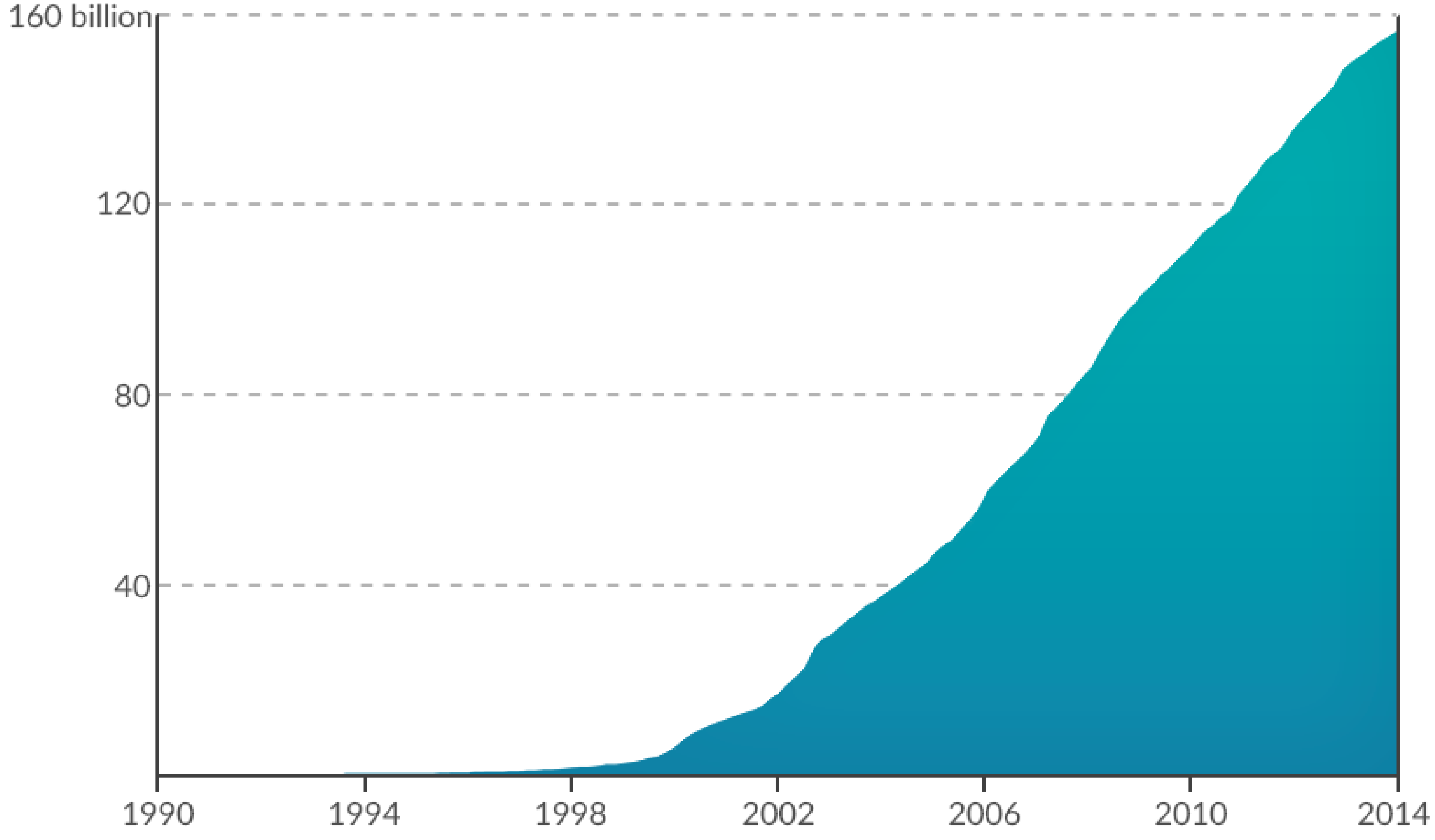
They Need You Out There!



Up to Present Day

Rising Amounts of Data to Study

- Exponential Growth of NIH base pairs through December 2013
- Data is to be assembled into sequences (jigsaw puzzle pieces).



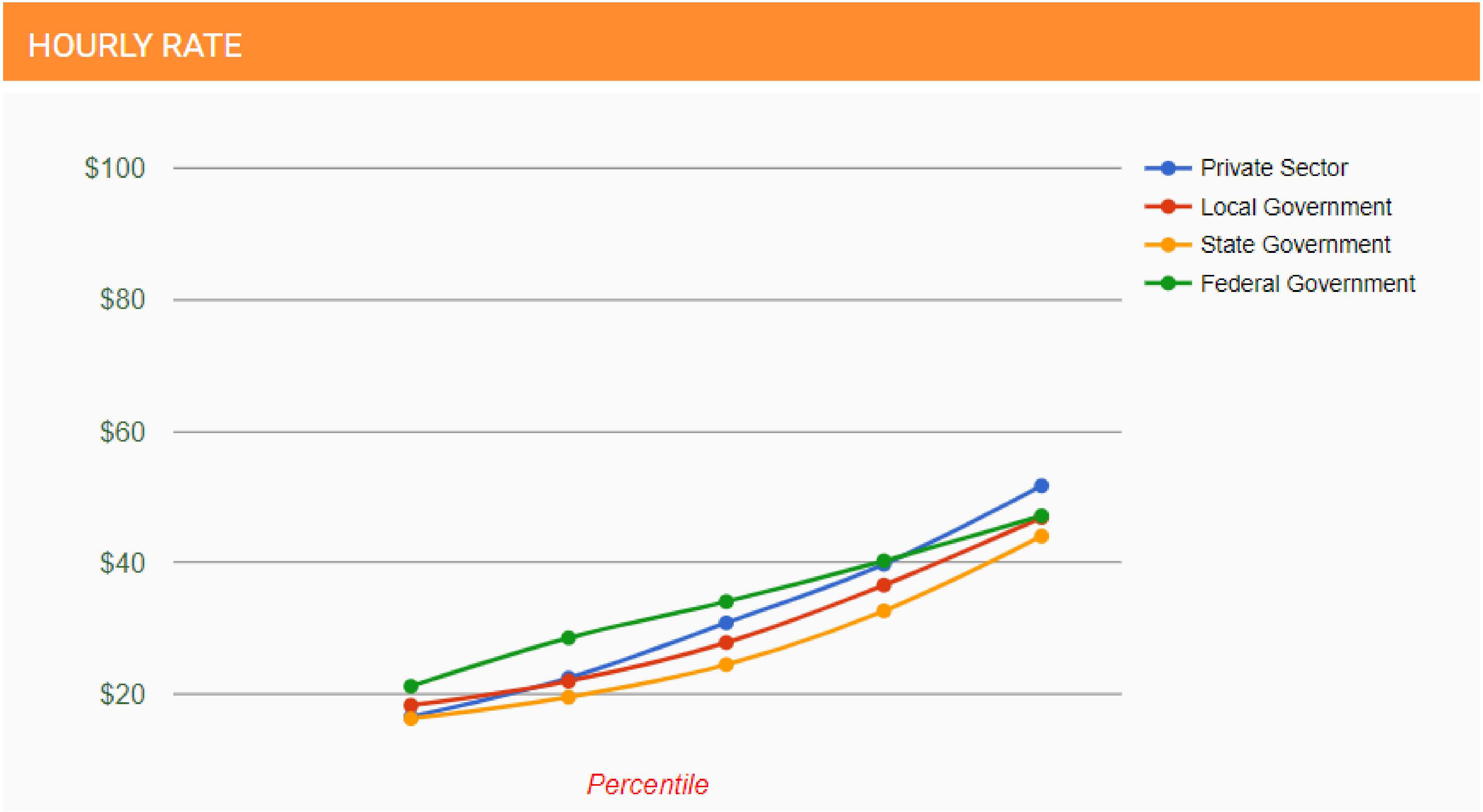
Researchers and Scientists

Bioinformatics Scientists are generally graduate-educated research scientists whose work involves the development of computer and technology-aided solutions to problems in biochemistry and biological research.

Primary focus areas may include genomics and proteomics.These professionals may be required to create, maintain and utilize databases of complex biodata,and utilize exsiting publicly available databases containing similar information.

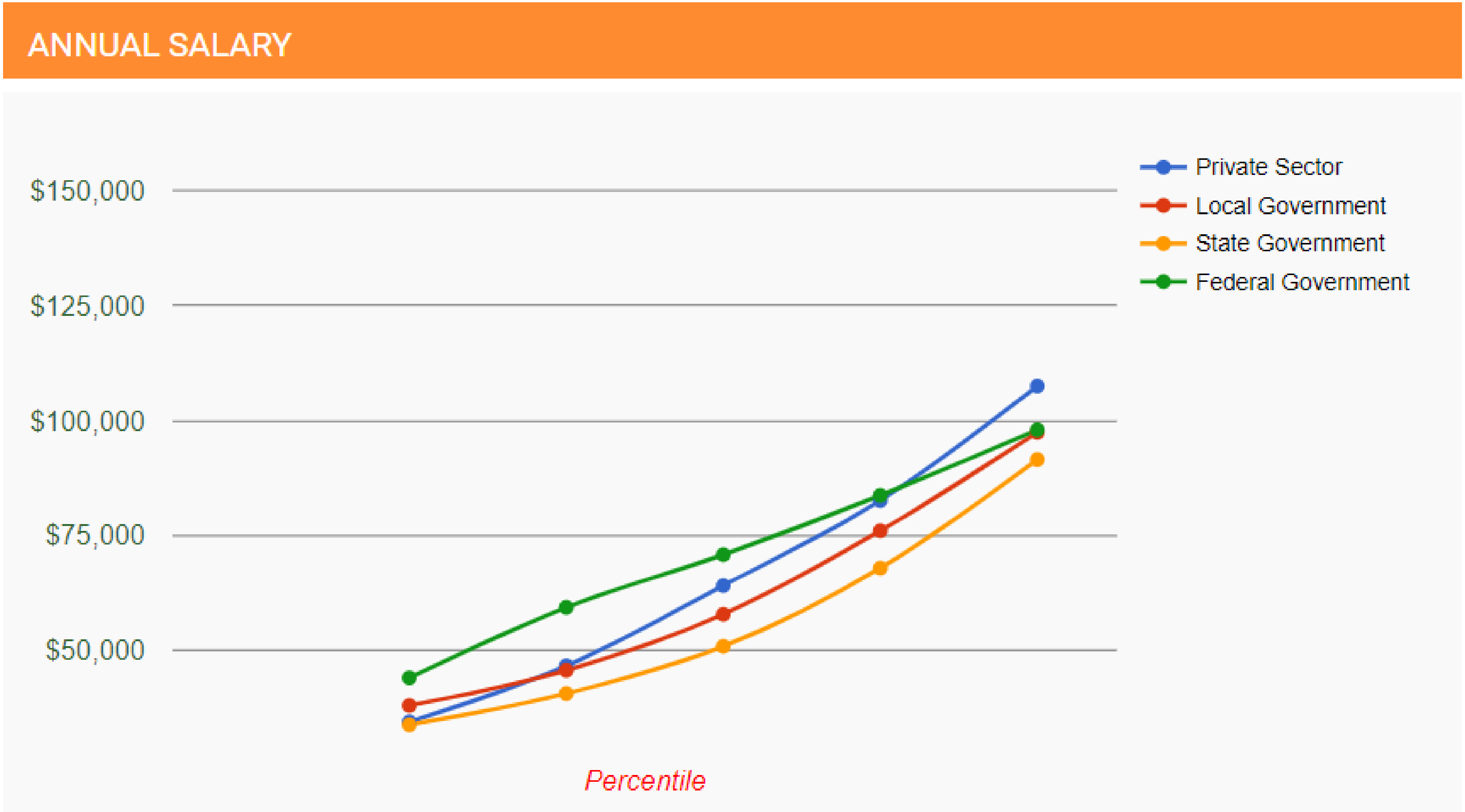
Go online to read more about careers!
https://www.recruiter.com/salaries/bioinformatics-scientists-salary/
Aug 2019

Wages



https://www.recruiter.com/salaries/bioinformatics-scientists-salary/
Aug 2019

Salaries

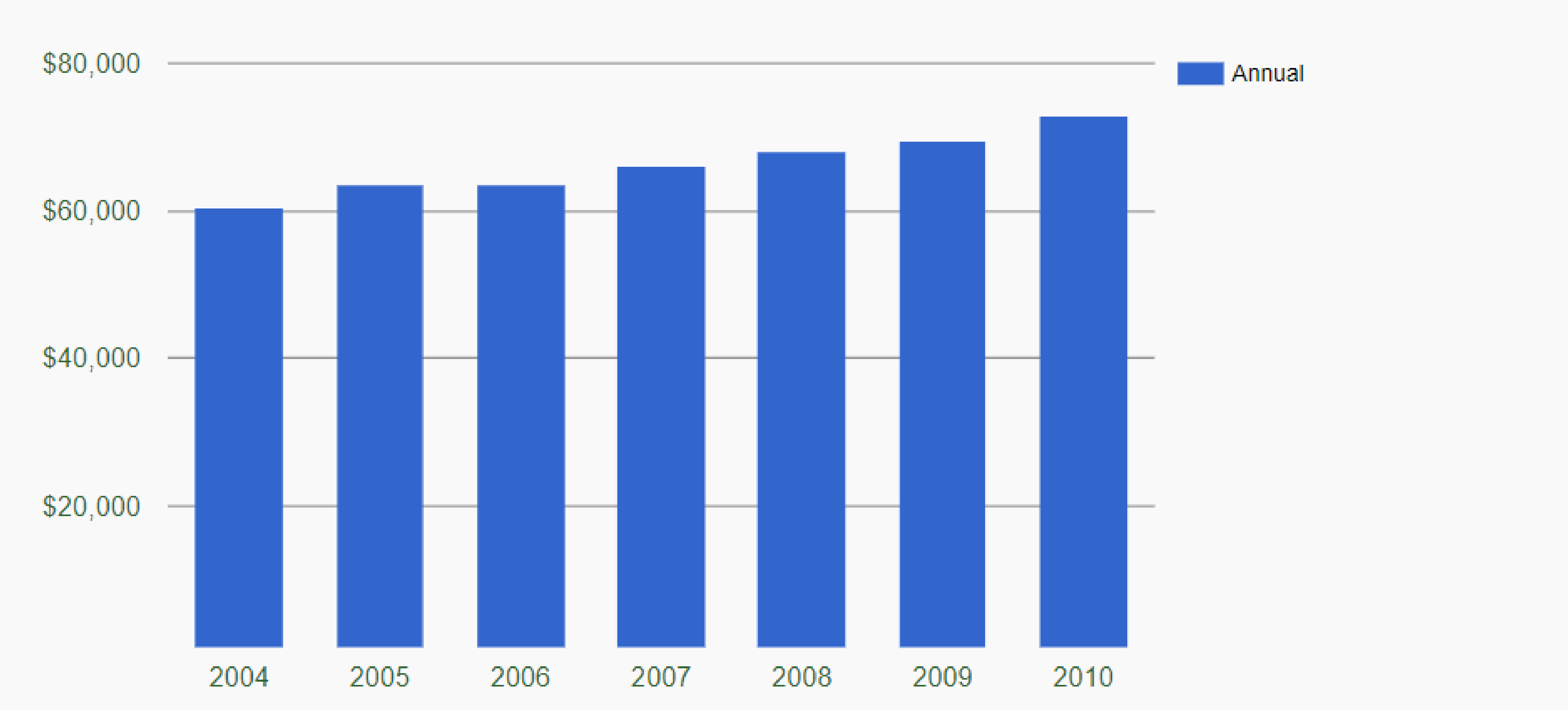


https://www.recruiter.com/salaries/bioinformatics-scientists-salary/
Aug 2019

Trends

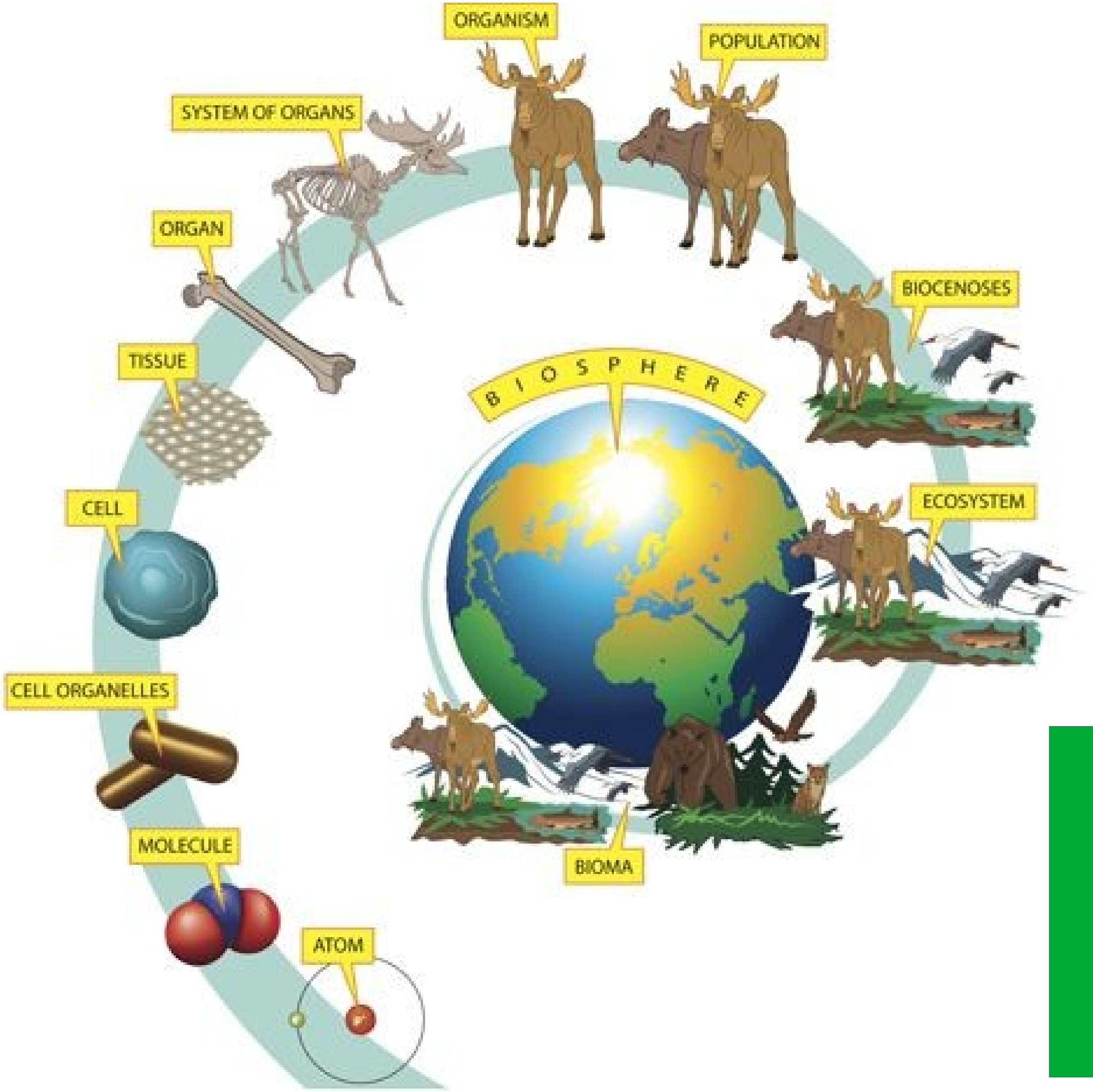
SALARY TREND

The annual compensation for this career has gone up since 2004. Salaries have increased by an average of 20.87 percent nationwide in that time.



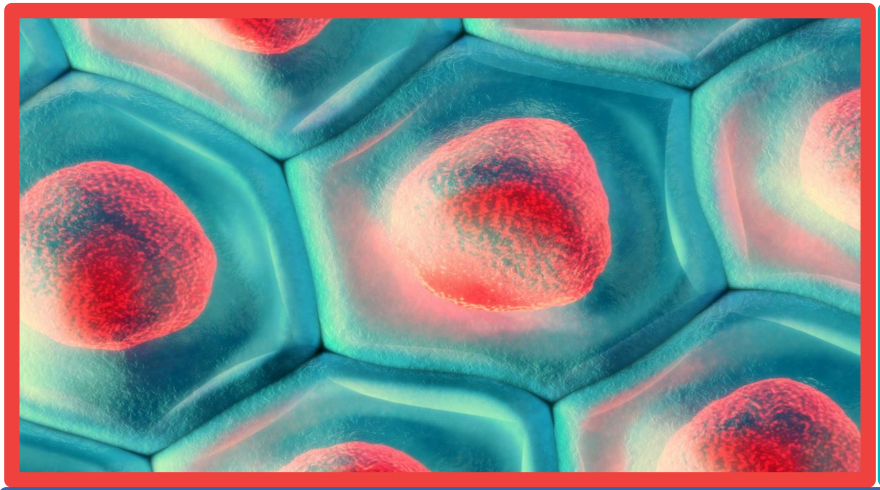
https://www.recruiter.com/salaries/bioinformatics-scientists-salary/
Aug 2019

Some Background in Biology



What's a MAJOR Commonality Here?

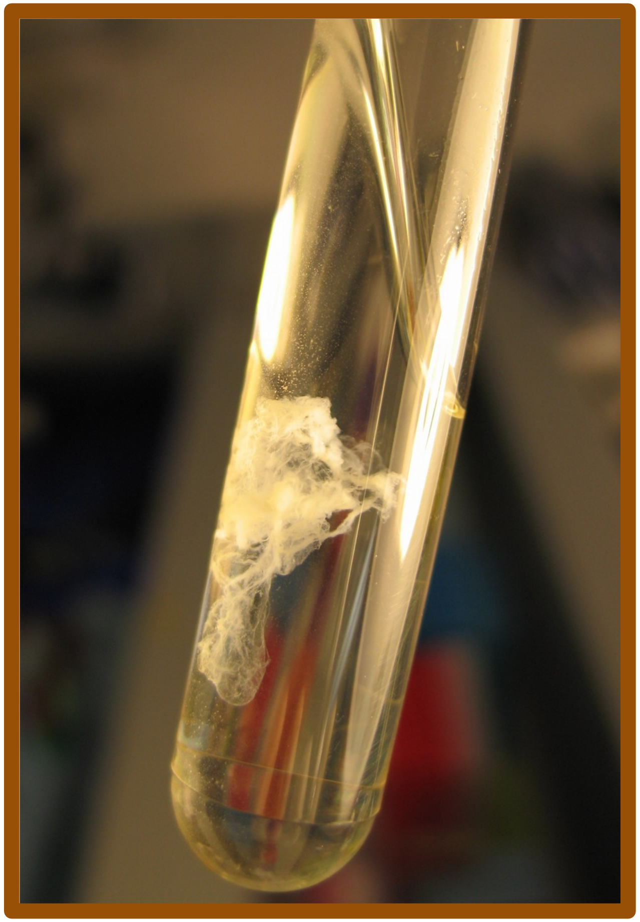
Organisms Have Genetic Systems!
And DNA is Often the Genetic Language



What Does Natural DNA look Like?

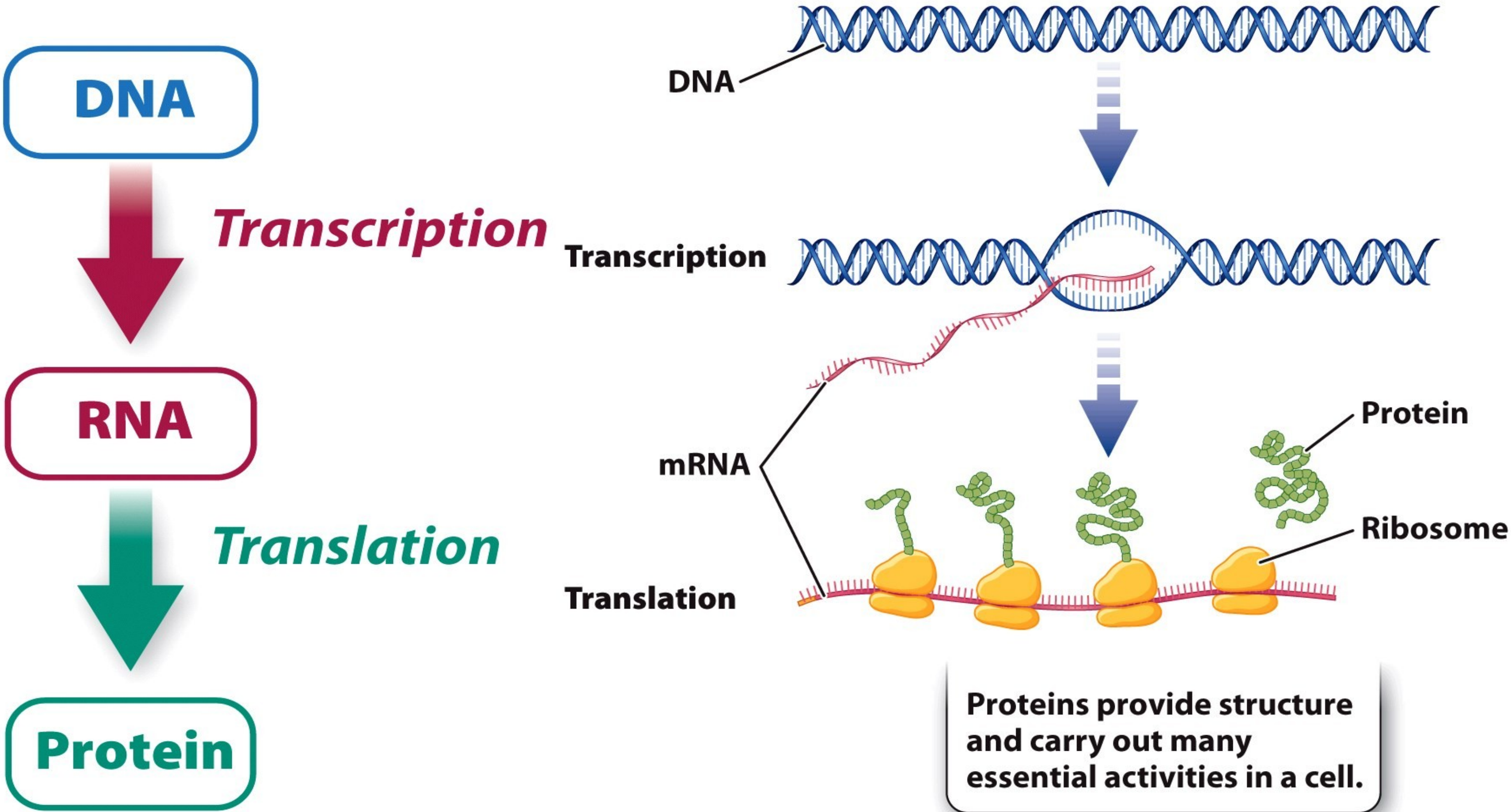
NOT like this!

Like this!



Like this!

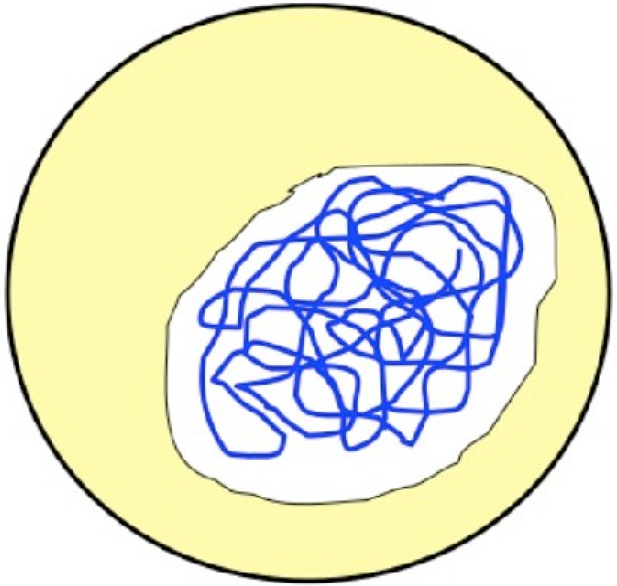
Central Dogma of Molecular Biology

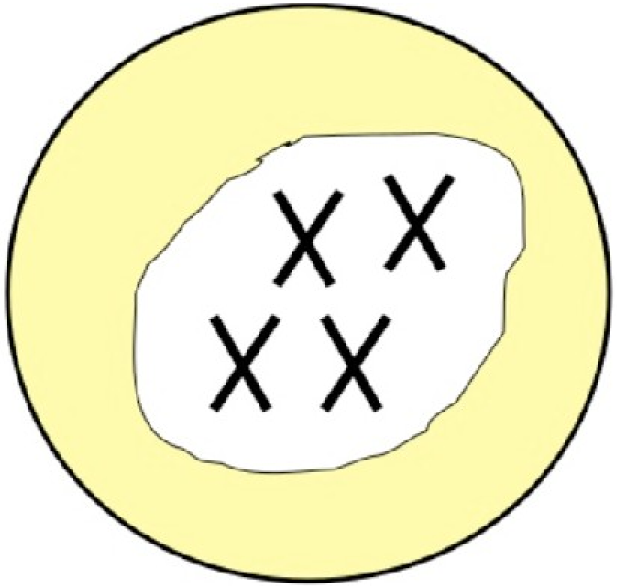


Dogma: *a principle, or set of principles, laid down by an authority as incontrovertibly true.*

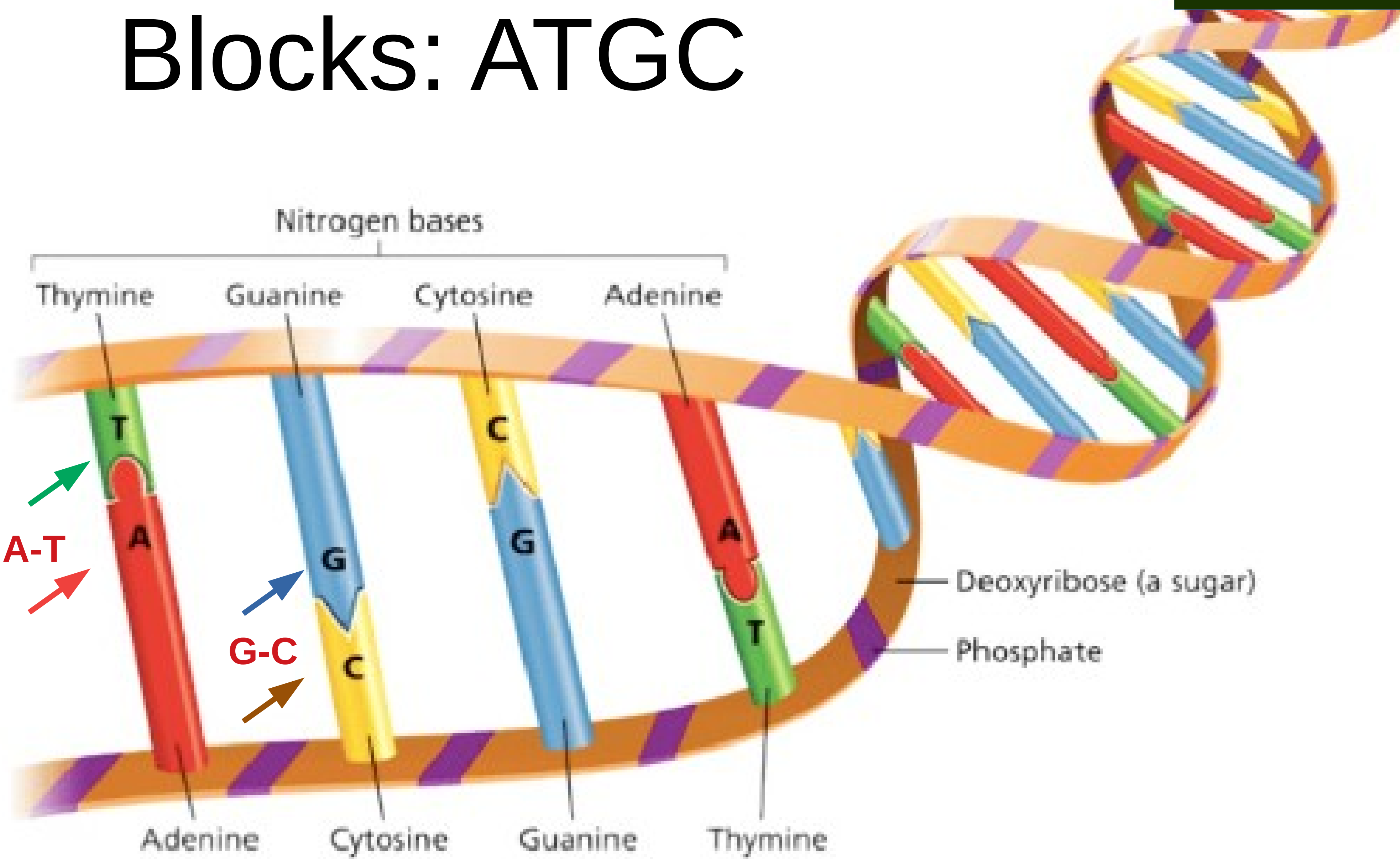
What is DNA?

- Found in the nucleus of a cell in two different structures: chromatin and chromosomes
- Genetic Material (Life's *blue prints*)
- *Written* inherited characteristics (genes)



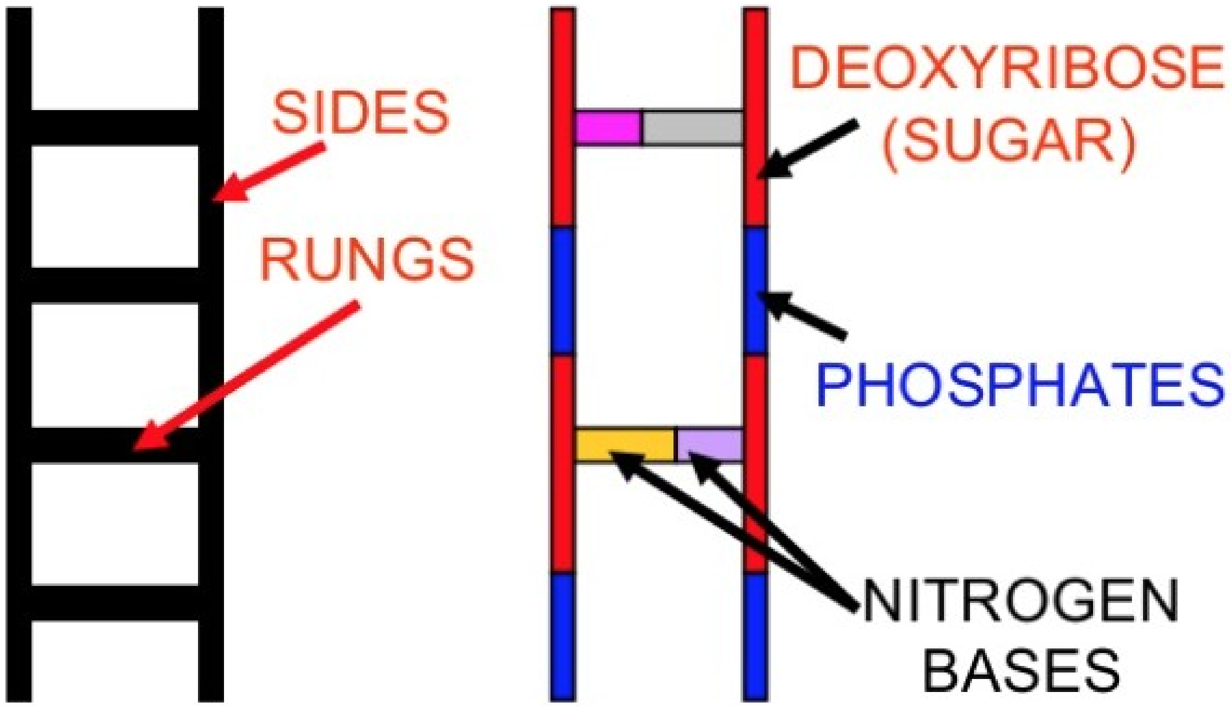


Molcular Building Blocks: ATGC



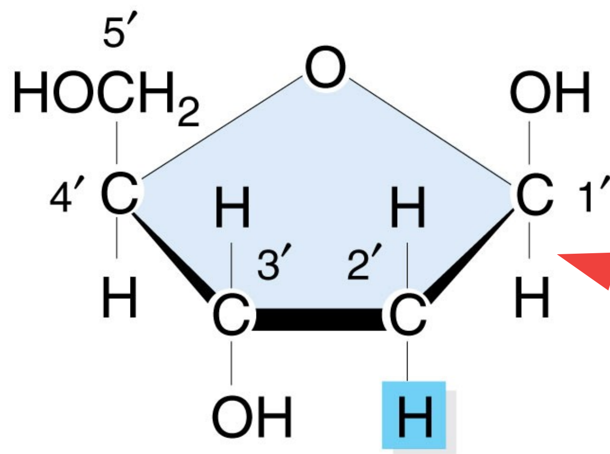
DNA Structure

- Formed like a twisted ladder
- There are two sides of the ladder
- Sugar (deoxyribose)
- Phosphates
- Alternating
- Rungs of the ladder
- Nitrogenous bases



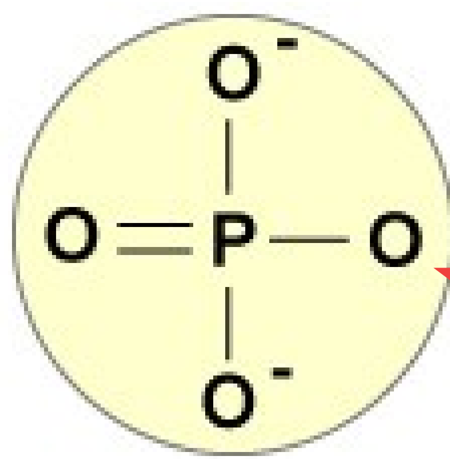
DNA Structure

Deoxyribose **S**ugar



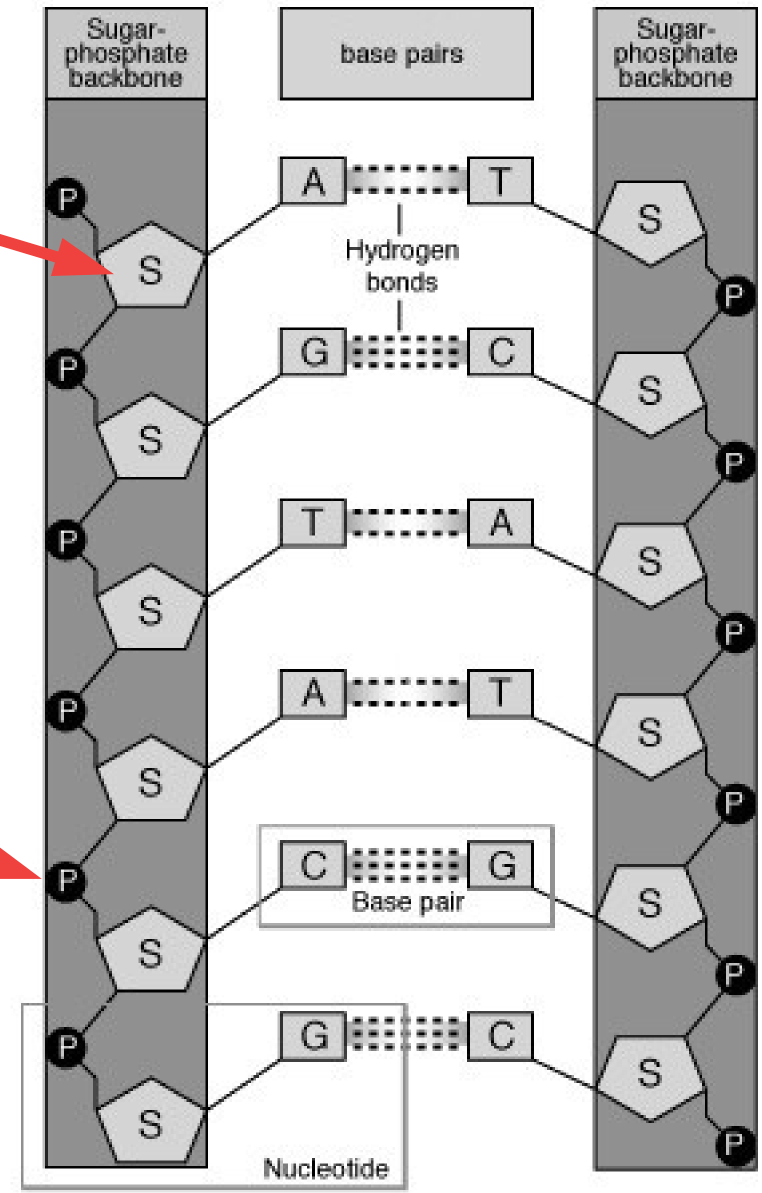
Deoxyribose

Phosphate



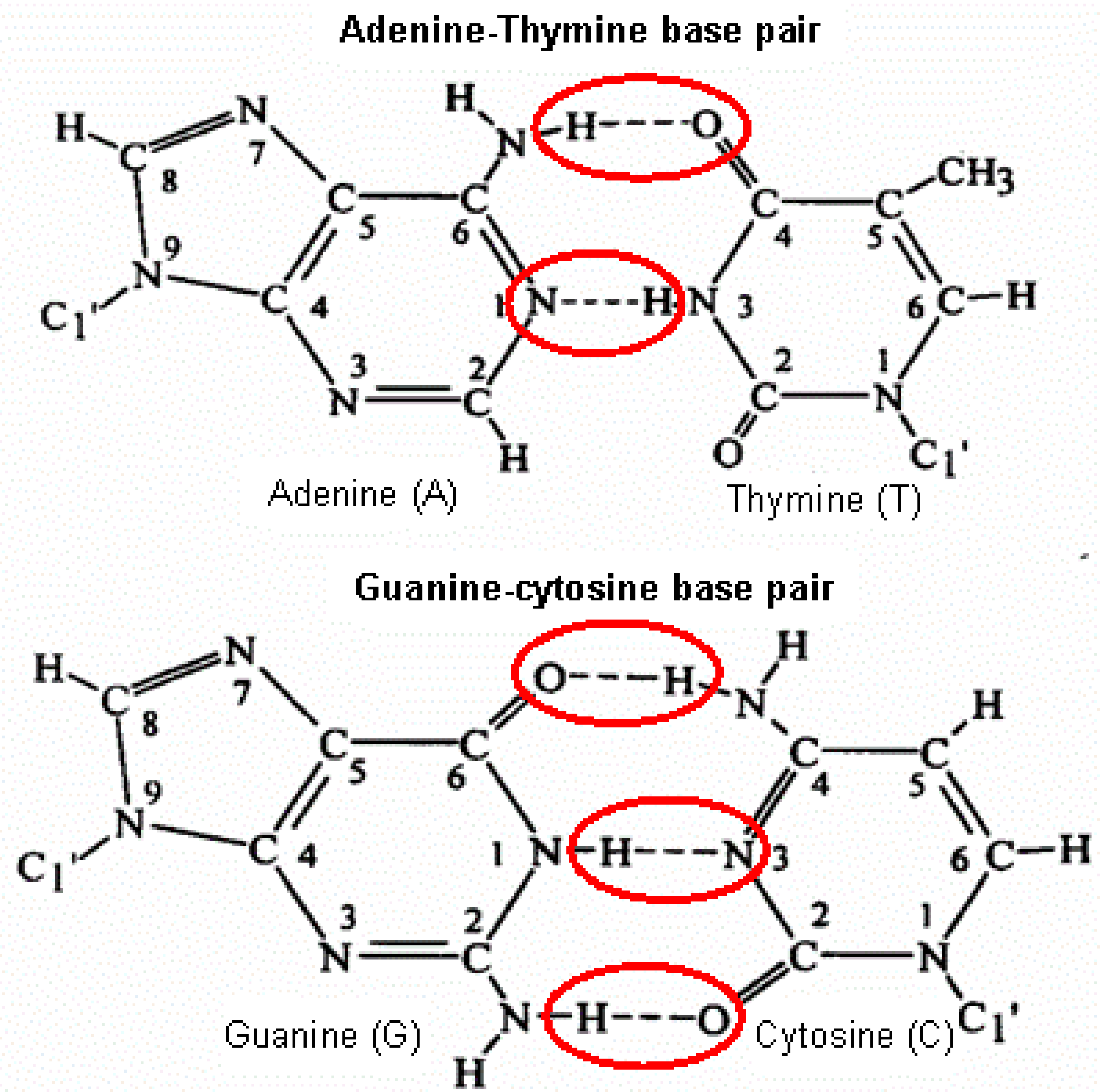
Phosphate group

Phosphorus surrounded by oxygens



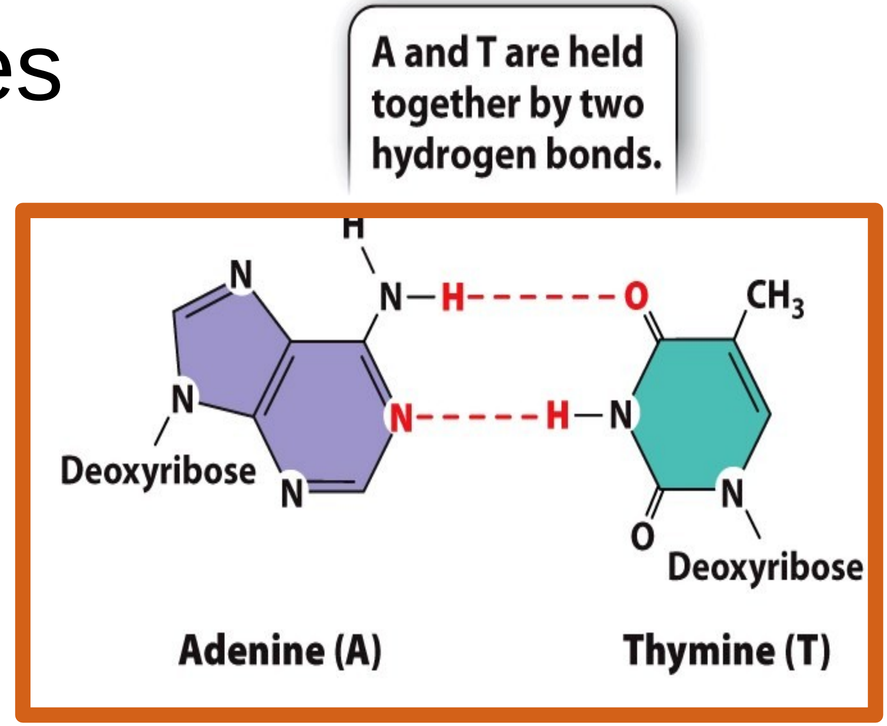
Base to Base Bonds:
How do nitrogenous bases pair?

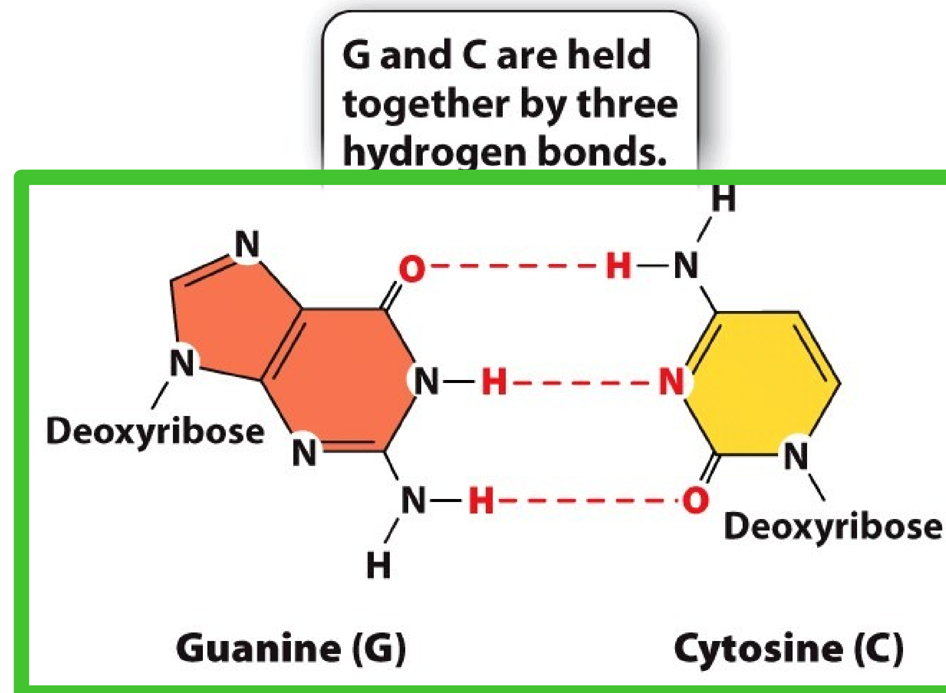
- Base-Specific bonding
- Preserves distance between (DNA's) backbones
- Hydrogen bonds
- Key to replication
- A-T's have two bonds
- G-C's have three bonds



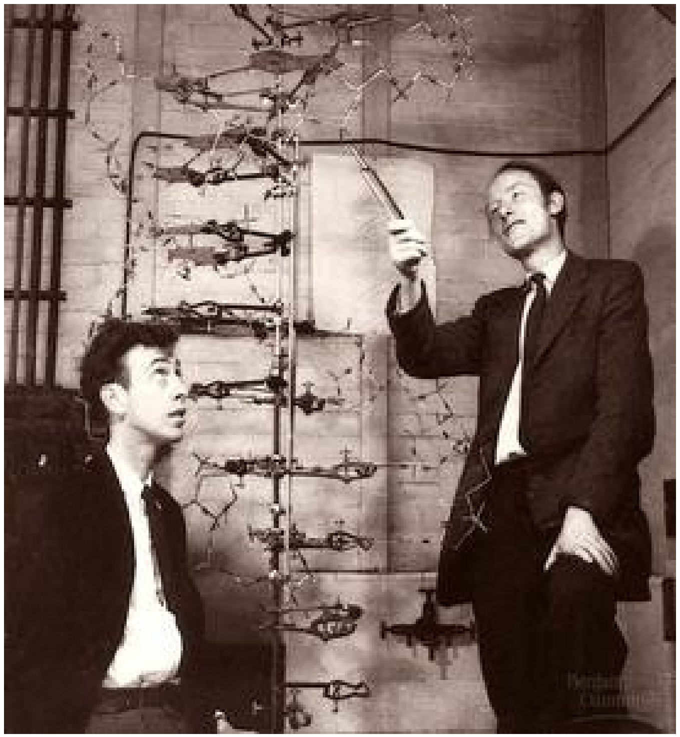
Purines and Pyrimidines

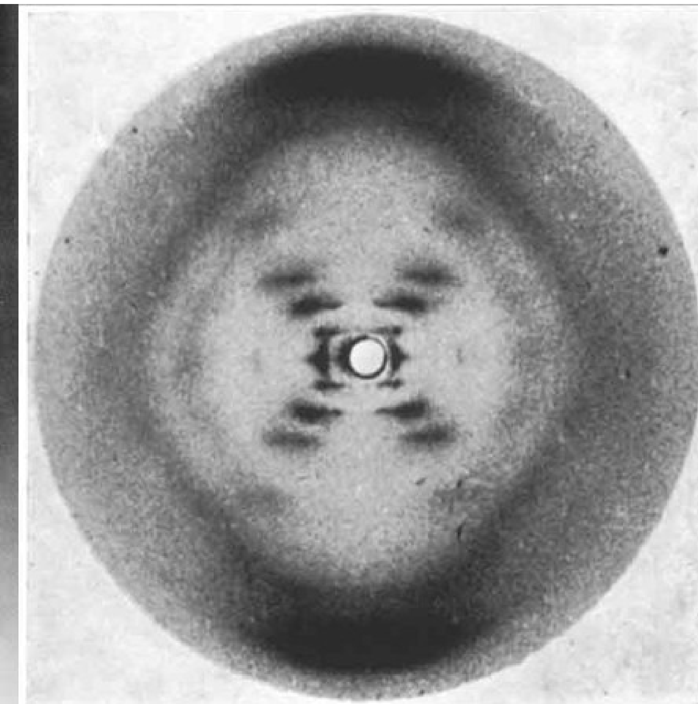
- Purines and Pyrimidines are nitrogenous bases that comprising the two different types of nucleotide bases in DNA and RNA.
- The two-carbon nitrogen ring bases (adenine and guanine) are purines, while the one-carbon nitrogen ring bases (thymine and cytosine) are pyrimidines.
- **Purines**: adenine and guanine
- **Pyrimidine**: thymine, cytosine, and uracil
- Purines include a number of biologically important compounds, such as adenosine, caffeine, uric acid, and the two bases adenine and guanine, which are components of DNA and RNA.





DNA Double Helix: Discovery of Structure



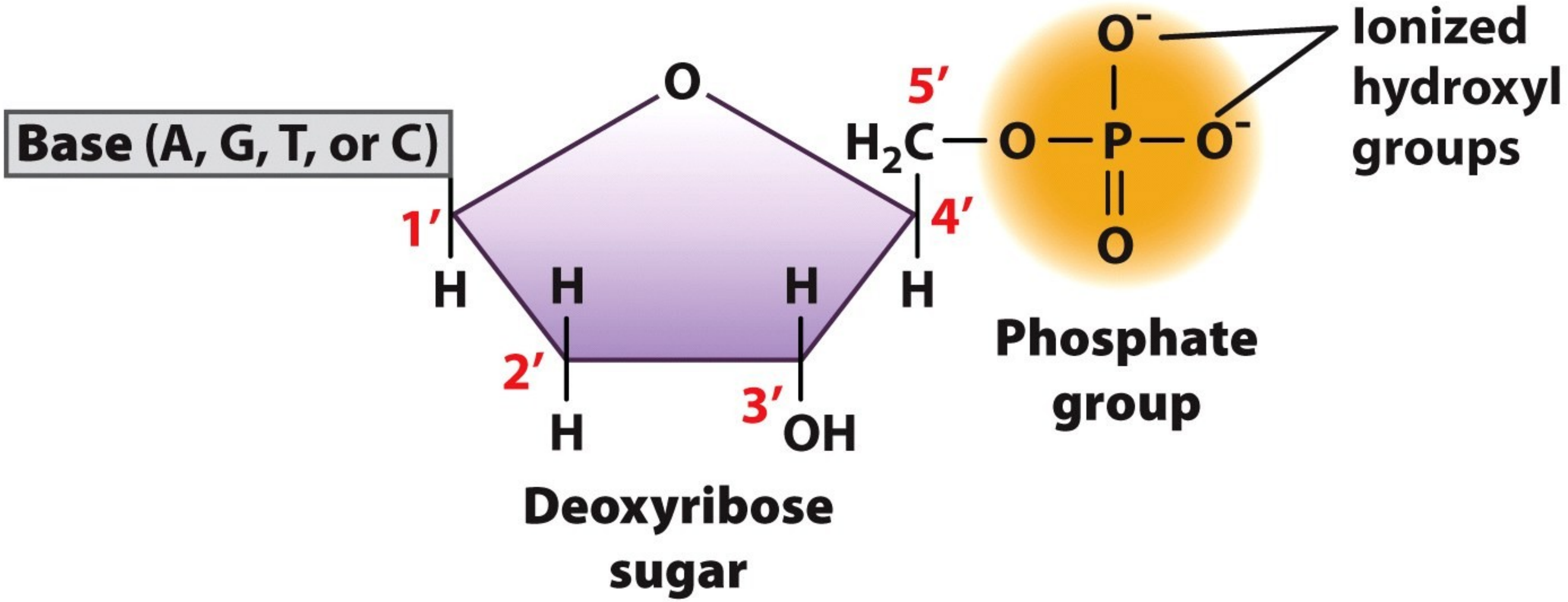


Watson and Crick, 1953

Rosalind Franklin and her data from x-ray crystallography

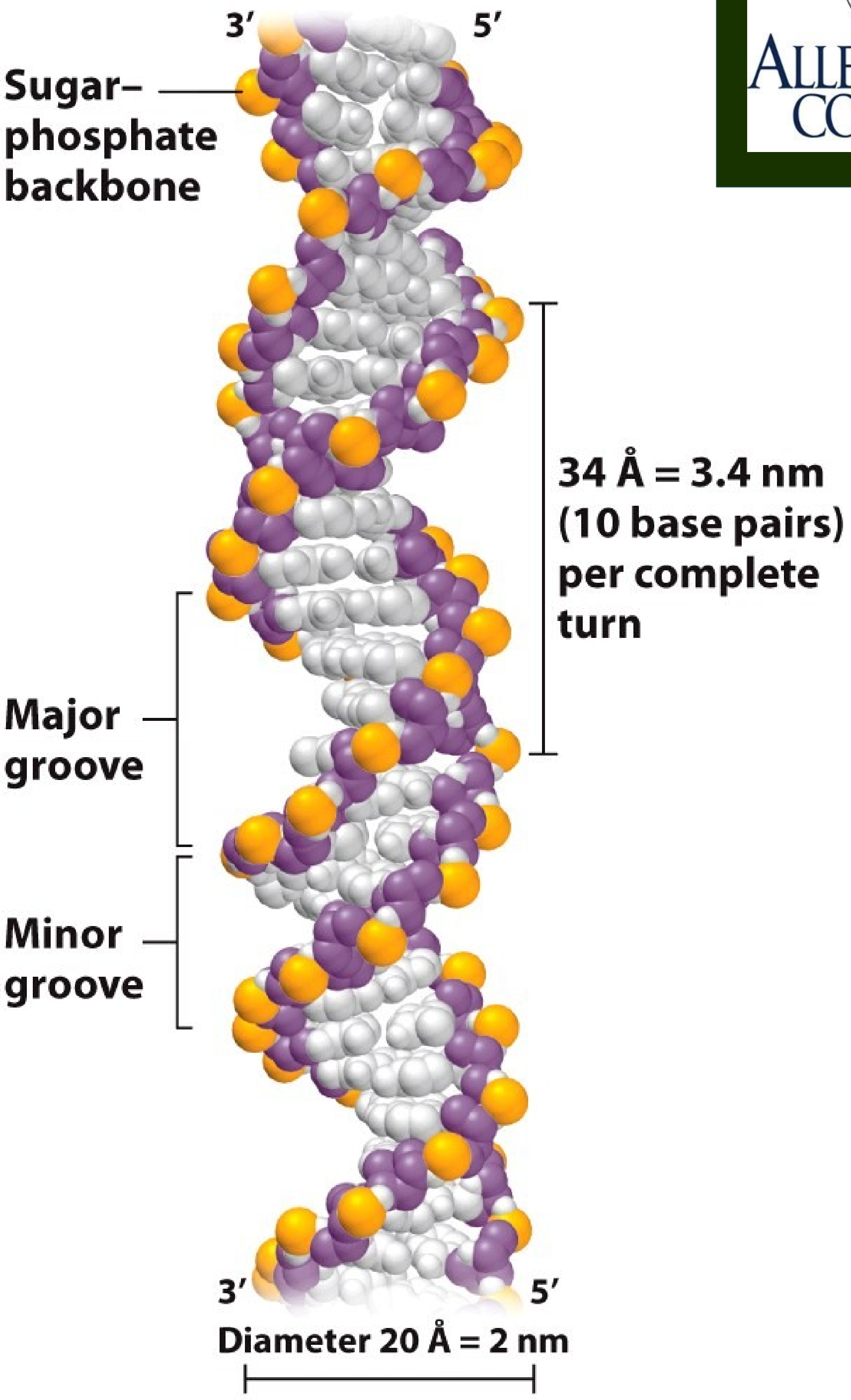
Nucleotide Structure

- *The support of rungs in the ladder.*
- One sugar
- One phosphate group
- One base (purine or pyrimidine)



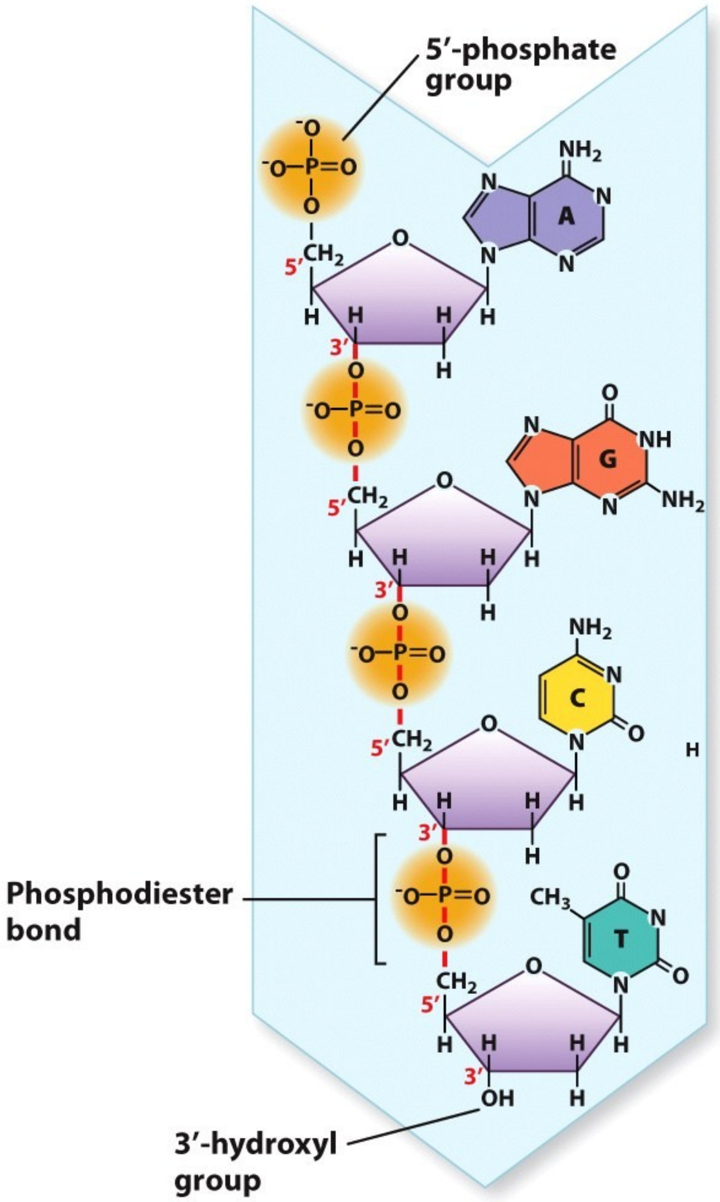
Structure of DNA

- Double-stranded
- Diameter – 2nm
- Helix
  - Complete turn = 10bp, 3.4nm
  - Major groove
  - Minor groove



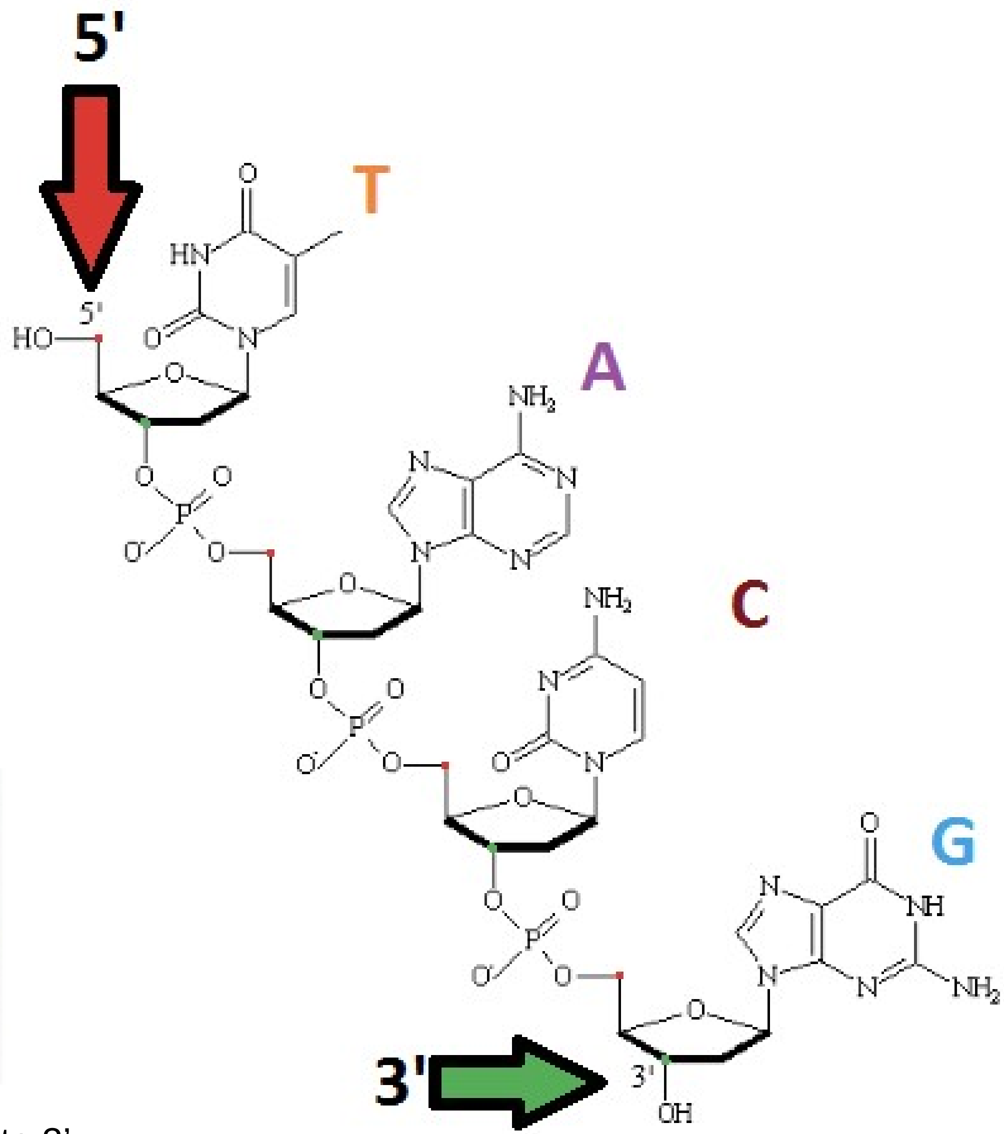
Structure of DNA

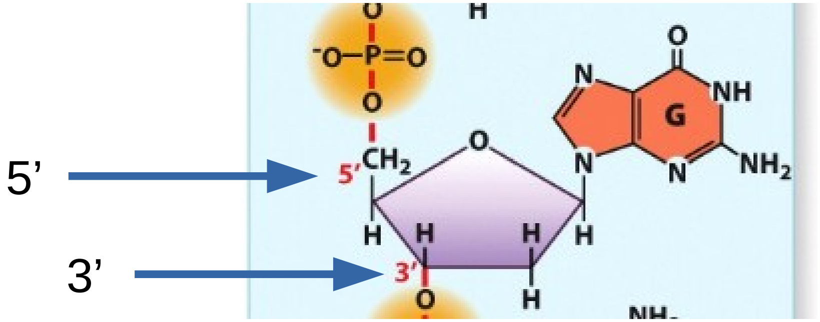
- Nucleotides are joined by phosphodiester bonds
  - phosphate to sugar
  - covalent bonds
- Polarity
  - 5’ end – phosphate group
  - 3’ end – hydroxyl group
- AntiParallel: DNA read in 5’ to 3’ direction



Five and Three Prime Ends

- A key feature of all nucleic acids is that they have two distinctive ends: **The 5' (5-prime) and 3' (3-prime) ends.**
- This terminology refers to the 5’ and 3’ carbons on the sugar.
- For both DNA and RNA, the 5' end bears a phosphate, and the 3' end a hydroxyl group.

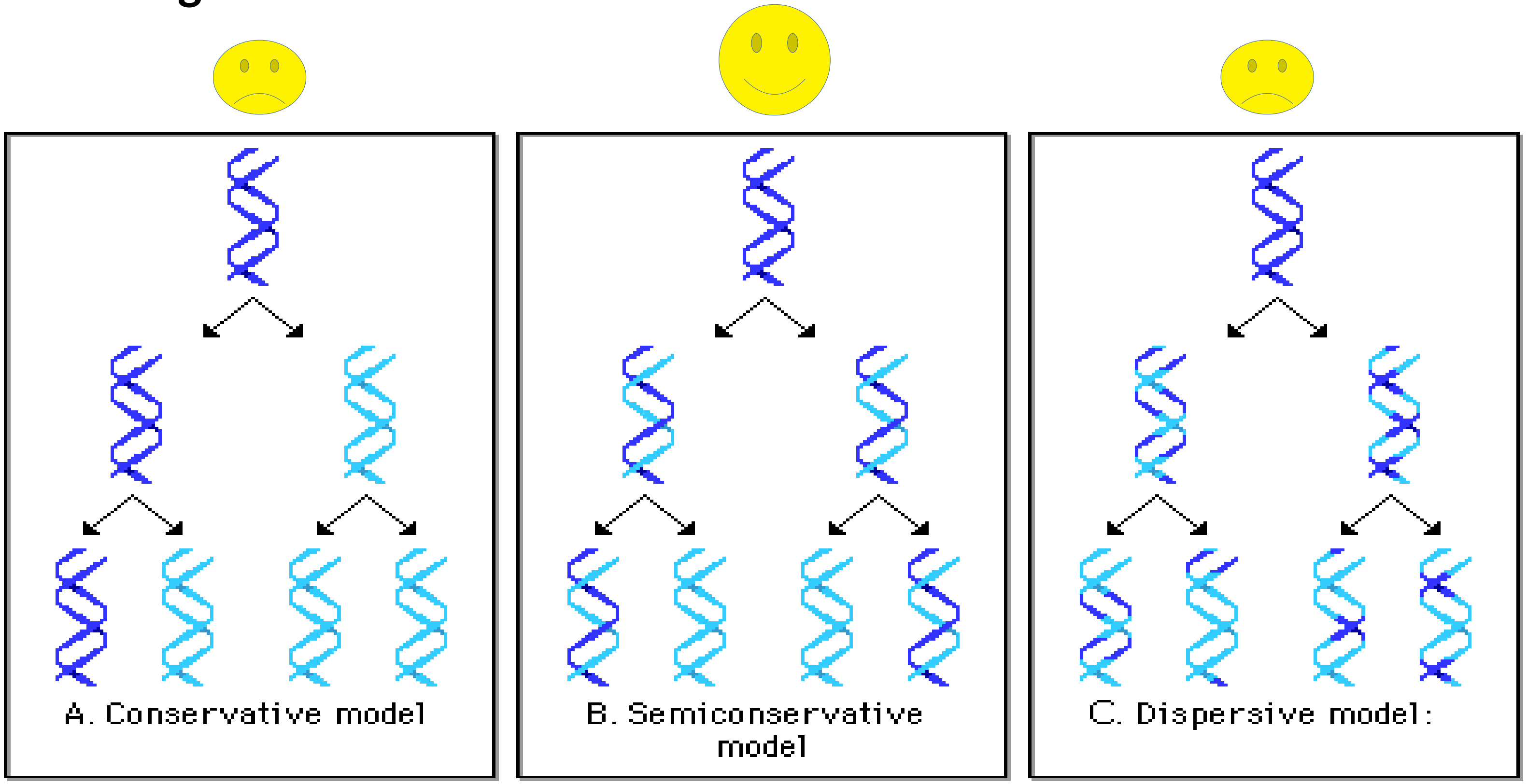




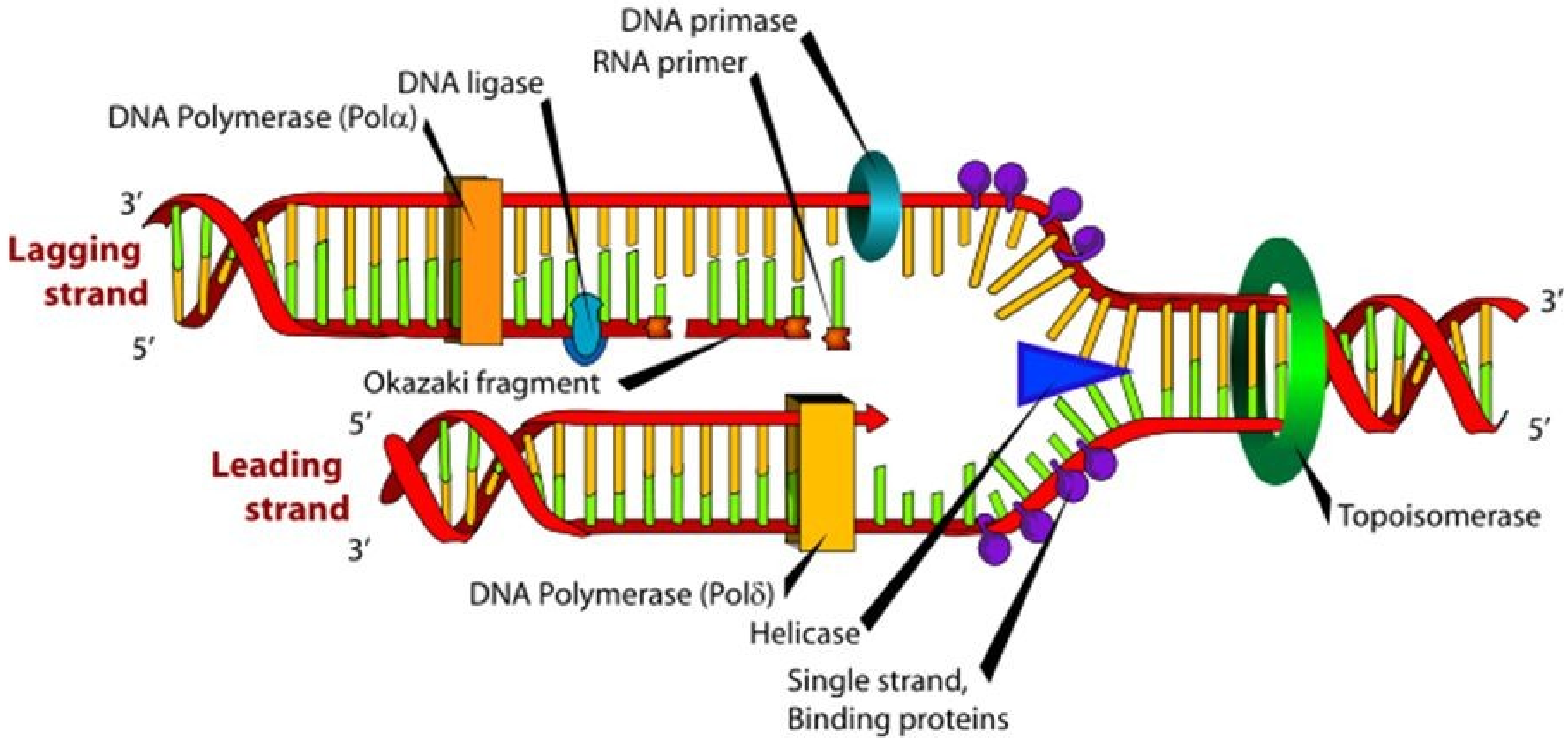
Note the ordering of carbons, hence 5’ to 3’

Proposed Mechanisms of Replication Process

1. The two sides of the parent molecule unwind/unzip
2. Daughter strands are synthesized using parent strands as templates
3. Parent/daughter duplex winds back together
   - **Semi-conservative: a 2nd gen helix composed of parental strand and one daughter.**

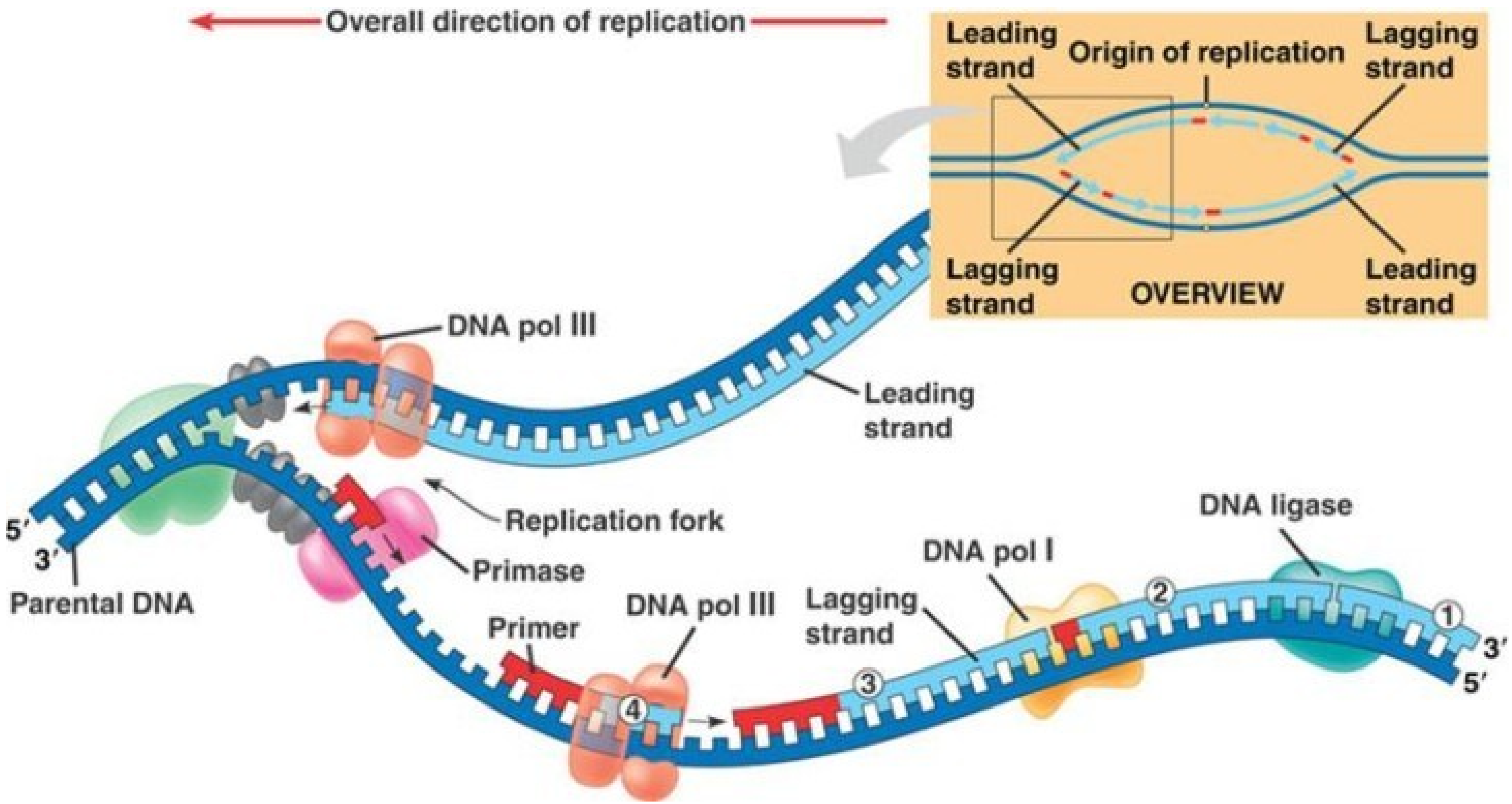


DNA Replication Process



http://www.hhmi.org/biointeractive/dna-replication-schematic

DNA Replication Enzymology



http://www.hhmi.org/biointeractive/dna-replication-basic-detail

http://highered.mheducation.com/sites/0073525324/student_view0/chapter20/dna_replication_fork.html

Review Question 1

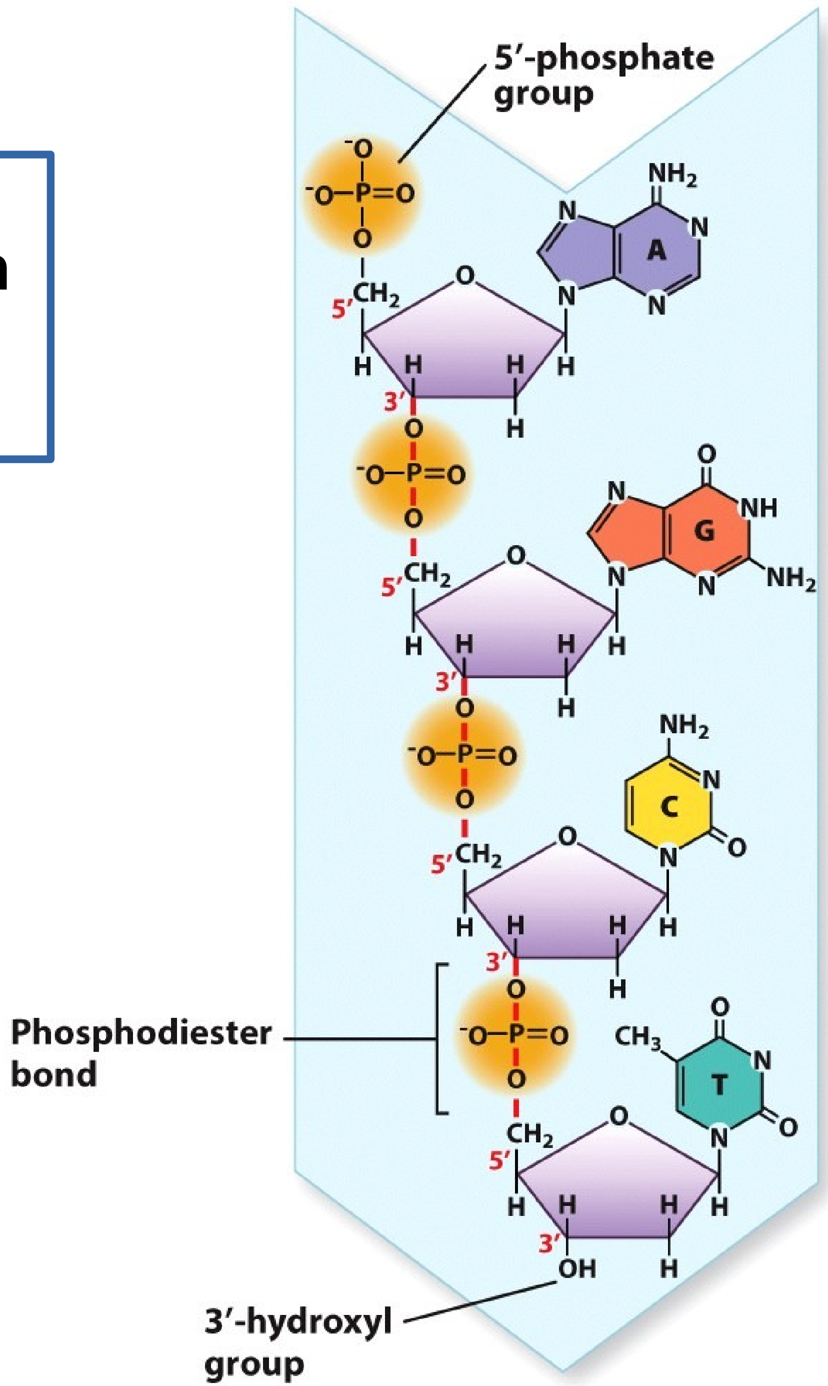
In the DNA sequence 5'–AGCT–3', the phosphodiester linkage between the adenine and the guanine connects:

A. The 2' end of the adenine to the 4' end of the guanine.

B. The 5' end of the adenine to the 3' end of the guanine.

C. The 5' end of the guanine to the 1' end of the adenine.

D. The 3' end of the adenine to the 5' end of the guanine.



Review Question 2

In the DNA of certain bacterial cells, 16% of the nucleotides are adenine. What are the percentages of the other nucleotides in the bacterial DNA?

A. 34% thymidine, 34% guanine, 16% cytosine

B. 34% uracil, 16% guanine, 16% cytosine

C. 16% thymidine, 34% guanine, 34% cytosine

D. 34% thymidine, 16% guanine, 34% cytosine

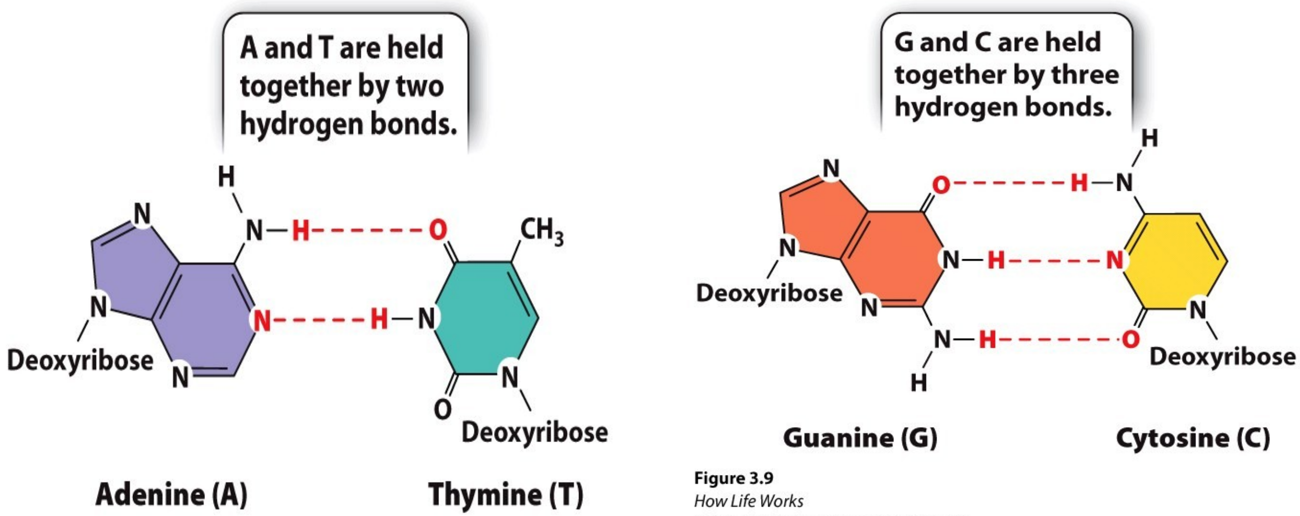
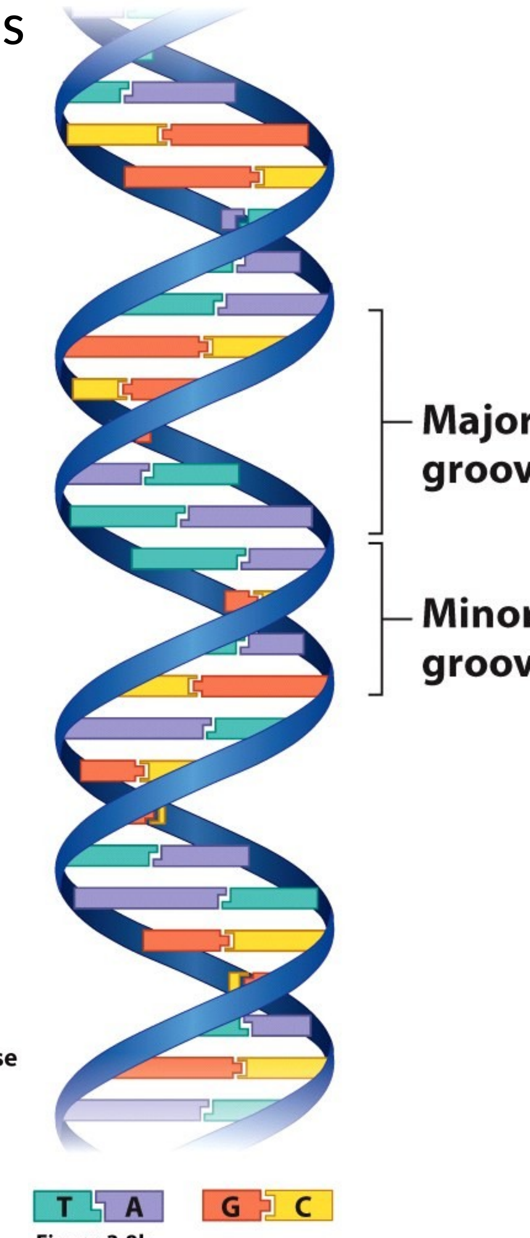


Figure 3.9
How Life Works



Review Question 3

1. DNA replicates in a semi-conservative manner. This means

   a) one daughter strand is synthesized as a large fragment while the other is synthesized in smaller fragments, both in the 5’-3’ direction
   b) every newly formed double-stranded DNA molecule consists of one parental strand and one daughter strand
   c) every newly formed double-stranded DNA molecule is comprised of two new daughter strands
   d) one daughter strand is synthesized as a large fragment in the 5’-3’ direction while the other is synthesized in smaller fragments in the 3’-5’ direction

Supporting Videos

- The Double Helix (Documentary about DNA discovery, 17 mins)
- http://media.hhmi.org/biointeractive/films/Double_Helix.html

- The Chemical Structure of DNA (3 mins)
- http://www.hhmi.org/biointeractive/chemical-structure-dna

- The Structure of DNA (6 mins)
- https://www.youtube.com/watch?v=o_-6JXLYS-k

- The def of 5' and 3' strands (1.5 mins)
- https://www.youtube.com/watch?v=qWZYpHSXvJo