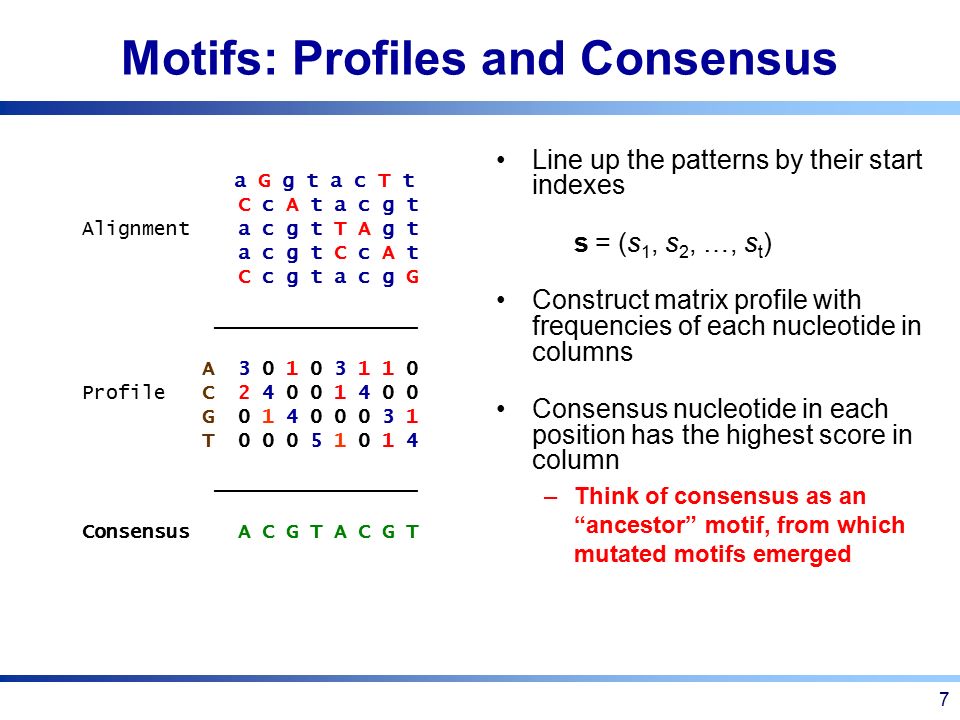
1. **DNA** – a self-replicating material which is present in nearly all living organisms as the main constituent of chromosomes. It is the carrier of genetic information.
2. **RNA** – a nucleic acid present in all living cells. Its principal role is to act as a messenger carrying instructions from DNA for controlling the synthesis of proteins, although in some viruses RNA rather than DNA carries the genetic information.
3. **Nucleotides** – a compound consisting of a nucleoside linked to a phosphate group. Nucleotides form the basic structural unit of nucleic acids such as DNA.
4. **Amino Acids** – a simple organic compound containing both a carboxyl (—COOH) and an amino (—NH2) group.
5. **Transcription** – begins when an enzyme called RNA polymerase (RNA pol) attaches to the DNA strand and produces complementary RNA.
6. **Translation** – occurs to convert RNA into proteins. Because the Ribosome examines DNA 3 bases at a time, we will often look for codons in DNA as we are examining it.
7. **Ribosomes** – a minute particle consisting of RNA and associated proteins found in large numbers in the cytoplasm of living cells. They bind messenger RNA and transfer RNA to synthesize polypeptides and proteins.
8. Given the following conversion chart between DNA and the 20 amino acids, convert a string of DNA to amino acids.
   * Use chart
9. Given an alignment of DNA showing a motif, construct the profile for the motif.



1. What is the Hamming distance between the following strings?
   * Use code
   * Number of differences in strings.
2. Which molecular machine is responsible for the replication of DNA?
   * DNA polymerases.
3. What is the Origin of replication?
   * A particular sequence in a genome at which replication is initiated.
4. What is the complexity of the FrequentWords algorithm?
   * O(|Text|^2 \* k)
5. What is the reverse complement of the following DNA string?
   * Use code
   * A -> T and G -> C, then reverse string.
6. Why is replication asymmetric?
   * DNA polymerases can only add nucleic acids to the 3'-OH end of DNA.
7. What happens during the deamination process?
   * The removal of an amino group from a molecule.
8. How could horizontal gene transfer impact the skew diagram when looking for an origin of replication?
   * If a gene from the forward half-strand of one bacterium is transferred to the reverse half-strand of another, then we will observe an irregularity in the skew diagram.
9. Which of the following are Neighbors with distance 3 of a string of DNA?
   * Use code
10. Gene A promotes gene B if:
    * When a transcription factor binds to a site on the DNA a little bit upstream of that DNA sequence. Some genes can control/promote/inhibit the transcription of other genes based on the proteins that they produce.
    * Gene B starts transcription when transcription factor X binds to B's binding site. And Gene A creates X when it's expressed. Since the expression of Gene A creates the transcription factor X, which leads to the Gene B being expressed, Gene A promotes Gene B.
11. What is the complexity of BRUTE FORCE MOTIF SEARCH?
    * O(n^t \* k \* t)
12. What is the complexity of the MEDIAN STRING problem?
    * O(4^k \* n \* t \* k)
13. GREEDY MOTIF SEARCH can find a motif much more quickly than BRUTE FORCE MOTIF SEARCH.  Why wouldn't you always use the GREEDY MOTIF SEARCH algorithm?
    * Because most greedy algorithms typically fail to find an exact solution of the problem and find an approximate solution instead.
14. How does Laplace’s Rule of Succession modify the GREEDY MOTIF SEARCH matrices?
    * Laplace’s Rule of Succession adds 1 to each element of COUNT(Motifs), updating the matrices.
15. What kind of an algorithm is RANDOMIZED MOTIF SEARCH?
    * Monte Carlo Algorithm – find quick approximate solutions that are not guaranteed to return the exact solution.
16. What problems can occur with using the GIBBS SAMPLER for motif search?
    * When there are islands of high-probability states, with no paths between them.
    * When all states have nonzero probability and there is only a single island of high-probability states.
    * Gibbs sampling replaces a single k-mer at each iteration.
    * Random patterns compete with the real motifs because there are so many errors.
17. Which DNA string is in the set of strings represented by this motif?
    * Use chart.