**Bioinformatics Worksheet #1: Bioinformatic Vocabulary – KEY**

The field of bioinformatics, just like other sciences, has vocabulary that is specific to the discipline. Understanding that terminology is key to successfully working through the process of identifying the proteins contained within a sequence of DNA, encoded by the sequence of nitrogenous bases, which, when translated, becomes a series of amino acids (aa) that make up each protein’s primary structure. Work to define the vocabulary words below, both from your previous material on genetics and the lecture provided today on the field of bioinformatics.

**Operon –**A functional genomic unit containing multiple genes transcribed by the same promoter.

**BLAST – Stands for B**asic **L**ocal **A**lignment **S**earch **T**ool. It is an algorithm used for comparing sequence information, either nucleotide or amino acid, with other sequences, typically from a database. Allows researchers to search genes or proteins by sequence to find potential matches or similar sequences.

**Prokaryote –** Unicellular organisms characterized by the lack of organelles.

**Codon –** A triplet of nucleotides that tRNA attaches to during translation. Corresponds to a specific amino acid

**Transcription –** The process of creating a complementary RNA sequence from a section of DNA.

**Translation –** The process of creating a peptide from an RNA transcript.

**Reading Frames (and why are there 6 for any sequence) –** a sequence of DNA split up into non overlapping triplets. There are 3 starting from the 1st, 2nd, and 3rd nucleotides from one end and three more from the other. Each produces a different sequence of amino acids

**Open Reading Frame –** The section of a reading frame that contains no stop codons

**Promoter-** Binding site for RNA polymerase to start transcription

**Shine Delgarno Consensus Sequence – B**inding site for the 16S rRNA, allowing for the active site to be situated over the start codon. Located about 8 nucleotides upstream of a start codon and is typically AGGAGG, though there is variation

**Central Dogma of Genetics–** Describes the flow of genetic information from DNA->RNA via transcription and RNA -> Protein via translation.