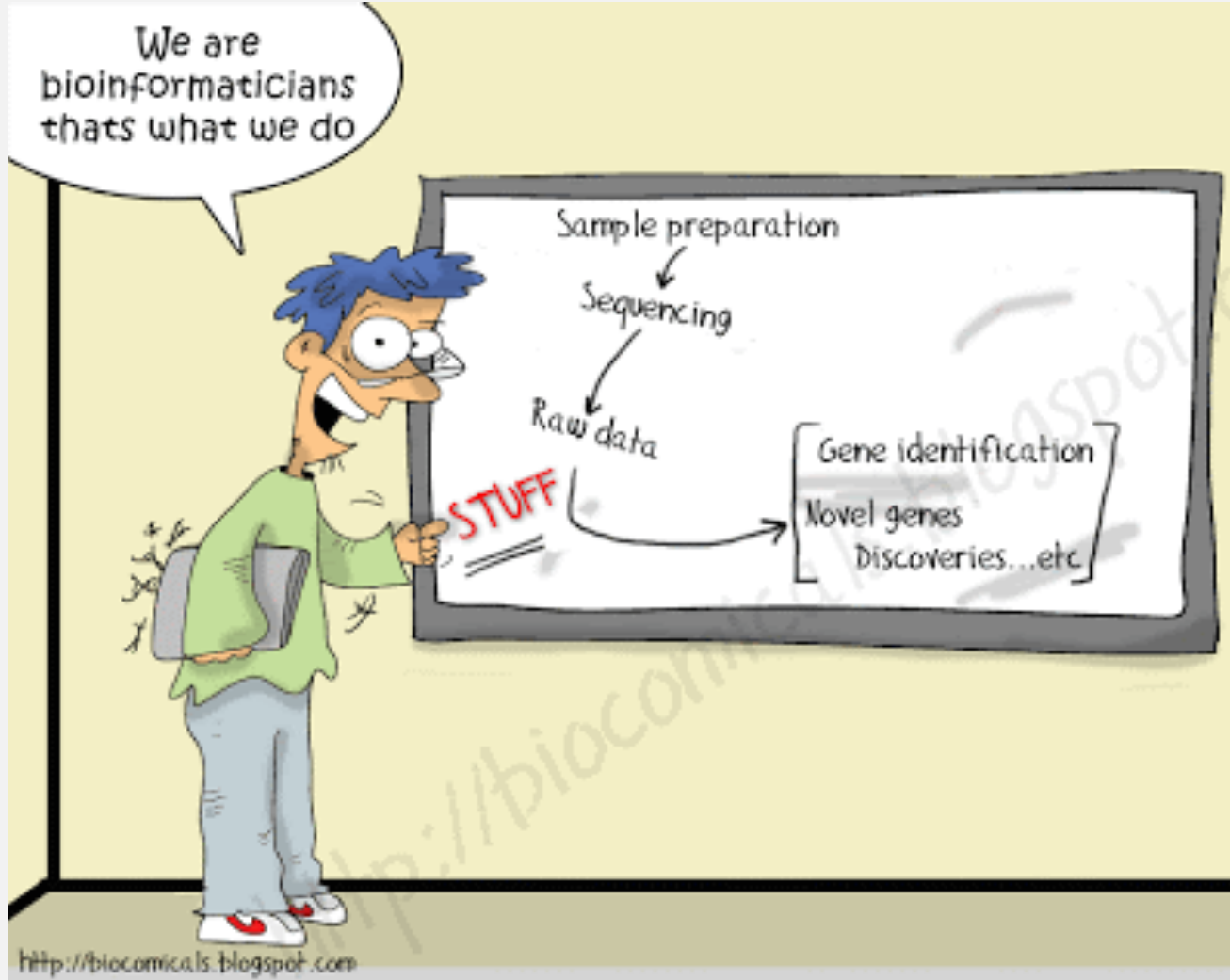


# BIOINFORMATICS INTERNSHIP

Session nine



## FUN FACT

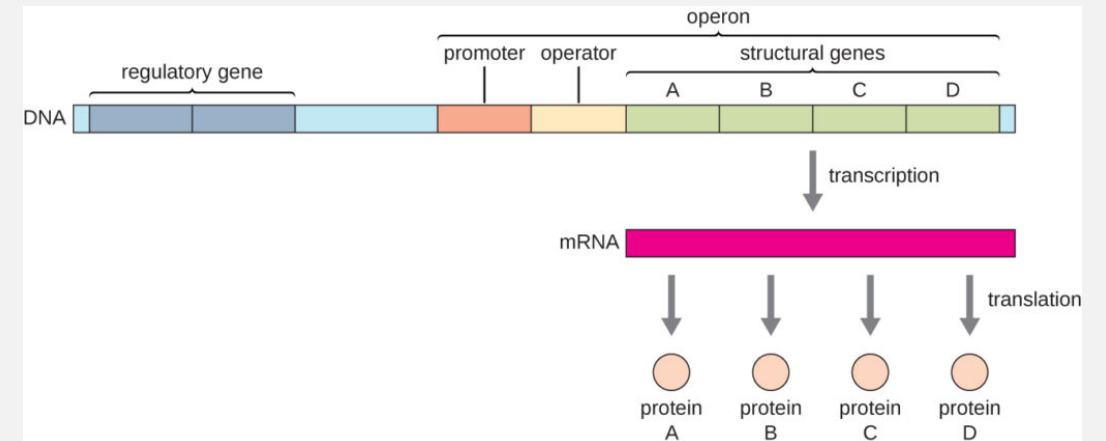
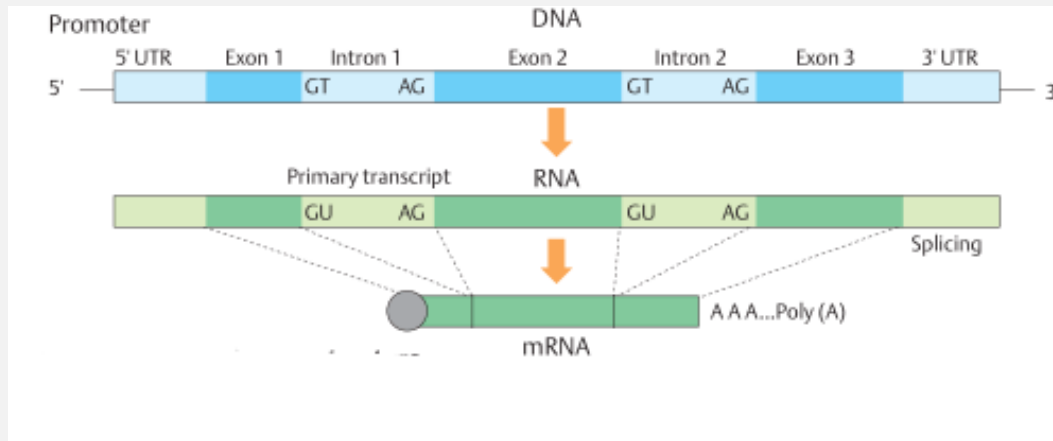
- What is “STUFF”?

## LETS REVIEW INFO ABOUT OUR DATA

- Long reads from illumina (NGS)
- Short reads from nanopore (NGS)
- Assemblies from Cushaw (short read mapping)

# NUCLEOTIDES

- Structural units that pair with one another to form DNA
- Made of sugar, a phosphate group, and a nitrogenous base
  - Each are different based upon orientation of compound bonds
- sequenced DNA tells us the order of the DNA, which makes up a single gene or a whole genome



# GENES

- A gene is a small section of DNA within a genome, that stores information
- The genes that we are viewing are purposed to make proteins, Spidroins
  - Are located in the silk production glands (can we name all ??)
- Can be up to 3 million bp

# AMINO ACIDS

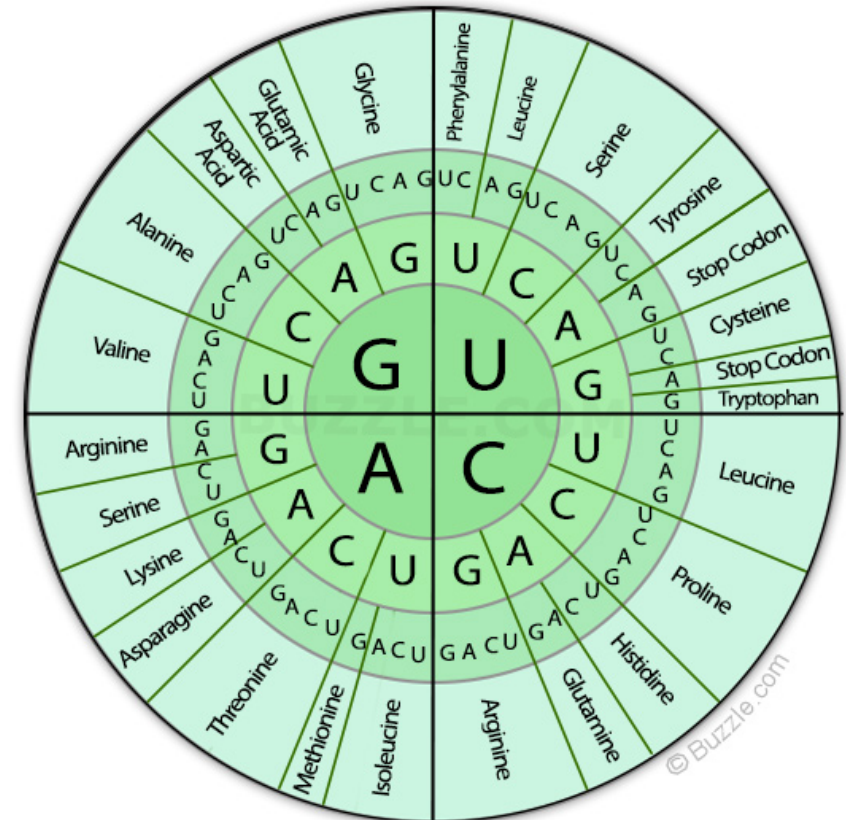
Organic compounds that contain amine and carboxyl functional groups.

A side chain differentiates.

There are 20 aa.

Encoded by RNA base sequence (a combination of GACU).

Multiple amino acids bind together and form a protein.



To decode the codon, move from the center circle towards the periphery.

|   |               |   |            |
|---|---------------|---|------------|
| A | Alanine       | M | Methionine |
| C | Cysteine      | N | Asparagine |
| D | Aspartic Acid | P | Proline    |
| E | Glutamic Acid | Q | Glutamine  |
| F | Phenylalanine | R | Arginine   |
| G | Glycine       | S | Serine     |
| H | Histidine     | T | Threonine  |
| I | Isoleucine    | V | Valine     |
| K | Lysine        | W | Tryptophan |
| L | Leucine       | Y | Tyrosine   |

---

Data from International Union of Pure and Applied Chemistry and International Union of Biochemistry.<sup>(70)</sup>

## AMINO ACID

# PROTEINS

- A molecule that carries out a particular duty
  - Composed and determined by amino acids
- Spidroin
  - Main component in spider silk
  - MaSp are large proteins (~3500 aa composition)
  - ~90% of protein sequence is repeat



# FORKING

- [See on github](#)