Bio4250 – Evolutionary Genetics **Dr Carr** **02 Sept 2022**

Lab #1 - (Re)familiarization with the structure of Deoxyribonucleic Acid (DNA)

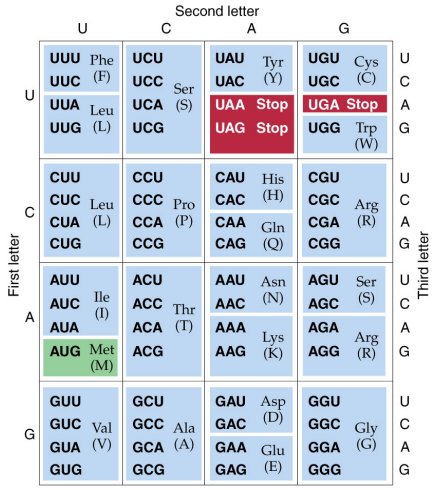
Modern evolutionary genetics centers on inferences made from the **DNA** molecule. Famously, **DNA** is a **double-stranded helix**, whose **sugar-phosphate (deoxyribose-PO_4) backbone** strands run in **anti-parallel directions**, and are held together by hydrogen bonds (**H-bonds**) between four bases, **Adenine**, **Cytosine**, **Guanine**, and **Thymine** (**A**, **C**, **G**, & **T**). This much of the structure is easily appreciated in flat drawings, with the work already done. It is less obvious in three-dimensional models, where structure must be appreciated from molecular properties.

In this lab, you will construct three-dimensional models of the **DNA** molecule from solid **3D** printer models of the **sugar-phosphate backbone** and the **A**, **C**, **G**, & **T** subunits. The purpose is to reinforce your understanding of **DNA** structure as it pertains to molecular evolution of the molecule. The lab will also introduce **Single-Nucleotide Polymorphism** (**SNP**) variation as the basis of molecular evolution.

1. Familiarize yourself with the **building blocks**.
    a. Identify the **5'** and **3'** ends of the **deoxyribose-PO_4** backbones. What features identify these?
    b. Identify each of the colored molecules with the correct bases. What features allow you to recognize them? [***Note***: the base colors conform to the standard convention]
    c. Identify the appropriate **base pairs**. What features allow them to pair, and how does this relate to question 1.b ?
2. Construct a **3D** model of an **ATG Start** triplet.
    a. Attach an **A** base to the **deoxyribose-PO_4** molecule in the **correct 5'-3' orientation**. How do you recognize the correct orientation? What is this molecule called?
    b. Add the **T** and **G** bases at the correct end of the **A**. How do you know which is the correct end?
    c. **Check your construction with the Instructor**.
3. Add the appropriate bases for the **complementary strand**, with the *proper strand orientation*, to form a double-stranded 3bp molecule. [***Note***: base pairs should join & separate easily; **deoxyribose-PO_4** molecules will require a bit of twisting to the proper orientation to join each other. Don't force a fit].
4. Construct a **multi-triplet DNA molecule** as instructed [***Note***: this may or may not start with the **ATG**].
    a. Why are the **3bp runs** called '***triplets***' rather than '***codons***' ?
    b. Recall that separation between adjacent base pairs is **0.34 nm**, and that one complete helical turn of **3.4 nm** comprises **10 bp**. As best as you can, adjust the backbone to achieve this.
        i. Recall that **DNA** is a **right-handed** helix: how do you arrange & recognize this ?
    c. Add the appropriate bases as in Step 3 to form a **dsDNA** molecule as instructed.
        i. Try constructing each strand *separately* (pay attention to **5'-3'** orientation) and / or
        ii. Try adding *successive* base pairs to the appropriate end.
    d. ***If** you come up with a clever means of achieving a 10bp period, share it with the class.
5. **Join** the separate molecules from each working group into a **single molecule**.
    a. Identify the **coding** and **non-coding** strands, and the **5'-3'** orientation. What do you look for? [Have all groups done it right?]
    b. Identify the **amino acid polypeptide** encoded by the molecule, beginning with the start triplet. Use the **universal genetic code** [next page].

6. **SNP variation:**
    a. The gene coding for the **β-globin** chain of Hemoglobin (**Hb**, standard from **Hb-A**) is known to exist in a variety of allelic forms, some of which result in severe anemias. The most famous of these is **Hb-S**, which results in **Sickle-Cell Anemia** in West African populations and their descendants in North America. The allele is due to a **2nd position transversion** in the **6th triplet**.
        i. **Make a model** of the first six triplets of **Hb-A**.
        ii. **Alter the model** to demonstrate the **SNP** that results in **Hb-S**.
        iii. **NOTE TO SELF**: Given only the amino acids, the base sequence of the **DNA** molecule may be ambiguous, with two or four **SNPs** coding the same amino acid substitution. **Probably**, *give students the base sequence of the first six triplets*, AND have them discover the ambiguity of the **SNP** substitution themselves.
    b. The table below left gives the amino acid sequences of the **β-globin** molecule for seven allelic variants, including **Hb-A** and **HB-S,** at different positions in the molecule. For one of the last five allelic variants as assigned, construct an artificial **DNA** sequence based on positions **6**, **26**, **63**, & **121**.
        i. **NOTE TO SELF**: The Instructor can construct the **Standard** ["*Normal*"] sequence as a demonstration:

β-chain	Amino acid position						
	1	2	6	26	63	121	146
Normal	Val	His	Glu	Glu	His	Glu	His
Hb-S	Val	His	Val	Glu	His	Glu	His
Hb-C	Val	His	Lys	Glu	His	Glu	His
Hb-E	Val	His	Glu	Lys	His	Glu	His
Hb-M Saskatoon	Val	His	Glu	Glu	Tyr	Glu	His
Hb Zurich	Val	His	Glu	Glu	Arg	Glu	His
Hb-D β Punjab	Val	His	Glu	Glu	His	Gln	His



Questions

a. Amino acid residues ##6, 26, & 121 involve **SNP** variants of **GLU** triplets to three alternative coding triplets. What is the **DNA SNP** in each case?
b. Same question for residue #63, which involves **SNP** variants to *two* alternatives of the **HIS** triplet.
c. Which of these **SNP** variants might result in functional changes to the **β-chain**? Explain.
d. Why are the 3-letter combinations called '***triplets***' rather than '***codons***' ?
e. Why are the nucleotide variants called '***SNPs***' rather than '***mutations***'?
f. **FOR FURTHER THOUGHT**:
    i. The table at right is the **Universal Genetic Code**: how does it differ from a **DNA translation table**?
    ii. Why might the **GLU** triplet be especially prone to **SNP** variation? [How many amino acid substitutions are possible given a single SNP in the triplet?]

NOTE TO SELF: Construction of two six triplet molecules may be excessive. Given the purpose of the lab, skip Steps 4 & 5 and proceed directly to Step 6 for its focus on SNPs, with only middle four triplets that show SNP variation.