

Day 3: Control flow Practicals

Let's get down to business and start typing some real code. In order to solve the proposed exercises you are required to use the provided examples.

```
In [1]: from data import cytb, translations, acidAA, IUPAC_codes
```

In this module you will find 3 things:

- examples.sequences
- examples.translations
- examples.acid-aa

Just print() any one to see what's inside each of them.

Problem 1

Verify that there are no illegal characters in any of the sequences in *cytb*.

Tip:

Use the *IUPAC_codes* list that was imported from the data module to check if the sequence's nucleotides are legal.

```
In [ ]:
```

Problem 2:

Translate the sequences in *cytb* from nucleotides into aminoacids.

Tips:

Use the translations dictionary from the data module (*translations*);

Use a *for* loop to iterate through the sequences in *cytb*;

Translating the sequences into a list of aminoacids rather than a string will save you work later on.

```
In [ ]:
```

Problem 3:

Reverse and complement the sequences in *cytb*.

Tip:

Use slicing to reverse the sequence.

```
In [ ]:
```

Problem 4:

a) Find the aminoacids in *acidAA* in the sequences from *cytb*. Return the positions of both the amino acid (in the protein sequence) and the codon position (in the nucleotide sequence).

Tip:

Remember that an acid aminoacid may be present more than once.

```
In [ ]:
```

b) How frequent are the acid amino acids? Are they more or less frequent than what you would expect under a random distribution pattern (i.e., where every amino acid has the same probability of occurring)?

```
In [ ]:
```

c) Return the sequence's nucleotides from each sequence in *cytb* until any of the acid amino acids from *acidAA* is found; Return "Target

aminoacids not found" if neither of them is present in the sequence.

Tip:

A while loop is the simplest choice here, but feel free to use any other way

In []:

Problem 5:

Find all the equal sequences in *cytb* and "collapse" them into a single sequence (with the name of all the collapsed sequences).

In []: