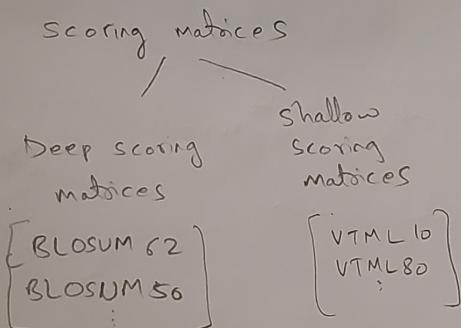


Bio python accepts both DNA & Protein sequences,
for DNA sequences → 'Identity' is the name of
the algorithm that generates a
distance scoring matrix

Dec 15/2021 | Wednesday | 14:38 | @ 106

Okay about "Identity" distance matrix generator

Identity: - it's a substitution model



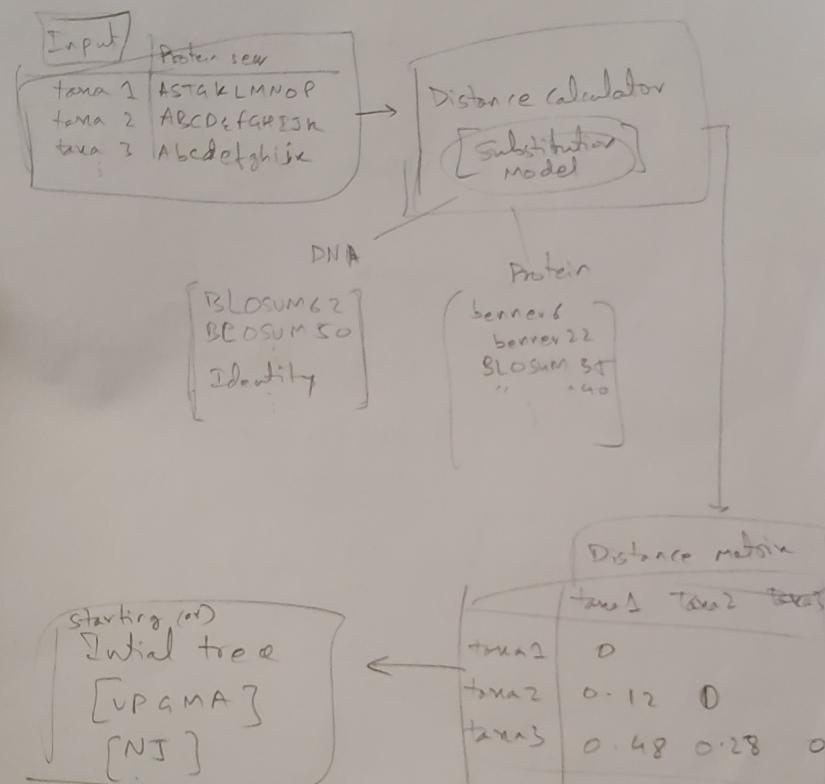
BLAST → Basic Local Alignment Search Tool
→ One of the most widely used bioinformatics program for
sequence matching

→ Algorithms that came before BLAST →
Smith-Waterman Algorithm

Guarantees best Sequence Alignment
but time consuming than BLAST

Distance calculator → needs ~~to~~ substitution model

↓
DNA → Protein



Starting (or)
Initial tree
[UPGMA]
[NJ]

Parsimony tree constructor

tree
searcher [NNI tree searcher]
↳ Parsimony scorer [Fitch algorithm]

starting tree [UPGMA, NJ)

Distance Matrix

Distance calculator [Substitution models]

↳ DNA [BLOSUM 30, ... Identity...]
↳ Protein [Brenner 6 ... BLOSUM--]

↳ .phy file

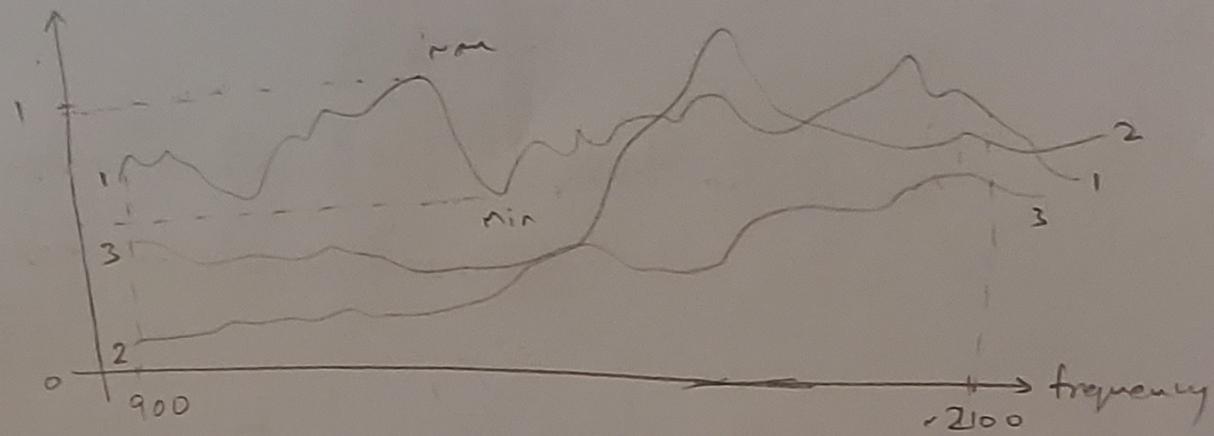
convert data into .phy file

→ protein sequence is the goal

→ so convert the reflectance data into protein sequence

(PRO)

SWIR 2



No. of wave proteins is $A \dots \rightarrow 20$ proteins

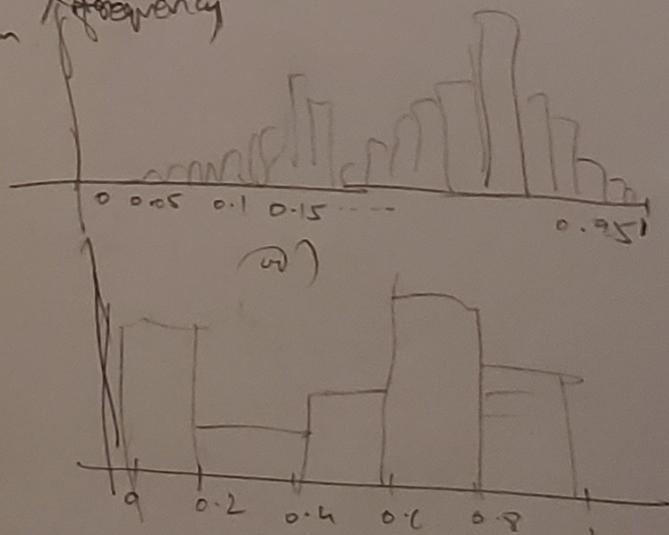
can divide $0-1$ into 20 parts

$$\Rightarrow \frac{1}{20} = 0.05 \rightarrow \text{step size}$$

$$\rightarrow 0, 0.05, 0.1, 0.15, \dots, 0.95, 1$$

next step \rightarrow build a histogram of values of reflectance ($0-1$)

example histogram \uparrow frequency



why? \leftarrow

Dec 16 2021 | Thursday | 10:05 | @ 106

- Histogram of all the reflectances of all the materials
are not plotted yet. ~~Individually yes, as a collage~~
No.
- Alternative :- download individual histograms and save, to
show it

DONE

next →

Convert all the reflectance Data to a .ply file

Steps → .ply file

5 → 10
Taxa1 → ABCDEFGHIJ → String → length 288
Taxa2 → " "
Taxa3 → " "
Taxa4 → " "
Taxa5 → "

~~(*)~~
convert each value of the reflectance array
into a letter and append to string.