

ANALIZA SIMILARITĂȚII GENETICE

Obiectiv:

- Căutarea secvențelor de aminoacizi de interes.
- Realizarea tabelului de frecvență pentru variabile calitative (frecvența absolută & frecvența relativă)
- Calcularea indicatorilor de similaritate genetică.
- Interpretarea rezultatelor indicatorilor de similaritate.

Cerințe:

1. Determinați frecvența absolută și relativă a aminoacizilor corespunzători lanțului alfa al hemoglobinei la om și cal.
2. Realizați analiza similarității lanțului alfa al hemoglobinei la om și cal. Calculați și interpretați următorii indicatori: A. Indicele Simpson, B. Indicele de diversitate Simpson, C. Reciproca Indicelui Simpson, D. Indicele de diversitate al lui Shanon.

1. Determinați frecvența absolută și relativă a aminoacizilor corespunzători lanțului alfa al hemoglobinei la om și cal

- Realizați o căutare în National Center for Biotechnology Information protein database (<http://www.ncbi.nlm.nih.gov/>) pentru a identifica secvențe de aminoacizi pentru lanțurile alfa ale hemoglobinei la om (CAA23752) și cal (P01958).
- Copiați secvențele de aminoacizi din fișierul FASTA. Realizați un fișier *.html cu structura tabelară a secvențelor de aminoacizi (alpha1-hu.html, respectiv alpha1-ho.html).
- Realizați un fișier **Excel** cu denumirea *Hemo-alpha-1.xls*. Copiați coloanele cu secvențele de aminoacizi într-o foaie de calcul denumită *Date*.
- Obțineți tabelul de frecvență absolută prin utilizarea opțiunii PivotTable and PivotChart Report ... (meniul Data).

Instrucțiuni:

- Realizarea fișierului *.html:
 - Copiați secvența de aminoacizi într-un nou fișier Notepad.
 - Salvați fișierul ca fișier *.html:



File name:	apfa-ho.html	Save
Save as type:	All Files	Cancel
Encoding:	ANSI	

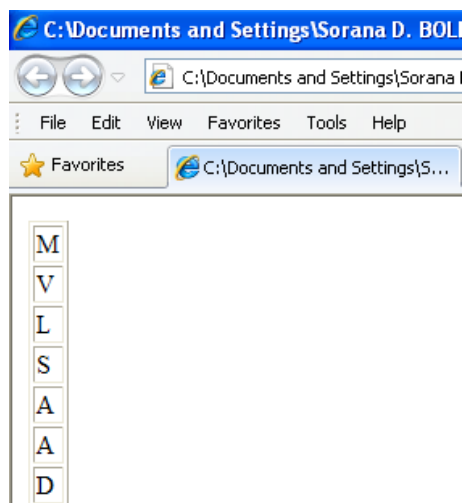
- Realizați structura tabelară a secvenței de aminoacizi după cum urmează:

```

apfa-ho - Notepad
File Edit Format View Help
<table border =1><tr><td>
M<tr><td>
V<tr><td>
L<tr><td>
S<tr><td>
...
L<tr><td>
T<tr><td>
S<tr><td>
K<tr><td>
Y<tr><td>
R
</Table>

```

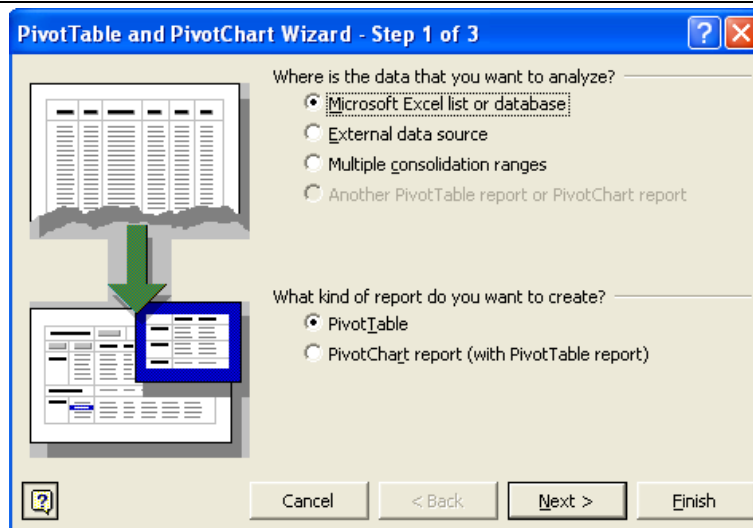
- Deschideți documentul *.html cu Internet Explorer: fișierul trebuie să fie de forma:



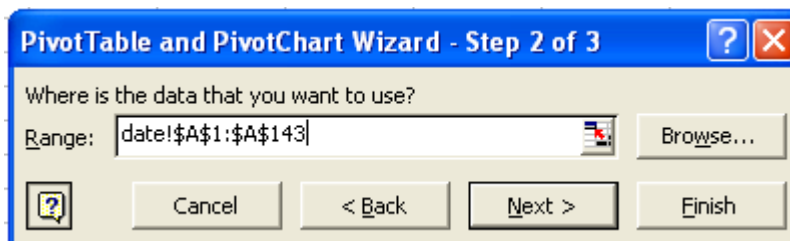
- Realizarea tabelului de frecvență în Microsoft Excel:
 - Copiați coloana cu aminoacizi pentru prima specie în coloana A a unei foi de calcul denumită Date. Inserați un nou rând și denumiți coloana în funcție de specie:

	A
1	aa-cal
2	M
3	V
4	L
5	S

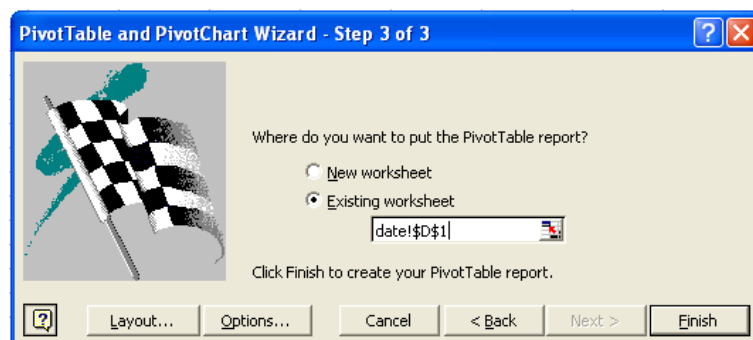
- Pentru realizarea tabelului de frecvență urmați pașii:



- selectați coloana cu date:



- definiți celula din foaia de calcul Date unde doriți să afișați tabelul de frecvență:



- veți obține următoarea structură:



- urmând instrucțiunile din imagine tabelul de frecvență va avea următoarea structură:

D	E
Count of aa-cal	
aa-cal	Total
A	16
C	1
D	9
E	3
F	7
G	10
H	10
K	11
L	21
M	2
N	4
P	6
Q	1
R	3
S	13
T	9
V	12
W	1
Y	3
Grand Total	142

- Pentru a obține frecvența relativă:

G	H	I
	Frecvența absolută	Frecvența relativă
A	16	=H3/H\$22
C	1	
D	9	
E	3	
F	7	
G	10	
H	10	
K	11	
L	21	
M	2	
N	4	
P	6	
Q	1	
R	3	
S	13	
T	9	
V	12	
W	1	
Y	3	
Total	142	

- Tabelul de frecvență trebuie să arate ca și cel din figura următoare:

G	H	I
AA	Frecvența absolută	Frecvența relativă
A	16	0.11
C	1	0.01
D	9	0.06
E	3	0.02
F	7	0.05
G	10	0.07
H	10	0.07
K	11	0.08
L	21	0.15
M	2	0.01
N	4	0.03
P	6	0.04
Q	1	0.01
R	3	0.02
S	13	0.09
T	9	0.06
V	12	0.08
W	1	0.01
Y	3	0.02
Total	142	1

- Procedați similar pentru a obține tabelul de frecvență și pentru cea de-a doua specie.

2. Realizați analiza similarității lanțului alfa al hemoglobinei la om și cal.

- Creați în același fișier Excel utilizat pentru prima cerință o nouă foaie de calcul denumită *Similaritate*. Copiați în această foaie de calcul, într-o singură structură tabelară frecvențele absolute.
- Calculați pentru fiecare specie, utilizând formulele de calcul din curs, indicatorii ceruți la punctele A-D.
- Interpretați rezultatele obținute.
- Salvați fișierul.

Trimiteți fișierul creat ca fișier atașat la adresa de e-mail: sbolboaca@gmail.com.

La subiect: Prenume, NUME, Master BiodivBioconserv – Laborator 5.