Analiza Similarițății Genetice

Obiectiv:

- Căutarea secvențelor de aminoacizi de interes.
- Realizarea tabelului de frecvenţă pentru variabile calitative (frecvenţa absolută & frecvenţa 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 (<u>http://www.ncbi.nlm.nih.gov/</u>) 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*.
- Obţineţi tabelul de frecvenţă absolută prin utilizarea opţiunii PivotTable and PivotChart Report ... (meniul Data).

Instrucțiuni:

- Realizarea fişierului *.html:
 - Copiați secvența de aminoacizi într-un nou fișier Notepad.
 - Salvaţi fişierul ca fişier *.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:

	🗾 a	ıpfa-l	ho - Not	epad		
	File	Edit	Format	View	Help	
	<ta M<t V<t L<t S<t< th=""><th>ble r><t r><t r><t< th=""><th>border :d> :d> :d> :d></th><th>' =1></th><th>≺tr>∢</th><th></th></t<></t </t </th></t<></t </t </t </ta 	ble r> <t r><t r><t< th=""><th>border :d> :d> :d> :d></th><th>' =1></th><th>≺tr>∢</th><th></th></t<></t </t 	border :d> :d> :d> :d>	' =1>	≺tr>∢	
•						
	L <t T<t S<t K<t Y<t< th=""><th>r><t r><t r><t r><t r><t< th=""><th>:d> :d> :d> :d> :d> :d></th><th></th><th></th><th></th></t<></t </t </t </t </th></t<></t </t </t </t 	r> <t r><t r><t r><t r><t< th=""><th>:d> :d> :d> :d> :d> :d></th><th></th><th></th><th></th></t<></t </t </t </t 	:d> :d> :d> :d> :d> :d>			

R

</Table>

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



- Realizarea tabelului de frecvență în Miscrosoft 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	\vee
4	L
5	S

• Pentru realizarea tabelului de frecvență urmați pașii:

PivotTable and PivotChart Wizard - Step 1 of 3				
	Where is the data that you want to analyze? Microsoft Excel list or database External data source Multiple <u>c</u> onsolidation ranges Another PivotTable report or PivotChart report			
	What kind of report do you want to create? Pivot <u>T</u> able PivotCha <u>r</u> t report (with PivotTable report)			
2	Cancel < Back <u>N</u> ext > <u>F</u> inish			

- selectați coloana cu date:

PivotTable and PivotChart Wizard - Step 2 of 3				? 🗙
Where is the data that you want to use?				
Range: date!\$A\$1:\$A\$143			Browse	
2	Cancel	< <u>B</u> ack	<u>N</u> ext >	Einish

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

PivotTable and PivotChart Wizard - Step 3 of 3					
	Where do yo C <u>N</u> Click Finish to	ou want to put ti lew worksheet xisting workshee date!\$D\$1 o create your Pi	ne PivotTable r st 	eport?	
Layout Op	tions	Cancel	< <u>B</u> ack	Next >	<u>F</u> inish

- veți obține următoarea structură:

	Drop Column Fields Here
Drop Row Fields Here	Drop Data Items Here

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

D	E
Count of aa-cal	
aa-cal 🗸 🔻	Total
A	16
С	1
D	9
E	3
F	7
G	10
Н	10
К	11
L	21
M	2
N	4
Р	6
Q	1
R	3
S	13
Т	9
V	12
W	1
Υ	3
Grand Total	142

• Pentru a obține frecvența relativă:

G	Н	l I
	Frecventa	Frecventa
	absoluta	relativa
A	16	=H3/H\$22 [
С	1	Ĩ
D	9	
E	3	
F	7	
G	10	
Н	10	
К	11	
L	21	
M	2	
N	4	
Р	6	
Q	1	
R	3	
S	13	
Т	9	
V	12	
W	1	
γ	3	
Total	142	

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

G	H	I
^^	Frecventa	Frecventa
	absoluta	relativa
A	16	0.11
С	1	0.01
D	9	0.06
E	3	0.02
F	7	0.05
G	10	0.07
Н	10	0.07
К	11	0.08
L	21	0.15
M	2	0.01
N	4	0.03
Ρ	6	0.04
Q	1	0.01
R	3	0.02
S	13	0.09
Т	9	0.06
V	12	0.08
W	1	0.01
γ	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.
- Salvaţi fişierul.

Trimiteți fișierul creat ca fișier atașat la adresa de e-mail: <u>sbolboaca@gmail.com</u>. La subiect: Prenume, NUME, Master BiodivBioconserv – Laborator 5.