



Лабораторна робота № 2

Тема: Кореляція виродженості генетичного коду з частотою зустрічності амінокислот у складі білків

Мета: ознайомлення з властивостями генетичного коду, еволюцією білків та кодуєчих генів

Завдання 1. Виродженість генетичного коду

Виродженість – одна з основних властивостей генетичного коду, яка полягає в тому, що всі амінокислоти, крім метіоніну та триптофану, кодується двома або більшою кількістю триплетів.

Коефіцієнт виродженості: $K = M / 61$, де M – число кодонів, що відповідають даній амінокислоті, 61 – загальне число значущих кодонів. За значенням коефіцієнту виродженості усі амінокислоти можна розбити на 5 груп:

Група	Амінокислоти	M	K	N
I	метіонин (мет.), триптофан (три)	1	0,0166	2
II	глутамінова кислота (глу), глутамін (глан), аспарагінова кислота (асп), аспарагін (асн), лізин (ліз), цистеїн (цис), фенілаланін (фен), тирозин (тир), гістидин (гіс)	2	0,0322	9
III	Ізолейцин (ілс)	3	0,0498	1
IV	гліцин (глі), аланін (ала), валін (вал), треонін (тре), пролін (про)	4	0,0664	5
V	лейцин (лей), серин (сер), аргінін (арг)	6	0,0996	3

Завдання 2. Підрахунок виродженості

Використовуючі надані послідовності амінокислот у білках потрібно розрахувати n_1 – кількість амінокислот, що відносяться до кожної з 5 груп. Потім необхідно підрахувати q_1 – теоретично очікуване число радикалів, що відносяться до цих груп. Для цього загальне число амінокислот у складі білку треба помножити на нормований множник $K \cdot N$: $q_1 = Z \cdot K \cdot N$, де Z – загальне число амінокислотних залишків у цьому білку.

Потім необхідно розрахувати величину відхилення результату, що спостерігається, від теоретично очікуваного: $d_1 = n_1 - q_1$.

Отримані дані занести до таблиці 1 окремо для кожного з запропонованих білків (див. кінець методичних вказівок).

Таблиця 1. Результати розрахунків виродженості коду



ВП 3 ЗАГАЛЬНОЇ ГЕНЕТИКИ



Перелік амінокислот	Число амінокисл. цього типу (M)	Число амінокис. даної групи (n ₁)	Теорет. очікуване число амінокис. даної групи (q ₁)	Відхилення	d ₁ ²	d ₁ ² /q ₁
мет						
три						
глх (глу+глан)						
асх (асп+асн)						
ліз						
цис						
фен						
тир						
гіс						
іле						
глі						
ала						
вал						
тре						
про						
лей						
сер						
арг						

Завдання 3. Оцінка кореляції виродженості генетичного коду з частотою зустрічальності амінокислот у складі білку.

Оцінити кореляцію виродженості генетичного коду з частотою зустрічальності амінокислот у складі білка за критерієм χ^2 :

$$\chi^2 = \sum d_1^2 / q_1.$$

Табличне значення χ^2 при рівні значущості 0,05 дорівнює 9,488. Якщо наше χ^2 буде більше табличного, то можна вважати доведеним існування кореляції виродженості генетичного коду з частотою зустрічальності амінокислот у складі білка.



ВП 3 ЗАГАЛЬНОЇ ГЕНЕТИКИ



Таблиця кодування амінокислот

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid (Aspartate)
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid (Glutamate)
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine
Asx	B	Aspartic acid or Asparagine
Glx	Z	Glutamine or Glutamic acid
Xaa	X	Any amino acid
TERM		termination codon

Білок 1.

MVSSFLLAEYKNLIVNPSEHFKISVNEDNLTEGPPDTLYQKIDTVLLSVISLLNEPNP
 DSPANVDAAKSYRKYLYKEDLESYPMEKSLDECSAEDIEYFKNVPVNVLPVPSDDY
 EDEEMEDGTYILTYDDEDEEEDDEEMDDE

Білок 2.

MTYELPKLPYTYDALEPNFDKETMEIHYTEKHHNIYVTKLNEAVSGHAELASKPGEEL
 VANLDSVPEEIRGAVRNHGGGHANHTLFWSSLSPNGGGAPTGNLKAAIESEFGTF
 DEFKEKFNAAAAARFGSGWAWLVVNNGKLEIVSTANQDSPLSEGKTPVLGLDVWE
 HAYYLFQNRPEYIDTFWNVINWDERNKRFDAAK



ВП 3 ЗАГАЛЬНОЇ ГЕНЕТИКИ



Білок 3.

MVSSFLLAEYKNLIVNPSEHFKISVNEDNLTEGPPDTLYQKIDTVLLSVISLLNEPNP
DSPANVDAAKSYRKYLYKEDLESYPMEKSLDECSAEDIEYFKNVPVNVLPVPSDDY
EDEEMEDGTYILTYDDEDEEEDEEVDDE

Білок 4.

MTYELPKLPYTYDALEPNFDKETMEIHYTEKHHNIYVTKLNEAVSGHAELASKPGEEL
VANLDSVPEEIRGAVRNHGGGHANHTLFWSSLSPNGGGAPTGNLKAAIESEFGTF
DEFKEKFNAAAAARFGSGWAWLVVNNGKLEIVSTANQDSPLSEGKTPVLGLDVWE
HAYYLKFQNRREYIDTFWNVINWDERNKRFDAAKMTYELPKLPYTYDALEPNFDK
ETMEIHYTEKHHNIYVTKLNEAVSGHAELASKPGEELVANLDSVPEEIRGAVRNHGG
GHANHTLFWSSLSPNGGGAPTGNLKAAIESEFGTFDEFKEKFNAAAAARFGSGWA
WLVVNNGKLEIVSTANQDSPLSEGKTPVLGLDVWEHAYYLKFQNRREYIDTFWN
VINWDERNKRFDAAKMTYELPKLPYTYDALEPNFDKETMEIHYTEKHHNIYVTKLNE
AVSGHAELASKPGEELVANLDSVPEEIRGAVRNHGGGHANHTLFWSSLSPNGGGA
PTGNLKAAIESEFGTFDEFKEKFNAAAAARFGSGWAWLVVNNGKLEIVSTANQDSP
LSEGKTPVLGLDVWEHAYYLKFQNRREYIDTFWNVINWDERNKRFDAAK