

Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Кафедра молекулярної і медичної біофізики

“ЗАТВЕРДЖУЮ”

Проректор з науково-педагогічної  
Роботи

Антон ПАНТЕЛЕЙМОНОВ



20\_\_ р.

Робоча програма навчальної дисципліни

БІОІНФОРМАТИКА

(назва навчальної дисципліни)

рівень вищої освіти другий (магістерський)

галузь знань 10 Природничі науки  
(шифр і назва)

спеціальність 105 Прикладна фізика та наноматеріали  
(шифр і назва)

освітня програма " Біофізика "  
(шифр і назва)

спеціалізація \_\_\_\_\_  
(шифр і назва)

вид дисципліни обов'язкова  
(обов'язкова / за вибором)

факультет радіофізики, біомедичної електроніки та комп'ютерних систем

2021 / 2022 навчальний рік

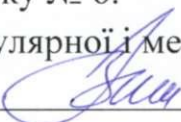
Програму рекомендовано до затвердження вченою радою факультету радіофізики, біомедичної електроніки та комп'ютерних систем.  
25 червня 2021 року, протокол №8.

**РОЗРОБНИКИ ПРОГРАМИ:**

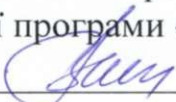
Олексій БОРІКОВ, кандидат біологічних наук, доцент кафедри молекулярної і медичної біофізики.

Програму схвалено на засіданні кафедри молекулярної і медичної біофізики.  
Протокол від 07 червня 2021 року № 6.

Завідувач кафедри молекулярної і медичної біофізики

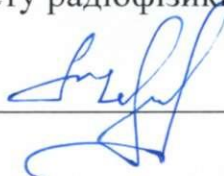
  
\_\_\_\_\_ Володимир БЕРЕЗ

Програму погоджено з гарантом освітньо-професійної програми «Біофізика».  
Гарант освітньо-професійної програми «Біофізика»

  
\_\_\_\_\_ Володимир БЕРЕЗ

Програму погоджено методичною комісією факультету радіофізики, біомедичної електроніки та комп'ютерних систем.  
Протокол від 17 червня 2021 року № 6.

Голова методичної комісії факультету радіофізики, біомедичної електроніки та комп'ютерних систем

  
\_\_\_\_\_ Леонід ЧОРНОГОР

## ВСТУП

Програма навчальної дисципліни «**Біоінформатика**» складена відповідно до освітньо-професійної програми підготовки

другий рівень вищої освіти / магістр

спеціальність 105 Прикладна фізика та наноматеріали

Освітньо-професійна програма «Біофізика»

### 1. Опис навчальної дисципліни

1.1. Мета викладання навчальної дисципліни «БІОІНФОРМАТИКА» полягає в ознайомленні з сучасними концепціями біоінформатики, дати студентам цілісне уявлення про структуру та методи аналізу біологічних послідовностей на молекулярному рівні та на рівні організмів, структуру та методи аналізу просторових структур біологічних молекул, структуру та методи комп'ютерного аналізу геномів, сформувати у студентів цілісний і системний погляд на організацію біологічної інформації на молекулярному рівні.

1.2 Основні завдання вивчення дисципліни «БІОІНФОРМАТИКА» - надання необхідного обсягу знань в області застосування біологічної інформації з метою її застосування для розуміння аналізу структури біологічних молекул, геномів організмів різного рівня організації.

1.3. Кількість кредитів – 4

1.4. Загальна кількість годин - 120

1.5. Характеристика навчальної дисципліни	
Нормативна / за вибором	
Денна форма навчання	Заочна (дистанційна) форма навчання
Рік підготовки	
2-й	-й
Семестр	
3-й	-й
Лекції	
12 год.	-год.
Практичні заняття	
24 год.	- год.
Самостійна робота	
84 год.	-год.

1.6. Заплановані результати навчання

*Знати:* основні концепції аналізу біологічних текстів, основні концепції відтворення і аналізу просторової організації біомолекул, основи організації цілих геномів та методи їх порівняльного аналізу.

*Вміти:* аналізувати та порівнювати біологічні тексти, працювати з банками даних біологічних послідовностей і просторових структур, здійснювати парне та множинне вирівнювання послідовностей, реконструювати просторову структуру, розраховувати поведінку і аналізувати особливості просторової структури білків, визначати білок-кодуєчі ділянки в нуклеотидних послідовностях, проводити філогенетичний аналіз, вміти цілісно і системно мислити.

## 2. Тематичний план навчальної дисципліни

**Розділ 1.** Основні поняття біоінформатики. Класична біоінформатика.

**Тема 1.** Визначення та предмет досліджень біоінформатики. Основні розділи біоінформатики. Особливості та труднощі обчислень в різних галузях біоінформатики. Застосування результатів досліджень біоінформатики в різних галузях біології та медицини.

**Тема 2.** Первинні біологічні дані. Банки та бази біологічних послідовностей. Основні характеристики банків даних. Формати даних біологічних послідовностей. – 1 години.

**Тема 3.** Парне вирівнювання біологічних послідовностей. Якість вирівнювання. Матриці амінокислотних замінів. Пошук гомологів в банках даних. Основні алгоритми пошуку гомологічних послідовностей.

**Тема 4.** Множинні вирівнювання біологічних послідовностей. Значення та алгоритми. Вимога до однозначності множинного вирівнювання.

**Розділ 2.** Структурна біоінформатика (комп'ютерна структурна біологія).

**Тема 5.** Структурна біоінформатика. Ієрархічність властивостей живих систем. Зв'язок “послідовність-просторова структура”. Особливості просторової організації структури біополімерів. Стабілізація просторової організації макромолекул.

**Тема 6.** Особливості представлення даних, що описують просторову структуру макромолекул. Формати файлів. Банки даних просторових структур. Методи визначення просторової структури макромолекул. Візуалізація молекулярних даних.

**Тема 7.** Методи відтворення просторової структури білків. Реконструкція *ab initio* (конформаційний пошук). Реконструкція за гомологією. Однозначний зв'язок між первинною та третинною структурою.

**Тема 8.** Структура білків та еволюція. Доменний склад і доменна архітектура. Родини та надродини білків. Порівняння просторової структури макромолекул та просторове вирівнювання. Передбачення і пошук потенційно активних сайтів на поверхні білків.

## 3. Структура навчальної дисципліни

Назви розділів і тем	Кількість годин			
	денна форма			
	усього	у тому числі		
л		прак	і р	
<b>Розділ 1. Основні поняття біоінформатики. Класична біоінформатика.</b>				
Тема 1. Мета, задачі і методи біоінформатики, загальні застосування.	1	1	-	
Тема 2. Банки та бази даних.	5	1	4	
Тема 3. Вирівнювання біологічних послідовностей. Парне вирівнювання	6	2	4	
Тема 4. Вирівнювання біологічних послідовностей. Множинне вирівнювання	6	2	4	
Разом за розділом 1	18	6	12	
<b>Розділ 2. Структурна біоінформатика (комп'ютерна структурна біологія).</b>				
Тема 5. Особливості просторової організації структури біополімерів.	1	1	-	
Тема 6. Зв'язок між послідовністю та просторовою	1	1	-	

структурою біомакромолекул.				
Тема 7. Банки даних просторових структур. Порівняння послідовностей та просторове вирівнювання.	6	2	6	
Тема 8. Структура білків та еволюція. Передбачення і пошук потенційно активних сайтів на поверхні білків.	6	2	6	
Разом за розділом 2	18	6	12	

#### 4. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Знайомство з банками даних GenBank, EMBL, Uniprot. Пошук послідовностей білків та нуклеїнових кислот за різними критеріями.	2
2	Знайомство з програмним забезпеченням, що використовується для вирівнювання послідовностей (ClustalX, T-Coffee, MEGA). Парне вирівнювання послідовностей білків різного ступеня консервативності з використанням різних вагових матриць амінокислотних замінів (BLOSSUM, PAM, Gonnet, матриця ідентичності).	4
3	Знайомство з системою SRS (Sequence Retrieve System). Пошук у банках GenBank, EMBL, Uniprot, GeneBee за допомогою SRS послідовностей нуклеїнових кислот та білків, гомологічних до заданих, з використанням серій алгоритмів BLAST і FASTA.	4
4.	Множинні вирівнювання послідовностей білків та нуклеїнових кислот за допомогою алгоритмів NJ і UPGMA (використовується програма ClustalX). Візуалізація вирівнювання (DS Vizualizer). Візуалізація філогенетичних дерев за допомогою TreeView та NjPlot. Порівняння дерев, побудованих на базі різних алгоритмів вирівнювання.	2
5.	Огляд форматів PDB, gro, mol2 та hin. Пошук в Protein Data Bank даних щодо просторових структур заданих білків за різними критеріями. Знайомство з засобами візуалізації молекулярних даних: RasMol, PyMol, DS Vizualizer, візуалізація білків, полінуклеотидів та низькомолекулярних сполук.	4
6.	Знайомство з настільними (MODELLER, SwissPDB Viewer) та веб базованими (3D-Jigsaw, SwissModel) засобами гомологічного моделювання білків. Пошук Міжнародному банку білкових структур білків, які є оптимальними матрицями для моделювання просторової структури заданих білків. Використання знайдених матриць для реконструкції просторової структури в MODELLER і SwissPDB Viewer.	4
7.	Визначення доменної архітектури заданих білків за допомогою веб-інструментів SMART, Pfam, PROSITE, HMM. Знайомство з банками даних SCOP та CATH. Визначення за допомогою SCOP та CATH типів упаковки заданих білків.	2
8.	Пошук та розшифровка генів за допомогою інструментів GENMARK та ORF Finder.	2
	Разом	24

## 5. Індивідуальні завдання

№ з/п	Види, зміст самостійної роботи	Кількість годин
1	Основні розділи біоінформатики. Особливості та труднощі обчислень в класичній та структурній біоінформатиці.	2
2	Банки даних біологічних послідовностей.	2
3	Формальне визначення парного вирівнювання. Якість вирівнювання. Вагові матриці.	2
4	Особливості просторової організації і структури біополімерів.	2
5	Методичні підходи до визначення просторової структури макромолекул.	2
6	Подвійне та множинне вирівнювання послідовностей.	4
7	Білок-білкові та білок нуклеїнові взаємодії. Біополімери як молекулярні мішені для дії терапевтичних препаратів.	5
8	Сучасний стан вивченості геномів. Повні та неповні геноми.	2
9	Пошук регуляторних послідовностей геномів.	4
10	Генні мережі.	2
Разом		27

## 6. Методи контролю

Контрольні роботи, практичні завдання.

## 7. Схема нарахування балів

Поточний контроль, самостійна робота, індивідуальні завдання								Контрольна робота, передбачена навчальним планом		Розрахунково-графічне завдання	Екзамен (залікова робота)	Сума
Розділ 1				Розділ 2				Розділ 1	Розділ 2			
T1	T2	T3	T4	T1	T2	T3	T4	Розділ 1	Розділ 2			
5	5	5	5	5	5	5	5	10	10	10	30	100

## 8. Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка	
	для чотирирівневої шкали оцінювання	для дворівневої шкали оцінювання
90 – 100	відмінно	зараховано
70-89	добре	
50-69	задовільно	
1-49	Незадовільно	не зараховано

## **Критерії оцінки успішності студентів при семестровому контролі та виконанні письмових робіт**

**Оцінку «відмінно» (5 балів – за завдання; 90-100 балів за курс у цілому)** отримує студент, якщо він:

- міцно засвоїв зміст навчальної дисципліни, наукових першоджерел і рекомендованої літератури;
- вміє повністю, глибоко і всебічно розкрити зміст матеріалу, поставленого завдання чи проблеми; комплексно вирішувати поставлені завдання чи проблему; правильно застосовує одержані знання з різних дисциплін для вирішення завдань чи проблем; послідовно і логічно викладає матеріал;
- висловлює обґрунтоване власне ставлення до тих чи інших проблем;
- чітко розуміє зміст і вільно володіє спеціальною термінологією; встановлює взаємозв'язок основних понять;
- грамотно ілюструє відповіді прикладами;
- вільно використовує набуті теоретичні знання для аналізу практичного матеріалу; демонструє високий рівень набутих практичних навичок.

Допускається декілька неточностей у викладенні матеріалу, які не приводять до помилкових висновків і рішень.

**Оцінку «добре» (4 бали – за завдання; 70-89 балів за курс у цілому)** отримує студент, якщо він:

- добре засвоїв основний зміст навчальної дисципліни, основні ідеї наукових першоджерел і рекомендованої літератури;
- аргументовано, правильно та послідовно розкриває основний зміст матеріалу;
- висловлює власні міркування з приводу тих чи інших проблем;
- точно використовує термінологію;
- має практичні навички з аналізу матеріалу.

Допускається декілька неточностей у використанні спеціальної термінології, похибок у логіці викладу теоретичного змісту або аналізу практичного матеріалу, несуттєвих та не грубих помилок у висновках та узагальненнях, що не впливають на конкретний зміст відповіді. Наявні неточності та помилки враховуються при визначенні оцінки за 100-бальною шкалою та відповідної літери В або С.

**Оцінку «задовільно» (3 бали – за завдання; 50-69 балів за курс у цілому)** студент отримує, якщо:

- у відповіді суть запитання в цілому розкрита, але зміст питання викладено частково; студент невпевнено орієнтується у змісті наукових першоджерел та рекомендованої літератури;
- матеріал викладений не завжди послідовно, висновки не ув'язані між собою;
- не вміє обґрунтовано оцінювати факти та явища, пов'язувати їх з майбутньою професійною діяльністю;
- при викладенні матеріалу, поясненні термінології та вирішенні практичних питань зроблені суттєві помилки.

**Оцінку «незадовільно» (менше 50 балів)** студент отримує, якщо:

- основний зміст завдання не розкрито; студент майже не орієнтується у наукових першоджерелах та рекомендованій літературі; не знає наукових фактів та визначень;
- допущені суттєві помилки у висновках;
- студент слабо володіє спеціальною термінологією;
- наукове мислення та практичні навички майже не сформовані.

## 9. Рекомендована література

### Основна література

1. Introduction to bioinformatics, 3rd edition by Arthur M. Lesk, Oxford university press, 2008, 432 pp., ISBN 978-0-19-920804-3.
2. Bioinformatics. Sequence, Structure and Databanks. A Practical Approach, Des Higgins, Willie Taylor, Oxford University Press, 2000, 272pp., ISBN 0199637903.
3. Bioinformatics. A Practical Guide to the Analysis of Genes and Proteins Andreas D. Baxeavanis, B. F. Francis Ouellette Wiley, John & Sons, Incorporated, 2004, 540 pp., ISBN 0471478784.

### Допоміжна література

1. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs 1997 <http://nar.oupjournals.org/cgi/content/full/25/17/3389>
2. Geoffrey J. Barton Protein Sequence Alignment and Database Scanning 1996 [http://barton.ebi.ac.uk/barton/papers/rev93\\_1/rev93\\_1.html](http://barton.ebi.ac.uk/barton/papers/rev93_1/rev93_1.html)
3. Altschul S.F. Applications and statistics for multiple high-scoring segments in molecular sequences *Karlin S.*, 1990 <http://www.pnas.org/cgi/reprint/90/12/5873>
4. Hoyle D. Shared Genomics: Developing an accessible integrated analysis platform for Genome-Wide Association Studies / Hoyle D., Delderfield M., Kitching L. [et al.] // AMIA Summits Transl Sci Proc. – 2010 Mar 1. – 2010. – P. 18-22.

## 10. Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. <http://beta.uniprot.org/> – SWISS-PROT|UniProt the protein sequence data bank, база даних UniProt
2. <http://www.ebi.ac.uk/uniprot/> – база даних UniProt на сервері Європейського інституту біоінформатики (European Bioinformatics Institute, EBI)
3. <http://www.expasy.org/sprot/> – бази даних Swiss-Prot, TrEmbl, UniProt на сервері ExPASy (Expert Protein Analysis System) Швейцарського Інституту Біоінформатики SIB
4. <http://www.rcsb.org/> – Protein Data Bank, база даних PDB.
5. <http://www.ncbi.nlm.nih.gov/> (<http://www.pubmed.com/>) – сервер Національного центру біотехнологічної інформації США (NCBI): бази даних GenBank, NCBI Protein Database, UniGene, HomoloGene та ін.
6. <http://cmm.info.nih.gov/modeling/> – сервер Центру моделювання молекул Національного Інституту Здоров'я НІН, США
7. <http://www.genebio.com/> – сайт компанії GeneBio (Geneva Bioinformatics S.A.), що розповсюджує інформацію з протеомних баз даних: SWISS-PROT, PROSITE, SWISS-2DPAGE та відповідне програмне забезпечення.