

Одеський національний університет імені І. І. Мечникова

(повна назва вищого навчального закладу)

Факультет/інститут біологічний факультет

Кафедра Кафедра мікробіології, вірусології та біотехнології

“ЗАТВЕРДЖУЮ”

Проректор з науково-педагогічної роботи

_____ (_____)

“ _____ ” _____ 20__ р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

БІОІНФОРМАТИКА

Рівень вищої освіти: перший (бакалаврський)

Спеціальність: 162 Біотехнології та біоінженерія

Інститут/факультет: Біологічний

Робоча програма складена на основі навчальної програми з дисципліни «Біоніформатика» .

Розробники: Васильєва Н.Ю., к.б. н., ст.н.спів. кафедри мікробіології, вірусології та біотехнології

Робоча програма затверджена на засіданні кафедри мікробіології, вірусології та біотехнології.

Протокол № ____ від. “ ____ ” _____ 20__ р.

Завідувач кафедри _____ (_____)
(прізвище та ініціали) (підпис)

Схвалено навчально-методичною комісією (НМК) біологічного факультету

Протокол № ____ від. “ ____ ” _____ 20__ р.

Голова НМК _____ (_____)
(прізвище та ініціали) (підпис)

Переглянуто та затверджено на засіданні кафедри мікробіології, вірусології та біотехнології

Протокол № ____ від. “ ____ ” _____ 20__ р.

Завідувач кафедри _____ (_____)
(прізвище та ініціали) (підпис)

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, спеціальність, спеціалізація, рівень вищої освіти	Характеристика навчальної дисципліни	
		денна форма навчання	заочна форма навчання
Загальна кількість: кредитів – 3 годин – 90 залікових модулів – 3 змістових модулів – 3 ІНДЗ* – _____ (вид завдання)	Галузь знань <u>16 Хімічна та біоінженерія</u> (шифр і назва) Спеціальність <u>162 Біотехнології та біоінженерія</u> (код і назва) Рівень вищої освіти: Перший (бакалаврський)	Нормативна / за вибором (ВНЗ/студента)	
		Рік підготовки:	
		4-й	
		Семестр	
		7-й	
		Лекції	
		20 год.	.
		Практичні, семінарські	
		34 год.	
		Лабораторні	
		-	-
		Самостійна робота	
		36 год.	-
		у т.ч. ІНДЗ*: -	
		Форма підсумкового контролю: іспит	

* – за наявності

2. Мета та завдання навчальної дисципліни

Мета курсу полягає в навчанні студентів основам аналізу експериментальних даних молекулярної біології: секвінування послідовностей (теоретичний огляд), пошук послідовностей у базах даних різного типу, проведення BLAST, проведення вирівнювання послідовностей різного типу т.д.

Ознайомлення студентів з такими методами біоінформатики як: методи організації інформації, методами обчислювальної математики і особливо використанню існуючого програмного забезпечення. Курс покликаний готувати фахівців в галузі “комп'ютерної біології”, який займається отриманням біологічно осмислених результатів при дослідженні вмісту банків даних.

Поетапне досягання поставленої мети формує фахові компетентності у здобувачів вищої освіти спеціальності 162 «Біотехнології та біоінженерія».

Завдання курсу:

- 1) введення студентів в курс сучасного стану науки у зв'язку із стрімким наростанням молекулярно-біологічних і молекулярно-генетичних даних, пов'язаного з масовою розшифровкою повних геномних послідовностей;
- 2) ознайомити студентів з основними банками та базами даних та засобами для роботи з ними;
- 3) ознайомити студентів з сучасним підходом до передбачення функцій генів і пошук структурних і функціональних особливостей геномів на основі порівняння багатьох геномів;
- 4) побудова на основі отриманих з баз даних послідовностей дерев філогенезу на основі математичних розрахунків моделей еволюції. Вивчення основних алгоритмів побудови дерев філогенезу.

Процес вивчення дисципліни спрямований на формування елементів наступних компетентностей:

а) загальних (ЗК):

1. Здатність застосовувати знання у практичних ситуаціях (K01).
2. Навички використання інформаційних і комунікаційних технологій (K04).
3. Здатність вчитися і оволодівати сучасними знаннями (K05).

б) фахових загальних (КФЗ):

1. Здатність використовувати знання з математики та фізики в обсязі, необхідному для досягнення інших результатів освітньої програми (K10).
2. Здатність використовувати сучасні автоматизовані системи управління виробництвом біотехнологічних продуктів різного призначення, їх технічне, алгоритмічне, інформаційне і програмне забезпечення для вирішення професійних завдань (K23).
3. Здатність використовувати комп'ютерну інформаційну базу даних у плануванні, проведенні та оптимізації біотехнологічних досліджень, володіти базовими навичками в галузі біоінформатики та математичного моделювання біотехнологічних процесів (K27).

Нормативний зміст підготовки бакалавра у процесі вивчення дисципліни, сформульований у термінах результатів навчання:

1. Вміти розраховувати основні критерії оцінки ефективності біотехнологічного процесу (параметри росту біологічних агентів, швидкість синтезу цільового продукту, синтезувальна здатність біологічних агентів,

економічний коефіцієнт, вихід цільового продукту від субстрату, продуктивність, вартість поживного середовища тощо) (ПР20).

2. Використовувати професійно-профільовані знання в галузі математичної статистики, біоінформатики й практичного використання комп'ютерних технологій для статистичної обробки експериментальних даних і математичного моделювання біотехнологічних явищ і процесів (ПР27)

На вивчення навчальної дисципліни відводиться 90 годин, що становить 3 кредитів ЄКТС.

3. Зміст навчальної дисципліни

Змістовий модуль 1. Загальні поняття і постулати біоінформатики

Тема 1. Біоінформатика: становлення галузі науки та основні поняття

Введення у біоінформатику. Історія розвитку та становлення сучасної біоінформатики. Інформаційний вибух в молекулярній біології і генетиці. Виникнення інформаційної біології. Актуальні завдання біоінформатики та інформаційної біології.

Тема 2. Математичний аналіз генетичного коду

Генетичний код як система. Особливості генетичного коду з точки зору біоінформатики. Статистичні залежності, які були знайдені при математичному аналізі генетичного коду. ORF. Пошук ORF у непримовній послідовності генетичного тексту.

Змістовий модуль 2. Біологічні бази даних. Знайомство з біологічними базами даних. Філогенія.

Тема 3. Біологічні бази даних

Основні бази даних. Бази даних послідовностей ДНК. Класифікація баз даних біологічних послідовностей. Поняття про основні бази даних та практичне ознайомлення з ними. Основи структур баз даних (записи, поля, об'єкти). Класифікація баз за способом заповнення (автоматичні, архівні, курирувані). Основні бази даних: GenBank, EMBL, SwissProt, NCBI, PDB. Засоби роботи з банками даних: SRS, Entrez. Програми для пошуку гомологів. BLAST.

Тема 4. Парне та множинне вирівнювання

Типи вирівнювання. Локальне та глобальне вирівнювання. Парне та множене вирівнювання. Алгоритм Нідлмана — Вунша. Програми для проведення вирівнювання різного типу: ClustlW, T-Coffe, Muscul, needle і т.д. Програми для аналізу генетичного тексту.

Тема 5. Методи молекулярної філогенії

Заміни нуклеотидів в послідовності ДНК. Синонімічні та несинонімічні заміни. Молекулярна еволюція: об'єкти і завдання. Розділи молекулярної еволюції. Теорія нейтральної молекулярної еволюції. Теорія мутаційного тиску. Методи молекулярної еволюції. Методи молекулярної філогенії. Визначення еволюційної відстані. Одне параметрична модель Джукса-Кантора та двох параметрична модель Кімури. Матриці амінокислотних замін.

Тема 6. Філогенія. Побудова дерев

Гомологи, дерева, еволюція. Що таке філогенетичне дерево. Дерево видів і дерево послідовностей. Довжини гілок і їхній зміст. Типи дерев. Алгоритми реконструкції філогенетичних дерев.

Змістовий модуль 3. Використання програми R для завдань біоінформатики

Тема 7. Статистичний аналіз генетичного тексту за допомогою програми R.

Використання пакетів seqinr та ad4. Завантаження даних за допомогою функції read.fasta. Функції, що сприяють статистичному генетичного тексту.

Тема 8. Парне глобальне та локальне вирівнювання послідовностей в програмі R

Парне глобальне вирівнювання послідовностей з використанням алгоритму Needleman-Wunsch. Використання пакету Biostrings. Використання додаткових функцій nucleotideSubstitutionMatrix, pairwiseAlignment та інших. Парне локальне вирівнювання послідовностей з використанням алгоритму Smith-Waterman.

Тема 9. Побудова філогенетичних дерев в R

Використання пакетів seqinr, phangorn та інші. Типи файлів, які використовують для побудови філогенетичних дерев в R. Побудова дерев методом Neighbor-Joining. Використання функцій read.alignment, dist.alignment, nj, plot. Використання додаткових функцій unrootedNJtree, cleanAlignment, rootedNJtree. Побудова дерев методом UPGMA. Використання функції upgma. Побудова дерев в R за допомогою кластерного аналізу. Використання пакету dendextend та інші.

4. Структура навчальної дисципліни

Назви тем	Кількість годин									
	Денна форма					Заочна форма				
	Усього	у тому числі				Усього	у тому числі			
		л	п/с	лаб	ср		л	п/с	лаб	ср
1	2	3	4	5	6	7	8	9	10	11
Змістовий модуль 1. Основні біоінформатики.напрями										
Тема 1. Біоінформатика: становлення галузі науки та основні поняття.	2	2								
Тема 2. Математичний аналіз генетичного коду	2	2								
Разом за змістовим модулем 1	4	4								
Змістовий модуль 2. Біологічні бази даних. Знайомство з біологічними базами даних. Філогенія.										
Тема 3. Біологічні бази даних.	14	2	6		6					
Тема 4. Парне та множинне вирівнювання	10	2	4		4					
Тема 5. Методи молекулярної філогенії	10	2	2		6					
Тема 6. Філогенія. Побудова дерев	16	4	6		6					

Разом за змістовим модулем 2	50	10	18		22					
Змістовий модуль 3. Використання програми R для завдань біоінформатики										
Тема 7. Статистичний аналіз генетичного тексту за допомогою програми R	12	2	6		4					
Тема 8. Парне глобальне та локальне вирівнювання послідовностей в програмі R	10	2	4		4					
Тема 9. Побудова філогенетичних дерев в R	14	2	6		6					
Разом за змістовим модулем 3	36	6	16		14					
Усього годин	90	20	34		36					

* – за наявності

5. Теми семінарських занять

Не передбачено.

6. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	БД NCBI. Основні структури цієї бази даних (записи, поля, об'єкти). Формування простих та складних запитів в базі даних NCBI. Редагування запитів.	6
2	Парне та множинне вирівнювання. Он-лайн програми для вирівнювання: ClustlW, T-Coffe, Muscul, needle і т.д.	4
3	Методи молекулярної філогенії	2
4	Гомологи, дерева, еволюція. Що таке філогенетичне древо. Древо видів і древо послідовностей. Довжини гілок і їхній зміст. Типи дерев. Алгоритми реконструкції філогенетичних дерев.	6
5	Статистичний аналіз генетичного тексту за допомогою програми R	6
6	Парне глобальне та локальне вирівнювання послідовностей в програмі R	4
7	Побудова філогенетичних дерев в R	6
	Разом	34

7. Теми лабораторних занять

Не передбачено.

8. Самостійна робота

Самостійна робота студентів під час вивчення даної дисципліни включає теоретичну підготовку з окремих тем спеціального курсу, що вивчаються в порядку самостійної роботи, а також підготовку до лабораторних занять, що вказані вище.

Самостійна робота студента над засвоєнням навчального матеріалу з курсу виконується навчальних аудиторіях, комп'ютерних класах, а також в домашніх умовах.

№ з/п	Назва теми/ види завдань	Кількість годин
1	Біоінформатика: становлення галузі науки та основні поняття.	2
2	Пошук літературних джерел. Студент повинен використовуючи бази PubMed підібрати сучасні статті відповідно темі своєї дипломної роботи	4
3	Робота з базами даних. Студент самостійно знайомиться з можливостями баз даних	4
4	Пошук гомологів завданої послідовності з використанням BLAST. Студент отримує послідовність у вигляді FASTA формату у викладача та проводить пошук гомологів у базах даних, використовуючи BLAST.	4
4	Вирівнювання. Ознайомитись з можливостями проведення множинного вирівнювання з використанням інших програм. Парне глобальне та локальне вирівнювання послідовностей в програмі R	4
5	Статистичний аналіз генетичного тексту за допомогою програми R	8
6	Побудова філогенетичних дерев. Використовуючи отримані навички студент повинен підібрати нуклеотидні (або амінокислотні) послідовності відповідно темі своєї дипломної роботи та побудувати філогенетичне дерево. Побудова філогенетичних дерев в R	10
	Разом	36

9. Індивідуальне навчально-дослідне завдання

Не передбачено.

10. Методи навчання

Під час лекцій застосовується словесний та наочний методи навчання. Лекційні заняття проводяться у формі бесіди, обговорення доповідей, дискусій, що сприяє розвитку творчої самостійності студентів, поглиблює їх інтерес до науки і наукових досліджень, виховує педагогічний такт, розвиває культуру мови, вміння та навички публічного виступу, участі в дискусії. Студенти знайомляться зі статтями у спеціальних виданнях, розв'язують проблемні ситуації.

11. Методи контролю

Поточне оцінювання студентів здійснюється під час опитування студентів на лабораторних заняттях. Модульний контроль з кожного змістового модуля здійснюється за допомогою виконання завдання на комп'ютері з використанням необхідних програм, що дозволяє проконтролювати рівень засвоєння матеріалу студентом.

Комп'ютерне тестування– результат сумарного контролю з теоретичних знань. Підсумкова робота, яка включає поетапне виконання комп'ютерної обробки послідовностей є контролем практичних навиків студентів після прослуховування йми курсу.

Рівень знань студента оцінюється за 100–бальною шкалою. Наведена таблиця визначає питому вагу оцінки з кожного модуля у загальній оцінці по дисципліні

12. Питання для підсумкового контролю

1. Які основні напрямки є в біоінформатиці?
2. Які біологічні бази даних ви знаєте?
3. Що означає термін «поміхостійкість» амінокислот?
4. Дайте визначення генам ортологам.
5. Що таке BLAST ? На підставі якого принципу він заснован BLAST.
6. До якого типу баз даних відноситься база даних NCBI. Поясніть чому
7. До якого типу моделей розрахунку нуклеотидних замін у послідовності відноситься модель Джукса-Кантора
8. За допомогою якої модифікації BLAST проводять пошук гомологів по нуклеотидної послідовності в базі даних
9. Яка з двох ланцюгів ДНК містить промотор
10. При якому типі вирівнювання застосовуються матриці PAM і BLOSUM
11. Який з розділів біоінформатики займається вивченням структури і функцій генів
12. Які типи вирівнювання ви знаєте
13. Дайте визначення терміну «слабка основа» у амінокислоти
14. Які програми використовуються для здійснення множинного вирівнювання
15. У якому типі запитів використовують булеві оператори
16. У чому сенс вирівнювання.
17. Які моделі розрахунку нуклеотидних замін в послідовності ви знаєте
18. Які є основні підходи для конструювання філогенетичних дерев ви знаєте
19. В якому напрямку генетичний код однозначний
20. Сформулюйте принципове положення мутаційної теорії
21. За допомогою якої програми виконується локальне вирівнювання двох послідовностей. З якою метою проводять вирівнювання такого типу

13. Розподіл балів, які отримують студенти

Поточне тестування та самостійна робота				Сума
Змістовий модуль 1	Змістовий модуль 2	Змістовий модуль 3		
30	30	40		100

Критерії та шкала оцінювання: національна та ECTS

Реалізація основних завдань контролю знань здобувачів вищої освіти в ОНУ досягається системними підходами до оцінювання та комплексністю застосування різних видів контролю. Згідно з діючою в університеті системою комплексної діагностики знань здобувачів вищої освіти, з метою стимулювання планомірної та систематичної навчальної роботи, оцінка знань здійснюється за 100-бальною системою, яка переводиться відповідно у національну шкалу («відмінно», «добре», «задовільно», «незадовільно») та шкалу європейської кредитно-трансферної системи (СКТС –A, B, C, D, E, FX, F).

За системою ОНУ	За шкалою ECTS	За національною системою	Визначення
90-100	A	5 (відмінно)	Повно та ґрунтовно засвоїв всі теми навчальної програми вміє вільно та самостійно викласти зміст всіх питань програми навчальної дисципліни, розуміє її значення для своєї професійної підготовки, повністю виконав усі завдання кожної теми та поточного модульного контролю в цілому.

85-89	B	4 (дуже добре)	Недостатньо повно та ґрунтовно засвоїв окремі питання робочої програми. Вміє самостійно викласти зміст основних питань програми навчальної дисципліни, виконав завдання кожної теми та модульного поточного контролю в цілому.
75-84	C	4 (добре)	Недостатньо повно та ґрунтовно засвоїв деякі теми робочої програми, не вміє самостійно викласти зміст деяких питань програми навчальної дисципліни. Окремі завдання кожної теми та модульного поточного контролю в цілому виконав не повністю.
70-74	D	3 (задовільно)	Засвоїв лише окремі теми робочої програми. Не вміє вільно самостійно викласти зміст основних питань навчальної дисципліни, окремі завдання кожної теми модульного контролю не виконав.
60-69	E	3 (достатньо)	Засвоїв лише окремі питання навчальної програми. Не вміє достатньо самостійно викласти зміст більшості питань програми навчальної дисципліни. Виконав лише окремі завдання кожної теми та модульного контролю в цілому.
35-59	FX	2 (незадовільно)	Не засвоїв більшості тем навчальної програми не вміє викласти зміст більшості основних питань навчальної дисципліни. Не виконав більшості завдань кожної теми та модульного контролю в цілому.
0-34	F	2 (незадовільно)	Не засвоїв навчальної програми, не вміє викласти зміст жодної теми навчальної дисципліни, не виконав модульного контролю, з обов'язковим повторним вивченням дисципліни.

14. Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	A	відмінно	зараховано
85-89	B	добре	
75-84	C		
70-74	D	задовільно	
60-69	E		
35-59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
0-34	F	незадовільно з обов’язковим повторним вивченням дисципліни	не зараховано з обов’язковим повторним вивченням дисципліни

14. Методичне та матеріально-технічне забезпечення

Методичне забезпечення: робоча програма навчальної дисципліни; навчальна програма дисципліни; мультимедійний супровід матеріалів лекцій та семінарських занять; план семінарських занять.

Матеріальне забезпечення: мультимедійний проектор Epson EB-X31 (введений в експлуатацію 10/2019 року), екран Protecta Matte White 180 (введений в експлуатацію 04/2002 року); комп'ютерний клас, мережа інтернет, free software

15. Рекомендована література

1. Васильева Н. Ю. Биоинформатика (Множественное выравнивание. Филогенетические деревья) Методическое пособие. — Одесса:Одесский национальный университет имени И. И. Мечникова, 2014. — 70 с
2. Леск, А. Введение в биоинформатику / А.Леск. —Москва: Бином. Лаборатория знаний, 2009. —318 с.
3. Козлов, Н.Н. Математический анализ генетического кода / Н.Н. Козлов. —Москва: Бином. Лаборатория знаний, 2012. —215 с.
4. Andrzej Polanski. Marek Kimmel. Bioinformatics. Springer Berlin Heidelberg New York. — 2007. — 389 p.P.M. Selzer, R.J. Marhуfer, A. Rohwer. Applied Bioinformatics. An Introduction. Springer-Verlag Berlin Heidelberg. — 2008. — 297 p.
5. Mount D.W. Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press. — 2001. —317 p.
6. Mount, David W. Bioinformatics: sequence and genome analysis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. — 2001. — 577 p.
7. Бернхард Хаубольд, Томас Вие. Введение в вычислительную биологию. эволюционный поход. — М. Ижевск:НИЦ «Регулярная и хаотическая динамика», Ижевский институт компьютерных исследований. — 2011. — 456 с.
8. Вейр Б. Анализ генетических данных: Пер. с англ. — М.: Мир. — 1995. — 400 с.
9. Лукашев В. В. Молекулярная еволюция и филогенетический анализ. Учебное пособие. — М.: БИНОМ. Лаборатория знаний. — 2009. — 256 с.
10. Основы биоинформатики : учеб. пособие / А. Н. Огурцов. — Х. : НТУ «ХПИ». — 2013. — 400 с.
11. Павлинов И. Я.. Введение в современную филогенетику (кладогенетический аспект). М.: изд-во КМК. — 2005.— 192 с.

17. Електронні інформаційні ресурси

База даних з літератури PubMed - <http://www.ncbi.nlm.nih.gov/pubmed>

База даних NCBI - <http://www.ncbi.nlm.nih.gov/>

База даних ENBL - <http://www.ebi.ac.uk/>

Програми для роботи з послідовностями ресурсу ENBL

<http://www.ebi.ac.uk/services>

Програма для здійснювання множинного вирівнювання

<http://www.ebi.ac.uk/Tools/msa/muscle/>

Програма для здійснювання множинного вирівнювання - <http://tcoffee.crg.cat/>

База даних UniProt - <http://www.uniprot.org/>

Пакет програм «Robust Phylogenetic Analysis For The Non-Specialist» -

<http://www.phylogeny.fr/>

Біоінформаційні ресурси для геноміки і протеоміки - www.expasy.org

Програми аналізу полінуклеотидних і поліпептидних послідовностей
-<http://blast.ncbi.nlm.nih.gov>

Програми множинного вирівнювання - www.genome.jp/tools/clustalw/

Форум з молекулярної біології - <http://molecularstation.com/>