

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Львівський національний університет імені Івана Франка  
Біологічний факультет  
Кафедра генетики та біотехнології

Затверджено  
на засіданні кафедри генетики та біотехнології  
(протокол № 17 від 29 серпня 2025 р.)

Завідувач кафедри.

  
Віктор ФЕДОРЕНКО

Силабус з навчальної дисципліни «Біоінформатика»  
що викладається в межах ОПП «Біотехнології та біоінженерія» першого  
(бакалаврського) рівня вищої освіти для здобувачів зі спеціальності 162  
«Біотехнології та біоінженерія»

Львів 2025

<b>Назва курсу</b>	Біоінформатика
<b>Адреса викладання курсу</b>	вул. Грушевського 4, 79005 Львів
<b>Факультет та кафедра, за якою закріплена дисципліна</b>	біологічний факультет, кафедра генетики і біотехнології
<b>Галузь знань, шифр та назва спеціальності</b>	Галузь знань – 16 Хімічна інженерія та біоінженерія, спеціальність – 162 Біотехнології та біоінженерія
<b>Викладачі курсу</b>	Богдан Омелянович Осташ (лекції) Степан Іванович Тістечок (практичні заняття)
<b>Контактна інформація викладачів</b>	<a href="mailto:bohdan.ostash@lnu.edu.ua">bohdan.ostash@lnu.edu.ua</a> 032 2394407 <a href="mailto:stepan.tistechok@lnu.edu.ua">stepan.tistechok@lnu.edu.ua</a> <a href="http://bioweb.lnu.edu.ua/employee/ostash-b-o">http://bioweb.lnu.edu.ua/employee/ostash-b-o</a>
<b>Консультації по курсу відбуваються</b>	<i>Очні консультації:</i> V семестр (2025 р), щовівторка, 11:30-13:00 <i>Онлайн-консультації:</i> у форматі “питання-відповідь” через електронну пошту, в робочі дні тижня, з 10:00-16:00; очікуйте на відповідь не пізніше ніж за три доби з моменту надходження питання
<b>Сторінка курсу</b>	<a href="https://bioweb.lnu.edu.ua/course/bioinformatyka">https://bioweb.lnu.edu.ua/course/bioinformatyka</a>
<b>Інформація про курс</b>	Курс розроблений так, щоб бакалаври набули фахові знання з аналізу нуклеотидних та амінокислотних послідовностей (НАП) за допомогою спеціалізованого програмного забезпечення. У курсі розглядаються біоінформатичні бази даних, основні математичні моделі НАП, методи попарного і множинного вирівнювань, моделей на їхній основі, зокрема великі мовні моделі, і як використовують у біотехнології. Курс включає теоретичний матеріал у вигляді лекцій і виконання лабораторних занять.
<b>Коротка анотація курсу</b>	Дисципліна «Біоінформатика» є нормативною дисципліною зі спеціальності 162 «Біотехнології та біоінженерія» для ОП бакалавра, яка викладається в 5 семестрі в обсязі 3 кредитів (за ECTS). У 2025 року у базах даних є інформація про гени більше 500 000 організмів, або більш ніж $21 \times 10^{12}$ пар нуклеотидів. Це колосальна кількість інформації про генетичну основу життя, і її кількість неупинно зростає. Систематичний аналіз усього масиву даних експериментальними методами нереалістичний як економічно так і технологічно. Це стимулювало розвиток методів комп'ютерного аналізу НАП, що допомогли упорядкувати дані у межах різних веб-сервісів. Найважливішим результатом синтезу молекулярної біології, геноміки, інформатики, комп'ютерних технологій і статистики – біоінформатики – стало те, що ця нова галузь перетворилася у самостійне знаряддя наукового відкриття. Курс “Біоінформатика” включає розділи, присвячені структурам баз даних; попарному і множинному вирівнюванню НАП; моделям, що узагальнюють множинні вирівнювання; основам філогенетичної реконструкції на основі; передбаченню функцій генів і пошукові мотивів у НАП; передбаченню тривимірних структур білків; аналізу молекул РНК.
<b>Мета та завдання курсу</b>	<b>Мета:</b> сформувати у слухачів курсу систему знань про основні бази даних нуклеотидних й амінокислотних послідовностей (НАП) та їхнього комп'ютерного аналізу, зокрема попарного і множинного

	<p>вирівнювання, моделей на основі вирівнювань, філогенетичної реконструкції на основі НАП.</p> <p><b>Завдання:</b> а) викласти концептуальні математичні та молекулярно-біологічні засади, на яких ґрунтується біоінформатика; б) ознайомити студентів з наявним арсеналом методів біоінформатики, їхніх можливостей і обмежень, останніх тенденцій розвитку дисципліни; в) сформулювати у студентів арсенал активних знань у галузі цієї дисципліни, тобто навчити їх формулювати наукові питання, відповіді на які можна шукати із залученням методів біоінформатики.</p>
<b>Література для вивчення дисципліни</b>	<p><b>Основна література:</b></p> <p><b>Базова:</b></p> <ol style="list-style-type: none"> <li>1. Осташ Б.О. Біоінформатика: аналіз генетичних послідовностей. Електронний підручник. Видавничий центр ЛНУ ім. Івана Франка, 2022, 232 стор. ISBN 978-617-10-0729-1. Доступ онлайн: <a href="http://dspace.lnlibrary.lviv.ua/handle/123456789/169">http://dspace.lnlibrary.lviv.ua/handle/123456789/169</a></li> <li>2. Higgs PG, Attwood TK. Bioinformatics and Molecular Evolution. Blackwell Publishing, Oxford, 2005. 398 p. ISBN 1-4051-0683-2.</li> </ol> <p><b>Допоміжна:</b></p> <ol style="list-style-type: none"> <li>3. Durbin R, Eddy S, Krogh A, Mitchison G. Biological Sequence Analysis. Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, Cambridge, 1998. 371 p. ISBN-13 978-0-521-62971-3</li> <li>4. Borodovsky M, Ekisheva S. Problems and Solutions in Biological Sequence Analysis. Cambridge University Press, Cambridge, 2006. 362 p. ISBN-13 978-0-521-61230-2</li> <li>5. Bioinformatics: a practical guide to the analysis of genes and proteins, 2<sup>nd</sup> Ed / AD Baxevanis, BFF Ouellette. – New York: John Wiley &amp; Sons, 2001. – 455 p.</li> <li>6. Allman ES, Rhodes JA. Mathematical Models in Biology. An Introduction. Cambridge University Press, Cambridge, 2003. 386 p.</li> <li>7. Pevsner J. Bioinformatics and functional genomics. 3<sup>rd</sup> edition. Wiley Blackwell, UK. – 2015- 1116 p. ISBN 978-1-118-58178-0.</li> <li>8. <a href="https://www.hsls.pitt.edu/obrc/">https://www.hsls.pitt.edu/obrc/</a></li> </ol>
<b>Тривалість курсу</b>	Один семестр
<b>Обсяг курсу</b>	Лекції – 32 год, практичні – 32 год, самостійна робота – 26 год; разом – 90 год (3 кредитів ECTS)
<b>Очікувані результати навчання</b>	<p>Після курсу бакалавр буде знати:</p> <ul style="list-style-type: none"> <li>- основні бази даних НАП, принципи функціонування алгоритмів вирівнювання НАП, і спектр питань щодо структури, функції та еволюції НАП, на які можна шукати відповіді із застосуванням біоінформатичних веб-ресурсів і програм;</li> <li>- методи філогенетичної реконструкції на основі НАП;</li> </ul> <p><b>вміти:</b></p> <ul style="list-style-type: none"> <li>- користуватись базами даних, де зберігаються нуклеотидні та амінокислотні послідовності, інформація про структури геномів, їхню експресію (транскриптомні дані – генні чіпи і RNAseq) та відповідні протеоми;</li> <li>- порівнювати НАП з гомологічними послідовностями за допомогою методів попарного і множинного вирівнювань; оперувати</li> </ul>

	<p>паттернами і профілями;</p> <ul style="list-style-type: none"> <li>- виявляти регуляторні послідовності у геномах – повтори, паліндроми, консервативні мотиви; передбачати функціональність гена (білка) на основі аналізу консервативних доменів і каталітичних центрів;</li> <li>- користуватись методами передбачення і моделювання структури білків; здійснювати молекулярно-філогенетичний аналіз.</li> </ul>
<p><b>ЗК, ФК, ПРН</b></p>	<p>Курс розроблено так, щоб сформувати у студентів <u>загальні і фахові компетентності</u>:</p> <p>ЗК01. Здатність застосовувати знання у практичних ситуаціях.</p> <p>ЗК04. Навички використання інформаційних та комунікаційних технологій.</p> <p>ЗК05. Здатність вчитися і оволодівати сучасними знаннями.</p> <p>ФК1. Здатність використовувати знання з математики та фізики в обсязі, необхідному для досягнення інших результатів освітньої програми.</p> <p>ФК2. Здатність використовувати ґрунтовні знання з хімії та біології в обсязі, необхідному для досягнення інших результатів освітньої програми.</p> <p>ФК14. Здатність використовувати сучасні автоматизовані системи управління виробництвом біотехнологічних продуктів різного призначення, їх технічне, алгоритмічне, інформаційне і програмне забезпечення для вирішення професійних завдань.</p> <p>ФК16. Здатність використовувати бази даних про геноми живих організмів, їхні протеоми, транскриптоми тощо при плануванні, проведенні та оптимізації біотехнологічних досліджень, використовувати методи біоінформатики для розробки біотехнологій.</p> <p>ФК17. Здатність планувати і проводити досліди з конструювання, вивчення, селекції та зберігання штамів мікроорганізмів – промислових продуцентів комерційно важливих біотехнологічних продуктів, у тому числі трансгенних організмів за допомогою методів клітинної і генетичної інженерії, аналізувати їхні результати, а також опрацьовувати способи їх ефективного використання в межах біотехнологічних виробництв.</p> <p><u>Програмні результати навчання:</u></p> <p>ПР1. Вміти застосовувати сучасні математичні методи для розв’язання практичних задач, пов’язаних з дослідженням і проектуванням біотехнологічних процесів. Використовувати знання фізики для аналізу біотехнологічних процесів.</p> <p>ПР24. Вміти користуватися базами даних, в яких зберігається інформація про геноми живих організмів, їхні протеоми, транскриптоми тощо при плануванні, проведенні та оптимізації біотехнологічних досліджень, використовувати методи біоінформатики для розробки біотехнологій.</p>

	<p>PR25. Вміти планувати і проводити дослідження з конструювання, вивчення, селекції та зберігання штамів мікроорганізмів – промислових продуцентів комерційно важливих біотехнологічних продуктів, у тому числі трансгенних організмів за допомогою методів клітинної і генетичної інженерії, аналізувати їхні результати, а також опрацьовувати способи їх ефективного використання в межах біотехнологічних виробництв.</p>
<b>Ключові слова</b>	Математичні моделі в біології, множинні і попарні вирівнювання, моделі Маркова, філогенетичні реконструкції, мовні моделі (AlphaFold, ESM-3), бази НАП, тривимірних структур білків.
<b>Формат курсу</b>	Очний або дистанційний
<b>Теми</b>	Наведено у таблиці
<b>Підсумковий контроль, форма</b>	Іспит в кінці семестру. Усний.
<b>Пререквізити</b>	Знання англійської мови на рівні, достатньому для перекладу наукових статей; необхідні знання з основ генетики, біохімії, зоології та ботаніки. Розуміння базових математичних понять (логарифм, частка значень, експонента, відсоткові величини) та теорії імовірностей та статистичного аналізу даних. Базові навички роботи з комп'ютером. Наявність комп'ютера/смартфона з підключенням до інтернету (для лабораторних занять у випадку дистанційного формату навчання).
<b>Навчальні методи та техніки, які будуть використовуватися під час викладання курсу</b>	Лекції, пояснення, виконання завдань на практичних заняттях.
<b>Необхідне обладнання</b>	Персональний комп'ютер, загальноживані комп'ютерні програми і операційні системи, підключення до мережі Інтернет
<b>Критерії оцінювання (окремо для кожного виду навчальної діяльності)</b>	<p>Виконання завдань упродовж семестру/іспит – 50/50</p> <p>За змістом перших дев'яти лекцій буде виконано поточний контроль знань у вигляді написання модуля. В модуль входять: визначення термінів (10 балів), два питання (по 6 б. кожне), одна схема чи таблиця, яку треба заповнити/зобразити (8 б.). Максимальна оцінка за модуль – 20 балів. Написання модуля обов'язкове.</p> <p>Ще 30 балів студент може набрати упродовж семестру за виконання шести контрольних завдань упродовж практичних занять (Бази даних, BLAST, MSA, WebLogo, Phylo1, Phylo2).</p> <p>Сумарний бал за модуль і за виконання контрольних завдань під час практичних занять має становити не менше 25, у такому співвідношенні: не менше 5 за модуль і не менше 20 за практичні.</p> <p>На іспиті (50 б.) – набір питань аналогічно до модуля; до термінів (10 б.), питань і схем (3 по 10 б.) додаються тести (2 по 5 б.). Письмова підготовка на протязі не більше 30 хв, далі усна відповідь. На іспит виноситься весь матеріал курсу</p> <p>Жодні форми порушення академічної доброчесності (<a href="https://www.lnu.edu.ua/wp-content/uploads/2019/06/reg_academic_virtue.pdf">https://www.lnu.edu.ua/wp-content/uploads/2019/06/reg_academic_virtue.pdf</a>) не толеруються.</p>
<b>Питання до екзамену</b>	1. Моделі, що описують частоту зустрічності “слів” у генетичних послідовностях – Бернуллі і ланцюги Маркова.

2. Консервативні і неконсервативні заміщення у попарних вирівнюваннях, способи їхнього оцінювання; підходи до уведення у вирівнювання прогалин
3. PubMed – стратегії пошуку бібліографічних даних, значення ідентифікаторів статей
4. Структура флет-файлу GenBank
5. Національний центр біотехнологічної інформації США (NCBI) – структура і функції
6. Теорія прийнятних точкових мутацій (PAM) М. Дейгоф
7. Опишіть процес побудови позиційно-специфічних рахункових матриць (PSSM).
8. Алгоритми динамічного програмування у попарному вирівнюванні.
9. Основні терміни і теоретичні засади попарного вирівнювання
10. Опишіть етапи прогресивного принципу множинного вирівнювання
11. Моделі еволюції нуклеотидних послідовностей
12. Моделі Маркова в аналізі генетичних послідовностей
13. Способи опису складності НАП та районів з низькою складністю. Біологічне значення НАП низької складності.
14. Моделі заміщення в кодонних послідовностях
15. Опишіть основні елементи сторінки результатів BLAST
16. Підходи до уведення прогалин у попарні й множинні вирівнювання
17. Принципи побудови і використання прихованих моделей Маркова в аналізі НАП
18. Статистична оцінка попарних вирівнювань. Теорема Карліна-Альтшуля
19. Як розуміти унітарні й емпіричні системи обчислення рахунків вирівнювань з точки зору теорії інформації? Наведіть приклади. Теорема Альтшуля
20. У рівнянні з теореми Карліна-Альтшуля число очікування  $E$  прямо пропорційне простору пошуку (розміру бази даних) і обернено рахунку вирівнювання  $S$ . Тобто, у більших базах імовірність виявити випадкові збіги мала б зростати. Однак на практиці більші бази даних приводять до вирівнювань з дуже низьким  $E$ . Чому?
21. Які підходи використовують до оцінювання розривів (прогалин) у попарних вирівнюваннях?
22. Основні параметри опису попарного вирівнювання:  $S$ ,  $S'$ ,  $E$ ,  $P$ , ідентичність  $ID$ , подібність  $SI$
23. НАП з точки зору інформації: інформаційний вміст, символічне повідомлення, частотний розподіл символів, імовірність зустрічності символів/слів у послідовності.
24. Параметри налаштування програми BLAST
25. PAM-матриці і спосіб їхньої побудови
26. Опишіть принципи побудови механістичних та емпіричних матриць заміщень; що відмінного і спільного між ними.
27. Принцип побудови матриць BLOSUM, їхні відмітні властивості (порівняно з PAM).
28. Відмінні і спільні риси паттернів, позиційно-специфічних матриць і прихованих моделей Маркова
29. Чому не існує не одна, а серія матриць PAM або BLOSUM? Як

	<p>обчислюють матрицю PAM250? Для яких потреб слід використовувати матрицю PAM20?</p> <p>30. Основні властивості моделей еволюції нуклеотидних послідовностей.</p> <p>31. Який біологічний зміст несе уведення розривів (прогалин) у попарні вирівнювання? Чи можна переставляти місцями позиції вирівнювання, і якщо так – то який біологічний процес відображає така маніпуляція НАП</p> <p>32. Принципи дотплот-аналізу НАП</p> <p>33. Що спонукало дослідників до розробки емпіричних підходів до оцінки вирівнювань амінокислотних послідовностей? Чому емпіричні підходи не набули поширення для нуклеотидних послідовностей?</p> <p>34. Спільне і відмінне в попарному і множинному вирівнюваннях</p> <p>35. Стисло опишіть моделі, які створюють на основі множинного вирівнювання НАП. Що в них відмінного?</p> <p>36. Еволюційні засади попарного вирівнювання. Основні терміни.</p> <p>37. Поясніть, що таке позиційно-незалежні та позиційно-специфічні рахунки вирівнювання, наведіть приклади їхнього використання в аналізі НАП</p> <p>38. Концепція псевдорахунків у біоінформатиці – приклади її використання</p> <p>39. Опишіть відомі вам способи множинного вирівнювання НАП</p> <p>40. Принцип функціонування алгоритму пошуку оптимального глобального вирівнювання (Нідельмана-Ванча)</p> <p>41. Відмінність між алгоритмами глобального і локального вирівнювання НАП</p> <p>42. За якими ознаками певну модель чи процес можна віднести до Марковського?</p> <p>43. Яку роль відіграють константи <math>\lambda</math> й <math>K</math> в рівнянні для обчислення числа <math>E</math>?</p> <p>44. Етапи побудови позиційно-специфічної рахункової матриці</p> <p>45. Моделі оцінки частот символів/слів у НАП</p> <p>46. Вичерпні та евристичні підходи до попарного вирівнювання НАП.</p> <p>47. Порівняйте позиційно-специфічні матриці й приховані моделі Маркова як методи опису множинних вирівнювань.</p> <p>48. ДНК- і білкові логотипи – принцип побудови, та інформація, яку він містить</p> <p>49. У чому полягає складність побудови множинних вирівнювань? Які є способи оцінки їхньої якості?</p> <p>50. Принцип функціонування алгоритму BLAST.</p> <p>51. Опишіть, як зміна налаштувань програми BLAST впливає на результат пошуку гомологів?</p> <p>52. Прогресивний та ітеративний принципи множинного вирівнювання</p> <p>53. Які фактори ускладнюють безпосереднє використання даних про частоти появи мутацій у ДНК для коректного опису її еволюції?</p> <p>54. Основні етапи побудови і використання позиційно-специфічних вагових матриць</p> <p>55. Можливості й обмеження методів попарного і множинного</p>
--	---

	<p>вирівнювання для виявлення гомологічних НАП</p> <p>56. Які дані містить сторінка результатів BLASTP?</p> <p>57. Основні етапи пошуку оптимального локального вирівнювання за алгоритмом Сміта-Уотермана</p> <p>58. Еволюційні засади філогенії</p> <p>59. Основні поняття у галузі філогенетичної реконструкції</p> <p>60. Яку інформацію містить філогенетичне дерево, а яку – не містить?</p> <p>61. Основні елементи філогенетичного дерева, взаємозв'язок між ними</p> <p>62. Вибір даних для філогенетичної реконструкції</p> <p>63. Що таке модель еволюції у філогенетичній реконструкції і яке її значення?</p> <p>64. Опишіть основні етапи філогенетичної реконструкції</p> <p>65. Методи статистичної оцінки філогенетичних дерев</p> <p>66. Підхід до філогенетичної реконструкції: з'єднання сусідів (NJ)</p> <p>67. Підхід до філогенетичної реконструкції: максимальної ощадності (MP)</p> <p>68. Підхід до філогенетичної реконструкції: максимальної вірогідності (ML)</p> <p>69. Чому дерево-провідник з прогресивних підходів до множинного вирівнювання не є філогенетичним деревом?</p> <p>70. У чому полягає суть курування вихідного множинного вирівнювання, що передує філогенетичній реконструкції?</p> <p>71. На чому ґрунтується суть пошуку мотивів ДНК?</p> <p>72. Принцип функціонування алгоритму максимізації очікування (EM)?</p> <p>73. Принцип функціонування програми MEME</p> <p>74. Поясніть поняття родина, фолд, клас білка (за системою SCOP)</p> <p>75. Поясніть принципи пошуку гомологів за допомогою програм BLAST й HHPred</p> <p>76. Чому порівняння первинних і третинних структур генетичних послідовностей дає різний результат?</p> <p>77. Практичне використання філогенетичного аналізу</p> <p>78. В чому полягає принцип зважування генетичних послідовностей у моделях множинних вирівнювань?</p> <p>79. Моделі заміщення в нуклеотидних і кодонних послідовностях, їхні особливості порівняно з моделями амінокислотних заміщень</p> <p>80. AlphaFold: принцип роботи, застосування у дослідженнях</p> <p>81. Програма Mol*Star</p>
<b>Опитування</b>	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу ( <a href="https://education-quality.lnu.edu.ua/">https://education-quality.lnu.edu.ua/</a> )

**Таблиця**

**Схема курсу «Біоінформатика». Форма навчання - денна**

Тиж-день	Тема занять (перелік питань)	Форма діяльності та обсяг годин	Додаткова література / ресурс для виконання завдань (за потреби)	Термін виконання
1	Вступ. Засади біоінформатики. Бази даних.	Лекції – 2 год. Практичні – 2 год. Самостійна робота – 3 год	<a href="http://rosalind.info/problems/list-view/">http://rosalind.info/problems/list-view/</a> .	1 тиждень

2,3	Математичні моделі	Лекції – 4 год, Практичні – 4 год. Самостійна робота – 3	<a href="https://www.bv-brc.org/">https://www.bv-brc.org/</a>	2 тижні
4,5,6	Попарне вирівнювання НАП	Лекції – 6 год, Практичні – 6 год, Самостійна робота – 3		3 тижні
7	Множинне вирівнювання НАП	Лекції – 2 год, Практичні – 2 год. Самостійна робота – 2		1 тиждень
8, 9	Моделі на основі множ. вирівнювань	Лекції – 4 год, Практичні – 4 год. Самостійна робота – 3	<a href="https://toolkit.tuebingen.mpg.de/tools/hhpred">https://toolkit.tuebingen.mpg.de/tools/hhpred</a>	2 тижні
10, 11, 12, 13	Філогенетична реконструкція	Лекції – 8 год, Практичні – 8 год. Самостійна робота – 5	<a href="https://phylogeny.fr">https://phylogeny.fr</a> <a href="https://evolution.berkeley.edu/evolibrary/search/topics.php?topic_id=15">https://evolution.berkeley.edu/evolibrary/search/topics.php?topic_id=15.</a>	4 тижні
14	Пошук паттернів та операторів	Лекції – 2 год, Практичні – 2 год. Самостійна робота – 3	<a href="https://meme-suite.org/meme/">https://meme-suite.org/meme/</a>	1 тиждень
15	Бази й аналіз білкових структур	Лекції – 2 год, Практичні – 2 год. Самостійна робота – 2	<a href="https://www.rcsb.org/3d-view">https://www.rcsb.org/3d-view</a>	1 тиждень
16	Бази й аналіз структур РНК	Лекції – 2 год, Практичні – 2 год. Самостійна робота – 2		1 тиждень

Автор



Богдан ОСТАШ

"Погоджено"

Голова методичної ради  
біологічного факультету  
Віталій ГОНЧАРЕНКО



"*Віталій*" *Віталій* 2025 р.

Гарант ОНЦ «Біотехнології та біоінженерія»  
Віктор ФЕДОРЕНКО



*29 08.* 2025 р.