

**ВІДГУК**  
**офіційного опонента**  
на дисертаційну роботу  
Рокицького Ігоря Володимировича  
**"БІОІНФОРМАТИЧНІ ПІДХОДИ ТА РЕПОРТЕРНА СИСТЕМА ДЛЯ**  
**ДОСЛІДЖЕННЯ ОСОБЛИВОСТЕЙ ВЖИВАННЯ КОДОНІВ У**  
**ГЕНОМАХ СТРЕПТОМІЦЕТІВ",**

представлену до захисту на здобуття наукового ступеня кандидата біологічних наук за спеціальністю 03.00.22 – молекулярна генетика

Кодон ТТА виконує важливу біологічну функцію, оскільки саме він впізнається лейциновою тРНК, а блокування роботи її гену *bldA* викликає порушення морфогенезу і вторинного метаболізму бактерій. Таким чином можна говорити про певний регуляторний механізм, що залежить від нормальної експресії гену *bldA* і наявності канонічного сайту - рідкісного кодону ТТА. Усе це передбачає високу точність процесу трансляції вищезгаданого кодону. Водночас, результати ряду досліджень вказують на те, що рідкісні кодони значно частіше схильні до помилкової трансляції, ніж популярні. Таким чином існує певний конфлікт між тим, що, з одного боку, кодон ТТА має бути рідкісним, щоб контролювати лише певні гени і, відповідно, транслюватися лише однією тРНК, і, з іншого боку, витримувати тиск притаманної рідкісним кодонам «ентропії» помилкової трансляції.

Водночас слід зазначити, що припущення про неточність рідкісних кодонів була визначена більше ніж 25 років тому назад на прикладі певного кола кодонів з модельних евкаріотичних організмів і бактерій нестрептоміцетного походження. Так, на сьогодні існують правдоподібні моделі трансляції, що дозволяють здійснювати досить точне передбачення декодування різних кодонів на основі послідовностей геномів. Водночас у випадку представників роду *Streptomyces* подібні дослідження практично відсутні.

Головною метою дисертаційної роботи Рокицького Ігоря Володимировича була розробка репортерної системи і web-сервісу, спрямованих на вивчення трансляційної регуляції генів вторинного метаболізму представників роду *Streptomyces*. Для вирішення поставленої задачі Автору, перш за все, було потрібно сформувати референтні вибірки ортологічних і функціонально споріднених генів *Streptomyces*, встановити для них оптимальні механістичні моделі нуклеотидних та амінокислотних заміщень. По друге, було необхідно визначити особливості контекстного вживання кодонів на основі аналізу геномів стрептоміцетів і дослідити особливості кодонних заміщень на прикладі функціонально різних класів генів.

Також перед автором стояло завдання встановити можливі варіанти трансляції та мітрансляції рідкісного лейцинового кодона ТТА у представників роду *Streptomyces*. Це, у

свою чергу, ґрунтувалось на застосуванні ансамблю різних біоінформатичних методів і лабораторної експериментальної оцінки концентрації тРНК в клітинах стрептоміцетів.

В результаті все це дозволило дисертанту розробити власну репортерну систему для дослідження особливостей вживання кодонів, а також, для дослідження рівня коректної трансляції рідкісного лейцинового кодона ТТА на прикладі представників роду *Streptomyces*.

**Зв'язок з науковими програмами, планами, темами.** Робота виконана у межах бюджетної теми БГ-41Нр “Універсальний генетичний механізм контролю продукції біологічно активних речовин стрептоміцетами” (№ державної реєстрації 0116U008070, 2016-2018 рр.), що виконувалась у науково-дослідній Лабораторії генетики, селекції та генетичної інженерії продуцентів антибіотиків (НДЛ-42) на кафедрі генетики та біотехнології Львівського національного університету імені Івана Франка.

**Наукова новизна дослідження та одержаних результатів, ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації.**

Дисертаційна робота Рокицького І.В. була виконана на сучасному науковому та методичному рівні і ґрунтується на значному матеріалі оригінальних досліджень і авторських розробок. Усі експерименти були виконані автором особисто або за його безпосередньою участю. Під час виконання дисертаційних досліджень автором було визначено оптимальні моделі заміщення нуклеотидних і амінокислотних залишків послідовностей, що походять з геномів і протеомів представників роду *Streptomyces*. Було доведено впорядкованість асоціації низки кодонів в досліджених геномах. На підставі отриманих результатів Автором було розроблено мережевий web-сервіс для візуалізації моделей кодонних заміщень. Застосовано на практиці оригінальний підхід для біоінформатичного прогнозування рівня помилкових трансляцій кодонів, а також запропоновано нову експериментальну модель для вивчення експресії рідкісного кодона ТТА представників роду *Streptomyces*.

Під час дослідження Автором було створено локальні репозиторії ортологічних послідовностей, що охоплюють функціонально відмінні класи генів та білків, залучених до вторинного метаболізму бактерій роду *Streptomyces*. Загалом, відібрана група послідовностей охоплювала ензими, транскрипційні фактори, а також білки мембранного комплексу.

Автором було відібрано оптимальні моделі, що описують заміщення, характерні для різних смислових послідовностей і відповідних продуктів їх трансляції. Показано, що групам послідовностей, об'єднаних гомологією функції, також притаманна подібність і за моделлю заміщень. Встановлено, що у випадку нуклеотидних послідовностей стрептоміцетів безсумнівно домінують дві моделі замін нуклеотидів: K3Pu (K81u, Three substitution types model and equal base freq.) та TVM (Transversion model, AG=CT and unequal base freq.).

Встановлено, що для кожної з досліджених амінокислотних послідовностей також була притаманна своя оптимальна модель.

Встановлено, що кодуючим ділянкам послідовностей геномів представників *Streptomyces* притаманні дев'ять позитивних дикодонних асоціацій: UAU-CUG, CUG-CGC, GUA-CGG, GAU-CCG, CUC-ACC, CUC-GCC, CUC-GGC, GAA-CUC, GAA-CUG. Статистична обробка результатів підтвердила, що вищезгадані дикодонні асоціації зустрічаються з частотою, вищою за випадкову. Також Здобувачем було виявлено дві негативні асоціації: CUC-CUG, CUC-GAG. Встановлено, що спільною рисою усіх досліджених асоціацій є те, що вони притаманні переважно лейциновим кодонам, які містять аденін або тимін у центральній позиції кодона.

На підставі результатів оригінальних досліджень було створено мережевий сервіс для визначення і візуалізації кодонних заміщень у масивах кодувальних послідовностей. Відповідно, автором розроблено рекомендації стосовно коректного аналізу результатів на підставі результуючих графіків. Із використанням розробленого інструменту було виявлено характерну для стрептоміцетів тенденцію до спрямованого збагачення генів GC-багатими кодонами.

На прикладі гену  $\beta$ -галактозидази *sco3479* із *Streptomyces albus* штаму J1074 було створено та апробовано репортерну систему, що дає змогу безпосереднього дослідження впливу різних мутацій чи факторів середовища на експресію рідкісного кодона ТТА.

Детальне вивчення матеріалів дисертації, автореферату, а також відповідних публікацій здобувача дає мені змогу підтвердити наукову новизну, оригінальність одержаних результатів, а також високій ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертаційному дослідженні.

#### **Практичне значення одержаних результатів.**

Перш за все, слід зауважити, що в результаті виконання дисертаційної роботи автором було розроблено мережевий web-сервіс для візуалізації моделей кодонних заміщень. Розроблений web-сервіс було застосовано на практиці для прогнозування рівня помилкових трансляцій кодонів. За результатами дисертаційного дослідження Автором було запропоновано нову експериментальну модель, яка, без сумніву, знайде подальше застосування в напрямку дослідження експресії рідкісного кодона ТТА у представників роду *Streptomyces*.

Також цінність отриманих результатів полягає у можливості їхнього використання для дослідження генетичних механізмів контролю продукції біологічно активних речовин стрептоміцетами. Отримані результати, створені штами *E. coli*, стрептоміцетів і рекомбінантні молекули ДНК вже знайшли використання у навчальному процесі кафедри

генетики та біотехнології Львівського національного університету імені Івана Франка. Залучення до навчального процесу було здійснено в межах положення про структурний підрозділ “Колекція культур мікроорганізмів - продуцентів антибіотиків”, затвердженого рішенням Вченої ради Львівського національного університету імені Івана Франка (протокол №15/2 від 24.02.2016 р.).

#### **Особистий внесок здобувача**

Викладені в дисертації результати було отримано автором особисто або за його безпосередньої участі.

#### **Повнота викладу результатів у наукових публікаціях.**

Загалом, за матеріалами дисертації опубліковано 9 наукових праць: 6 статей у фахових наукових журналах та 3 публікації у матеріалах закордонних та вітчизняних наукових конференцій.

#### **Структура дисертації**

Дисертація складається з вступу, огляду літератури, матеріалів і методів, результатів досліджень, обговорення результатів досліджень, висновків, списку використаних джерел (102 найменування) і трьох додатків. Роботу викладено на 130 сторінках машинописного тексту і проілюстровано 39 рисунками та 7 таблицями, а також наведено 26 математичних формул.

У роботі чітко сформульовано мету і завдання дослідження. Застосовані методи дослідження відповідають сучасному рівню і логічно обґрунтовані завданням дослідження. Зроблені висновки логічні і відповідають меті і завданням дисертаційного дослідження.

#### **Окремі дискусійні питання і зауваження до дисертації.**

Опонент не має принципових зауважень до роботи. Декілька запитань для дискусії та зауважень технічного характеру наведено нижче:

1. Має місце неправильна побудова огляду літератури. Домінує загальна інформація з значним відривом від теми і завдань дисертаційного дослідження. Детальний розгляд вагових матриць явно є зайвим і недоречним. Враховуючи тему дисертації, огляд літератури повинен починатися зі сторінки 41.
2. Має місце некоректне використання спеціальних символів, що застосовуються для опису вирівнювань біологічних послідовностей. Наприклад, на сторінці 34 (рисунок без назви) символ "\*" використовується Автором для позначення відмінностей, а не як маркер консервативних положень вирівнювання.
3. На першій сторінці Розділу 3 - «Результати досліджень», Здобувач стверджує: «У цій роботі для вивчення проблеми кодонного складу використано геноми бактерій роду *Streptomyces*. Їхні геноми досі не вивчались у цьому керунку, хоча містять виразні

- ПВК і тому мають бути доброю моделлю для таких досліджень.». Це не зовсім відповідає дійсності. Існує велика кількість робіт, починаючи з 90-х років 20-го сторіччя. Наприклад огляд: Leskiw B.K., Bibb M.J., Chater K.F. The use of a rare codon specifically during development? Mol Microbiol. 1991 Dec;5(12):2861-7. PMID: 1809830
4. В розділі 3 («Результати досліджень») присутні значні за об'ємом фрагменти тексту, які за сенсом і змістом мають бути в «Огляді літератури» або «Матеріалах та методах» досліджень.
  5. У мене є принципове питання до Рисунку 3.2, він також Рис. 1 автореферату. З викладеного матеріалу мені не зрозуміло, як автором було здійснено укорінення наведених філогенетичних дерев. Це серйозним чином може впливати на топологію останніх. І тим більше мені не зрозуміло, як це співвідноситься з пов'язаною з ними статтею автора в журналі «Вісник Львівського університету. Серія біологічна. 2016. Випуск 72. С. 75-81». У статті наводяться не вкорінені дерева і є значні відмінності (Рис.2) в порівнянні з результатами, що наведені в дисертації. Також звертаю увагу на некоректну вставку малюнків 3.1 і 3.2. В обох випадках малюнки обрізані по верхньому полю. Таким чином, зазначені дерева наведено частково.
  6. В продовження попереднього зауваження. Є також питання стосовно Рисунку 3.1. На мою думку, якщо ми врахуємо допустиме обертання вузлів, що є цілком допустимим, відмінності стають не на стільки очевидними. І найбільше зводяться до довжини гілок. Таким чином вважаю, що цей момент при реконструкції дерев з використанням різних моделей заміщення, GTR (A) та НКУ (Б), не був врахований належним чином.
  7. Слід зауважити, що автор постійно не дотримується прийнятої структури дисертації. Чимало методичних моментів внесено в результати. Також у розділі 3 - «Результати досліджень» присутні значні відступи, які явно були б доречні в Огляді літератури.
  8. Мені не зовсім зрозумілий виконаний здобувачем перебір різних програм для вирівнювання. Кожна із зазначених програм має своє призначення, а також нюанси додаткових і часто досить тонких налаштувань щодо конкретних задач дослідження. Все це викладено у відповідних посібниках користувача, і на мою думку, первинна задача і визначає вибір програми і методичного підходу. Особисто в мене склалося враження, що в кожному окремому випадку Автором було використано налаштування, що задані за замовчуванням. Якщо це так, то це також може служити джерелом помилок при подальшому аналізі. В тексті дисертації подібні нюанси не було відображено належним чином. Частково дисертант сам приходить до такого висновку на сторінці 101 (Розділ 4. «Обговорення результатів»)

9. Мені не зрозуміло для чого був використаний пакет Clustal Omega. Ця модифікація Clustal використовує приховані моделі Маркова і орієнтована на роботу з дуже великими вибірками послідовностей. Задач, що вимагають використання цього програмного пакету, я в дисертації не побачив.
10. Як видно з тексту дисертації, здобувачем показано існування різних оптимальних моделей синонімічних замін для різних кодуєчих послідовностей. Для мене не зовсім зрозуміло, як це узгоджується з даними роботи Kimura (1981) <https://doi.org/10.1073/pnas.78.1.454>, в якій показано, що швидкості замін синонімічних нуклеотидів приблизно рівні між собою у різних генів (Модель КЗР<sub>u</sub>) (навіть якщо швидкості заміни амінокислот змінюються, і це узгоджується з гіпотезою стосовно випадкового нейтрального мутаційного дрейфу молекулярної еволюції).
11. За даними літератури відомо близько 12-15-ти головних моделей, що описують механізми нуклеотидних замін. Деякі з них були розглянуті в огляді літератури. Також на підставі біоінформатичного дослідження автор наводить найкращі з них (КЗР<sub>u</sub>, GTR, JTT, WAG та LG), що підходять до тих чи інших об'єктів дослідження. Залишається не зрозумілим повний перелік моделей, які було проаналізовано Автором під час виконання дисертаційного дослідження.
12. У разі всіх наведених філогенетичних дерев слід зазначити недостатнє роз'яснення методологічної складової їх реконструкції і методу вкорінення. Цей факт сильно ускладнює оцінку правильності висновків, зроблених здобувачем.
13. У зв'язку з малим масштабом малюнки теплових карт (Рис. 3.4, Додаток А (1)) нечитабельні і, відповідно, не несуть жодної інформації. Крім того, у дисертації присутня неправильна нумерація посилань на додатки - у тексті присутні номери (наприклад, стор.71), а в самому Додатку використані буквені позначення (А, Б, ...).
14. В розділі 3.2. «Особливості контекстного вживання кодонів у геномах стрептоміцетів» згадується 50 видів роду *Streptomyces*, але ні в тексті дисертації, ні в додатку вищезгадані види не вказуються.
15. Розділ 3.3. «Візуалізація кодонних заміщень у геномах *Streptomyces*». Оскільки автор приділяє значний обсяг тексту дисертації опису розробленого веб-сервісу, вважаю що в роботі явно не вистачає алгоритмічної схеми цього інструменту. Це прийнята практика для робіт, в яких є елемент розробки програмного забезпечення.
16. Робота перевантажена ствердженнями, заснованими лише на отриманих автором цифрах і графічних даних. Вважаю, що в наукових працях це не припустимо. Якщо що-небудь стверджується, за цим має йти доказ або посилання на друковану роботу,

де це було доведено. На жаль, в більшості випадків таке підтвердження відсутнє. Наприклад, на стор. 80.

17. На жаль, робота перевантажена великою кількістю опечаток, помилок, невдалих виразів і не канонічних термінів: «не вивчались у цьому керунку», «прикладному вимірі.» «“горизонтальні” та “вертикальні” закономірності вживання кодонів», «позитивних дикодонних асоціацій», «форма рушійного добору на кодонному рівні» тощо. Створюється враження, що рукопис не вчитували.

### **Висновок:**

Незважаючи на численні зауваження, вважаю, що дисертаційна робота Рокицького Ігоря Володимировича «Біоінформатичні підходи та репортерна система для дослідження особливостей вживання кодонів у геномах стрептоміцетів» являє собою завершену наукову працю, присвячену важливій науковій і практичній проблемі. За структурою, актуальністю задач, теоретичним та методичним рівнем, новизною, фундаментальним та практичним значенням отриманих результатів дисертаційна робота цілком відповідає умовам п.11 «Порядку присудження наукових ступенів», затвердженого Постановою Кабінету Міністрів України від 24 липня 2013 р. №567, які пред'являються до кандидатських дисертацій, а її автор заслуговує на присудження наукового ступеня кандидата біологічних наук за спеціальністю 03.00.22 – молекулярна генетика.

### **Офіційний рецензент,**

Завідувач лабораторії біоінформатики та структурної біології,  
Відділу геноміки та молекулярної біотехнології,  
ДУ «Інститут харчової біотехнології та геноміки  
Національної академії наук України»  
кандидат біологічних наук,  
старший науковий співробітник

Карпов П.А.

Підпис зав. лабораторії біоінформатики та структурної біології,  
к.б.н., с.н.с. Карпова П.А.

ЗАСВІДЧУЮ

### **Вчений секретар**

ДУ «Інститут харчової біотехнології  
та геноміки НАН України»,  
к.б.н., с.н.с.

Пірко Я.В.

