

DOI <https://doi.org/10.30525/978-9934-26-226-5-68>

**APPLICATION OF MODERN BIOINFORMATICS METHODS  
TO SOLVE THE PROBLEMS OF SEARCHING  
FOR PATHOGENIC GENES AND PROTEINS  
OF THE DYSENTERY STRAIN E.COLI O104:H4**

**ЗАСТОСУВАННЯ СУЧАСНИХ БІОІНФОРМАТИЧНИХ  
МЕТОДІВ ДЛЯ ВИРІШЕННЯ ЗАДАЧ ПОШУКУ ПАТОГЕННИХ  
ГЕНІВ ТА БІЛКІВ ДИЗЕНТЕРІЙНОГО ШТАМУ  
E.COLI O104:H4**

**Sheiko V. I.**

*Doctor of Biological Sciences, Professor,  
Professor at the department of human  
biology, chemistry and methods  
of teaching chemistry  
Sumy State Pedagogical University  
named after A. S. Makarenko  
Sumy, Ukraine*

**Шейко В. І.**

*доктор біологічних наук, професор,  
професор кафедри біології людини,  
хімії та методики навчання хімії  
Сумський державний педагогічний  
університет імені А. С. Макаренка  
м. Суми Україна*

**Perekhodko K. M.**

*Master of Biotechnology  
and Bioengineering  
Postgraduate Student  
of the specialty 091 – Biology  
Mykola Gogol Nizhyn State University  
Nizhyn, Ukraine*

**Переходько К. М.**

*магістр Біотехнології та Біоінженерії  
здобувач спеціальності 091 – Біологія  
Ніжинський державний університет  
імені Миколи Гоголя  
м. Ніжин, Україна*

**Ivasenko A. Yu.**

*Master's Degree in Publishing and  
Printing Engineering  
Postgraduate Student  
of the specialty 091 – Biology  
Mykola Gogol Nizhyn State University  
Nizhyn, Ukraine*

**Івасенко А. Ю.**

*магістр інженер видавничо-  
поліграфічних виробництв  
здобувач спеціальності 091 – Біологія  
Ніжинський державний університет  
імені Миколи Гоголя  
м. Ніжин, Україна*

Науковий підхід біоінформатики в тому, що вона фокусується на створенні та застосуванні інтенсивних обчислювальних методів задля досягнення цієї мети. Основні зусилля дослідників в галузі біоінформатики спрямовані на вирішення завдань вирівнювання послідовностей, знаходження генів (пошук регіону ДНК, що кодує

гени), розшифровки геному, конструювання ліків, розробки ліків, вирівнювання структури білка, передбачення структури білка, передбачення експресії генів та взаємодій «білок-білок», повногеного пошуку асоціацій та моделювання еволюції [2].

Система BLAST є одним з найбільш широко використовуваних у біоінформатиці програм для пошуку послідовностей. Він звертається до фундаментальної проблеми досліджень у галузі біоінформатики. Використання евристичного алгоритму є набагато швидше ніж інші підходи, такі як обчислення оптимального вирівнювання.

Метою нашого дослідження було вивчення двох штамів лабораторного штаму *E.coli clinical laboratory strain K12* та патогенного штаму *E.coli O104:H4*. Так як обидва штами є досить вивченими в морфології та фізіології, було поставлено унікальний метод порівняння гомологічності (білків) клітинних стінок.

Об'єктом дослідження було біоінформативне дослідження за допомогою сучасних інтернет баз – мембранних білків (поверхневі ефектори TTCC білки).

Таблиця 1

Таблиця детекції генів

Детекція генів	Штами <i>E.coli</i>		
	<i>Escherichia coli</i> <i>O104:H4</i> (taxid:1038927)	<i>Escherichia coli</i> <i>O104:H4 str.</i> <i>01-09591</i> (taxid:1042803)	<i>clinical laboratory strain K12</i> <i>Escherichia coli</i>
rfb	+	+	-
eae	+	+	-
stx1	++	+	-
stx2	++	+	-
ehx	+	+	-
ген синтезу 140M D VirG	++	+	-

Для досягнення мети ми досліджували 2 контрольних зразки різних клінічних штамів *E.coli K12 EH297 та EH83333* для усунення помилок, виділених з різних біотопів організму людини, згідно з наданими даними BLAST. Спочатку експерименту була зроблена кагорта вибірки

частоти народження штаму в біоті людини (кишковик і слизова оболонка) і кількість опрацьованих даних в біоінформатичних базах. Потрібно взяти до уваги, що біоінформатичний аналіз патогенності штамів *E. Coli* бере початок у вивченні генів патогенності, які впливають на синтез патогенних екзотоксинів і мембранних білків. Тому була обрана стратегія – біоінформатична ідентифікація цих генів за допомогою BLAST/FASTA, і вже другим етапом у нас було підтвердження знаходження патогенних мембранних білків безпосередньо у *E. coli* O104:H4 і відповідно – повна їх відсутність у *clinical laboratory strain*.

В таблиці 1 було показано, що спецефічний ген *Stx* включає дві основні імунологічно відмінні форми (*Stx1* і *Stx2*), з мінорними варіантами *Stx2* (від *Stx2a* до *h*). На відміну від жодних генотипних відмінностей, *Stxs* мають багато загальних властивостей, включаючи молекулярну структуру, ферментативну активність, рецептор-специфічність і внутрішньоклітинний трафік.

Отримані дані свідчать про велике генетичне різноманіття факторів патогенності у STEC-штамів різних серогруп і дозволяють глибше зрозуміти та пояснити зв'язок генотипу патогену з його потенційною небезпекою для людини та її клінічної значущістю для інфекціоністів чи клінічних мікробіологів. Реалізовані в проекті авторські методи з експрес ідентифікації та знаходження кодуючих генів «патогенності» – унікальні, що підтверджено біоінформатичним дослідженням на базі BLAST. Комплекс підходів до ідентифікації генів застосовувався до аналізу геномних послідовностей *E. coli* O104:H4 та непатогенного клінічного *laboratory strain E. coli* K12. Згідно з отриманими даними та порівняння з *E. coli* O157:H7 (максимально патогенним), наш досліджуваний штам *E. coli* O104:H4 – придбав плазмиду, що кодує СТХ-М-15, бета-лактамазу класу А розширеного спектру, що надає стійкість до цефтазидим, яка відсутня за даними BLAST у клінічній *laboratory strain E. coli* (непатогенній).

Встановлено, що за допомогою біоінформатичних наукових програм вдалося визначити ключові ланки ідентифікації патогенного *E. coli* O104:H4 (STEC-культура) та непатогенної клінічної *E. coli* K12 за специфічними генами патогенності *rfb*, *eae*, *stx1*, *st*. Було показано, що ентеротоксигенні кишкові палички O104:H4 мають веротоксини (шигоподібні токсини) *Stx1* та *Stx2* (або тільки *Stx2*), білок інтимін, відповідальний за адгезію збудника до епітеліальних клітин кишечника, ентерогемолізін. Перелічені фактори патогенності детермінуються відповідно генами *rfb*, *eae*, *stx1*, *stx2*, *ehx*. Зазначені гени, а також *rfb* гени, відповідальні за синтез соматичного

O-антигену, є майбутніми генами-мішенями для діагностичних біоінформатичних систем, що використовуються під час ідентифікації патогенних штамів. Було визначено, що саме *E.coli* штам O104:H4 містив контрольні відповідно небезпечні гени патогенності *gfb*, *eae*, *stx1*, *stx2*, що контролюють відповідно типів, ентерогемолізину і це було ключовим сегментом експерименту. У ряді перевірочних вибірок експерименту було доведено роль та вплив даних генів на синтез патогенних блоків-білків на клітинній стінці.

### Література:

1. С.В. Горобець, О.Ю. Горобець, І.В. Дем'яненко, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ КПІ ім. Ігоря Сікорського. *Біоінформатика практикум*. 2020. 7-10 стр.
2. С.В. Горобець, О.Ю. Горобець, Т.А. Хоменко, Київ, НТУУ «КПІ», 2018. Основи біоінформатики. Підручник для студентів напряму підготовки 6.051401 «Промислова біотехнологія» факультету біотехнології і біотехніки, Національний технічний університет України «Київський політехнічний інститут», 72-73 стр.
3. Арчаков А. І. Геноміка, протеоміка та біоінформатика – науки XXI століття. *Фармацевтичний вісник*. 2017. № 9 (208). 16-17 стр.
4. Benson, D., I. Karsch-Mizrachi, D. Lipman, J. Ostell, B. Rapp and D. Wheeler. GenBank/BLAST. *Nucleic Acids Research*. 2016. P. 15-18.