

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ МОЛЕКУЛЯРНОЇ БІОЛОГІЇ І ГЕНЕТИКИ**

Гулковський Роман Владиславович



УДК 575.11+577.21

ГЕННІ МУТАЦІЇ У ДІТЕЙ З ІНТЕЛЕКТУАЛЬНОЮ НЕДОСТАТНІСТЮ

03.00.22 – молекулярна генетика

Автореферат
дисертації на здобуття наукового ступеня
кандидата біологічних наук

КИЇВ - 2016

Дисертацією є рукопис.

Роботу виконано у відділі геноміки людини Інституту молекулярної біології і генетики НАН України, м. Київ та на кафедрі загальної та молекулярної генетики Навчально-наукового центру “Інститут біології” Київського національного університету імені Тараса Шевченка, м. Київ.

Науковий
керівник

доктор біологічних наук, професор
Сиволоб Андрій Володимирович,
Навчально-науковий центр “Інститут біології”
Київського національного університету імені Тараса
Шевченка.

Офіційні
опоненти:

доктор біологічних наук, старший науковий співробітник
Макух Галина Василівна,
ДУ “Інститут спадкової патології НАМН України”, м. Львів,
провідний науковий співробітник відділення діагностики
спадкової патології;

доктор медичних наук, старший науковий співробітник
Вайсерман Олександр Михайлович,
ДУ «Інститут геронтології імені Д.Ф. Чеботарьова» НАМН
України, завідувач лабораторією епігенетики.

Захист дисертації відбудеться “11” квітня 2016 р. о 10³⁰ годині на засіданні спеціалізованої вченої ради Д 26.237.01 Інституту молекулярної біології і генетики НАН України за адресою: 03680, м. Київ-680, вул. Заболотного, 150.

З дисертацією можна ознайомитись у бібліотеці Інституту молекулярної біології і генетики НАН України, 03680, м. Київ-680, вул. Заболотного, 150.

Автореферат розіслано “ 10 ” березня 2016 р.

Вчений секретар
спеціалізованої вченої ради
кандидат біологічних наук



І.В. Крупська

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми.

Актуальною задачею сьогодення є дослідження генів, пошкодження яких спричинюють спадкові захворювання людини. Нині відомо кілька тисяч спадкових патологій людини, які асоційовані із мутаціями в одному гені-детермінаторі (моногенні патології). Крім того, були сформовані уявлення про мультифакторні захворювання (МФЗ) та мультифакторні патологічні стани. Передбачається, що в патогенез МФЗ залучена низка мутацій різних функціонально пов'язаних генів (генні мережі), кожен з яких є фактором спадкової схильності. Подальші дослідження генетичних чинників моногенних та мультифакторних патологій допоможуть з'ясувати особливості генетичних програм розвитку та функціонування організму людини, молекулярну природу патогенезу захворювань та старіння.

Приблизно в 50% спадкових патологій людини відмічаються симптоми порушень функцій центральної нервової системи (ЦНС). Значна частина цих патологій асоційована з генетичними факторами, а для інших етіологія та генетична детермінованість до сьогодні не встановлені. Це стосується також і порушень когнітивних функцій, які є одними з найбільш поширених патологій в світі. Молекулярно-генетичні механізми розвитку цих патологій досі вивчені недостатньо через надзвичайну клінічну та генетичну гетерогенність та складність організації нервової системи. Попри блискучі досягнення в галузі нейрогенетики, найважливішою проблемою залишається вивчення ролі різних генів в процесі формування та роботи мозку.

Для більшості генів, що експресуються в ЦНС функціональна роль досі досліджена недостатньо. В цьому аспекті надзвичайно перспективними є дослідження геномних реорганізацій та мутацій (поліморфізмів) генів-кандидатів як чинників патогенезу нейродегенеративних та нейропсихічних патологій. В останні роки розроблені новітні методи повногеномного скринінгу, які дозволяють картувати геномні локуси та ідентифікувати гени-кандидати різних патологічних станів. Визначення генів, білкові продукти яких залучені до розвитку та функціонування ЦНС, є важливим як з точки зору отримання нових важливих фундаментальних знань в галузі нейрогенетики, так і з точки зору дослідження генетичних основ нейродегенеративних захворювань та порушень інтелектуального розвитку (інтелектуальної недостатності - ІН).

Інтелектуальна недостатність є однією з найбільш поширених форм спадкової патології і частіш за все приводить до інвалідності у ранньому дитячому віці. Частота ІН в популяції становить 1-3% (Leonard and Wen, 2002; Maulik et al., 2011). Цей розлад значно впливає на якість життя самої хворої людини, її родини і суспільства в цілому. Близько 75% таких хворих потребують соціальної та освітньої підтримки протягом усього життя (Mefford et al., 2012).

Значна частина випадків ІН асоційована з генетичними факторами, а в решті випадків (~60%) її генетична детермінованість до сьогодні не встановлена (Moeschler and Shevell, 2014). Не зважаючи на останні досягнення в вивченні молекулярних основ деяких форм ІН, наше розуміння патогенезу таких порушень все ще значно обмежене. Тому ідентифікація нових генів-детермінаторів, факторів

спадкової схильності та генних мереж, залучених в патогенез ІН, має ключове значення для подальшого з'ясування молекулярних і фізіологічних механізмів розвитку клітин нервової системи та реалізації когнітивних функцій. Крім того, це забезпечить вирішення проблем діагностики, прогнозування перебігу захворювань, ефективного медико-генетичного консультування та пошуку шляхів терапевтичної корекції соматичних і нервово-психічних розладів при даній патології.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота відповідає основному плану науково-дослідних робіт відділу геноміки людини Інституту молекулярної біології і генетики НАН України і виконувалась в рамках бюджетної теми «Алельний поліморфізм генів, залучених до фенотипу моногенних та мультифакторних патологій» (шифр теми 2.2.4.13, № держ. реєстрації 0110U000695, 2011-2015 рр.). Роботу також виконано в рамках отриманого на конкурсних засадах проекту «Поліпшення діагностики розумової відсталості у дітей Центральної Східної Європи та Центральної Азії за допомогою генетичної характеристики та біоінформатики/статистики» (проект 7 Рамкової програми ЄС CHERISH, номер угоди про фінансову підтримку 223692, 2009-2012 рр.).

Мета і завдання дослідження. Метою роботи є з'ясування ролі мутацій, виявлених в ході повноекзомного секвенування в генах *EPHA1*, *PUS3*, *ZNF527*, *SCEL* та *Cborf223*, а також поліморфізму нуклеотидної послідовності гена *LIF* як генетичних факторів патогенезу інтелектуальної недостатності у дітей.

Для досягнення мети були поставлені основні завдання дослідження:

1. За даними клініко-генеалогічного аналізу та дослідження відомих генетичних чинників патогенезу, а також за результатами повногеномного скринінгу патогенних Copy Number Variants, створити дослідну групу пацієнтів з інтелектуальною недостатністю.
2. Базуючись на результатах повноекзомного секвенування хворих з родини № 094, провести аналіз виявлених змін нуклеотидної послідовності у двох неідентичних братів з інтелектуальною недостатністю та їх батьків.
3. Провести аналіз асоціації виявлених мутацій з патогенезом інтелектуальної недостатності серед членів досліджуваної родини № 094.
4. Провести біоінформатичний аналіз консервативності амінокислотної послідовності білкових продуктів досліджуваних генів в місцях локалізації ідентифікованих мутацій та аналіз третинної структури нормальних та мутантних білкових продуктів відповідних генів.
5. Провести біоінформатичний аналіз поліморфізму нуклеотидної послідовності генів-кандидатів розвитку інтелектуальної недостатності мультифакторної природи з метою вибору функціонально значущих поліморфних варіантів.
6. Розробити методики детекції алельних варіантів генів *EPHA1*, *PUS3*, *ZNF527* та *LIF* на основі полімеразної ланцюгової реакції.
7. Проаналізувати розподіл генотипів та алелів за досліджуваними мутаціями генів *EPHA1*, *PUS3*, *ZNF527* та *LIF* в популяційній вибірці населення України.
8. Вивчити асоціацію алельних варіантів генів *EPHA1*, *PUS3*, *ZNF527* та *LIF* з розвитком легкої інтелектуальної недостатності.

Об'єктом дослідження є спадкова природа інтелектуальної недостатності у дітей.

Предмет дослідження – гени-детермінатори та молекулярно-генетичні чинники спадкової схильності до розвитку інтелектуальної недостатності.

У роботі використовувались наступні **методи дослідження**: виділення та очищення геномної ДНК, полімеразна ланцюгова реакція, аналіз поліморфізму довжини рестрикційних фрагментів, секвенування, біоінформатичне моделювання та статистична обробка даних.

Наукова новизна одержаних результатів. Вперше ідентифіковано міссенс мутації с.1891G>A в 11-му екзоні гена *EPHA1* та с.212A>G в 1-му екзоні гена *PUS3*. Вперше отримано дані про розповсюдження виявлених нами у пацієнтів з інтелектуальною недостатністю алельних варіантів с.1475G>A та с.1891G>A гена *EPHA1*, с.212A>G гена *PUS3* та с.806_808CAT>TGTGCA гена *ZNF527* у популяційній вибірці населення України. Отримано моделі за гомологією просторових структур нормальних та мутантних білкових продуктів генів *EPHA1* та *PUS3*. Отримано дані на користь того, що генотип, до складу якого входять мутації с.1475G>A та с.1891G>A в гені *EPHA1* є генетичним чинником інтелектуальної недостатності моногенної природи. Встановлено, що носійство алельних варіантів 1475A гена *EPHA1* та 4524G гена *LIF* є фактором спадкової схильності до розвитку інтелектуальної недостатності мультифакторної природи. Показано, що носійство алельних варіантів 3022C та 1399A гена *EPHA1* має протекторний ефект, що зумовлює знижений відносний ризик розвитку інтелектуальної недостатності у дітей.

Практичне значення одержаних результатів. Ідентифіковано *EPHA1* як новий потенційний ген-детермінатор інтелектуальної недостатності та розроблено методики детекції мутантних варіантів с.1475G>A та с.1891G>A, що створює передумови для покращення диференційної діагностики інтелектуальної недостатності. Розроблено діагностичні методики для генетичного тестування маркерів спадкової схильності до розвитку інтелектуальної недостатності, що базуються на аналізі поліморфізму довжини рестрикційних фрагментів продуктів специфічної ампліфікації *in vitro* ділянок послідовності генів *EPHA1* та *LIF*.

Особистий внесок здобувача. Весь обсяг експериментальної частини дисертації, пошук та обробку літературних даних виконано автором особисто. Здобувачем особисто проведено: біоінформатичний аналіз досліджуваних мутацій та поліморфних варіантів, дизайн праймерів, підбір температурно-часових режимів проведення полімеразної ланцюгової реакції (ПЛР), а також параметрів електрофоретичного розділення продуктів ПЛР та умов проведення аналізу поліморфізму довжини рестрикційних фрагментів (ПДРФ) для детекції поліморфних варіантів досліджуваних генів. Аналіз клінічних даних проведено спільно з І.М. Бриль, Г.М. Бичковою, Н.О. Афанасьєвою, Е.Й. Пацкун, Н.О. Зимак-Закутньою, С.Ю. Логущ, Л.Ю. Силаєвою (лікарі-генетики медико-генетичних центрів України). Планування досліджень, обговорення, аналіз, інтерпретацію отриманих даних і підготовку публікацій до друку здійснено разом із науковим керівником.

Автор висловлює слова щирої вдячності завідувачу відділу геноміки людини Інституту молекулярної біології та генетики НАН України, д.б.н., проф.

Лівшиць Л.А. за допомогу в розробці стратегії досліджень, виборі підходів до проведення аналізу та узагальненні результатів досліджень, а також всім співробітникам відділу геноміки людини Інституту молекулярної біології і генетики НАН України за підтримку і сприяння у виконанні досліджень. З усіма перерахованими науковцями автор має спільні публікації.

Апробація результатів дисертації. Основні положення роботи доповідались на вітчизняних та зарубіжних з'їздах та конференціях: Європейській конференції з генетики людини (Париж, Франція 2013), Європейській конференції з генетики людини (Мілан, Італія 2014), ВІО 2014 Congress (Варшава, Польща 2014), ІХ Міжнародній конференції молодих науковців "Біологія: від молекули до біосфери" (Харків, Україна 2014), ХІ Міжнародній науковій конференції студентів і аспірантів "Молодь і поступ біології" (Львів, Україна, 2015), Європейській конференції з генетики людини (Глазго, Шотландія, Велика Британія 2015).

Публікації. За темою дисертації опубліковано 11 друкованих праць, зокрема – 5 статей у фахових наукових виданнях та тези 6 доповідей на вітчизняних та міжнародних конференціях і з'їздах.

Структура та обсяг дисертації. Дисертація складається із вступу, огляду літератури, матеріалів і методів дослідження, результатів експериментальних досліджень, обговорення та узагальнення результатів досліджень, висновків та списку використаних джерел. Роботу викладено на 145 сторінках стандартного друкованого тексту, проілюстровано 22 рисунками та 12 таблицями. Список використаної літератури охоплює 255 найменувань.

ОСНОВНИЙ ЗМІСТ

Матеріали та методи дослідження

Матеріалом дослідження були зразки крові індивідів з 95 сімей, включаючи хворих, для яких за результатами клінічного та психіатричного обстеження було встановлено діагноз ідіопатична інтелектуальна недостатність, їх батьків та сибсів. Зразки були надані медичними закладами України: Державною установою "Інститут педіатрії акушерства та гінекології НАМН України" (м. Київ), Державною установою "Інститут спадкової патології НАМН України" (м. Львів), Кримським медико-генетичним центром обласної лікарні (м. Сімферополь), обласним медико-генетичним кабінетом Закарпатської обласної клінічної лікарні ім. Андрія Новака (м. Ужгород), Хмельницьким міським перинатальним центром (м. Хмельницький), обласним центром медичної генетики обласної клінічної лікарні (м. Полтава). Контрольна група складалась з 250 неспоріднених донорів крові з різних регіонів України. Участь в дослідженні та забір крові проводились за умови інформованої згоди.

Виділення та очищення препаратів ДНК проводилося шляхом гідролізу лізатів клітин протеїназою К з наступною фенольною екстракцією (Маниатис и др., 1985). Ампліфікацію *in vitro* послідовностей ДНК ділянок досліджуваних генів та поліморфних локусів проводили за допомогою методу ПЛР (Saiki et al., 1988). Специфічний гідроліз ПЛР продуктів ендонуклеазами рестрикції проводили за умов, рекомендованих фірмою-виробником. Електрофоретичне розділення фрагментів

ДНК проводили в 1,5-2%-ому агарозному гелі в неденатуруючих умовах. Секвенування проводилося на капілярному секвенаторі ABI Prism 3110 Genetic Analyser (Applied Biosystem, США). Статистичну обробку даних проводили за загальноприйнятими методами статистики з використанням програмних пакетів “GenePop” (Rousset F., 2007) та OpenEpi (Sullivan K., 2009). Біоінформатичний аналіз нуклеотидної послідовності досліджуваних генів та поліморфних варіантів проводили з використанням інтернет-ресурсів та баз даних Blast (<http://blast.ncbi.nlm.nih.gov/>), GeneCards: The Human Gene Database (<http://www.genecards.org>), SNPinfo (<http://snpinfonihs.nih.gov/snpinfonihs/index.html>) та “1000 Genomes” (<http://browser.1000genomes.org/>). Множинне вирівнювання послідовностей білкових продуктів досліджуваних генів проводили з використанням веб-сервісу програми Homologene на сайті Національного центру біотехнологічної інформації США (<http://www.ncbi.nlm.nih.gov>) з налаштуваннями за умовчанням. Комп’ютерне моделювання просторової структури білків за гомологією проводили за допомогою web-ресурсу Swiss Model server (<http://swissmodel.expasy.org>).

Результати дослідження та обговорення

Аналіз асоціації ідентифікованих в ході повноекзомного секвенування мутацій з патогенезом інтелектуальної недостатності в родині № 094. За результатами клініко-генеалогічного, цитогенетичного та молекулярно-генетичного (аналіз гена *FMR1* та повногеномний аналіз патогенних Copy Number Variants) аналізу відомих чинників ІН для проведення секвенування екзому було обрано двох хворих сибсів (хлопчиків) з несиндромальною ІН неясної генетичної етіології та подібними психоневрологічними симптомами від здорових неспоріднених батьків (родина №094). За результатами секвенування екзому у обох пацієнтів з родини № 094 було виявлено зміни нуклеотидної послідовності: с.524C>T (rs141416270) та с.1135_1136insA (раніше не описана) в гені *SCEL*; с.585G>T (rs146391418) та с.586C>T (rs138503303) в гені *Cborf223*; с.806_808CAT>TGTGCA (rs386809049) в гені *ZNF527*; с.212A>G (раніше не описана) в гені *PUS3*; с.1475G>A (rs11768549) та с.1891G>A (раніше не описана) в гені *EPHA1*.

Проведене нами секвенування за Сенгером продуктів ПЛР генів *EPHA1*, *PUS3*, *ZNF527*, *SCEL* та *Cborf223* підтвердило результати повноекзомного секвенування пацієнтів та дозволило ідентифікувати генотипи їхніх батьків. З метою аналізу асоціації виявлених мутантних варіантів з ІН порівняльний аналіз змін нуклеотидної послідовності провели у хворих сибсів та їх здорових батьків.

Обидва пацієнти, а також їх здоровий батько виявилися гетерозиготними носіями мононуклеотидних замін (МНЗ) с.585G>T та с.586C>T в гені *Cborf223*, в той час як мати була гомозиготою за алелями дикого (рис. 1.). Водночас, обидва пацієнти та їх здорова матір ідентифіковані як гетерозиготні носії мононуклеотидної заміни с.524C>T (рис. 2а.) та мононуклеотидної інсерції с.1135_1136insA (рис. 2б.) в гені *SCEL*, а батько – як гомозигота за алелями дикого типу. Тому мононуклеотидна заміна с.524C>T (rs141416270) та мононуклеотидна інсерція с.1135_1136insA в гені *SCEL*, а також мононуклеотидні заміни с.585G>T (rs146391418) та с.586C>T (rs138503303) в гені *Cborf223*, були виключені нами зі списку потенційних кандидатів, залучених в патогенез ІН, оскільки виявилось, що один зі здорових

батьків має такі ж генотипи за мутаціями в генах *SCEL* та *C6orf223*, що і хворі сибси.

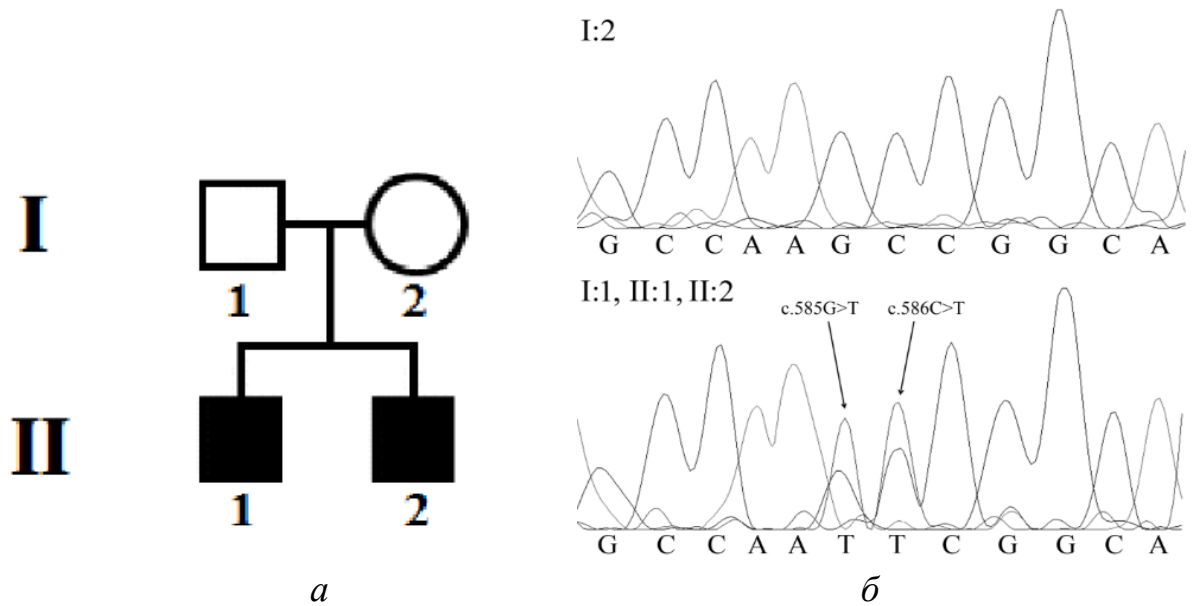


Рис. 1. Родовід (а) та хроматограма секвенування продукту ПЛР гена *C6orf223* зразків ДНК (б) членів родини № 094. II:1 – пацієнт 1, II:2 – пацієнт 2, I:1 – батько, I:2 – матір

Також було встановлено, що обидва пацієнти мають гомозиготний генотип за заміною c.806_808CAT>TGTGCA в гені *ZNF527*, в той час, як батько та мати є гетерозиготними носіями даної заміни (рис. 3а.). За заміною c.212A>G в гені *PUS3* обидва хворих сибси є гомозиготами, в той час, як батько є гетерозиготним носієм, а мати – гомозиготою за алелем дикого типу 212A (рис. 3б.). В даному випадку ми припускаємо, що в оогенезі матері мутація c.212A>G відбулася de novo.

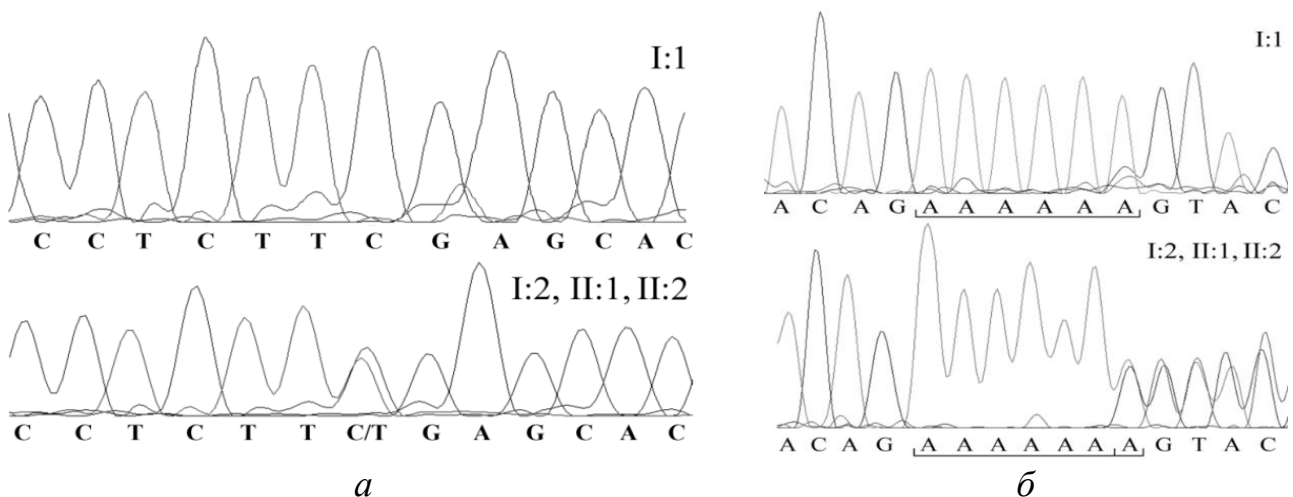


Рис. 2. Хроматограма секвенування продукту ПЛР локусів гена *SCEL* з мононуклеотидною заміною c.524 C>T (а) та мононуклеотидною інсерцією c.1135_1136insA (б) зразків ДНК членів родини № 094: II:1 – пацієнт 1, II:2 – пацієнт 2, I:1 – батько, I:2 – матір

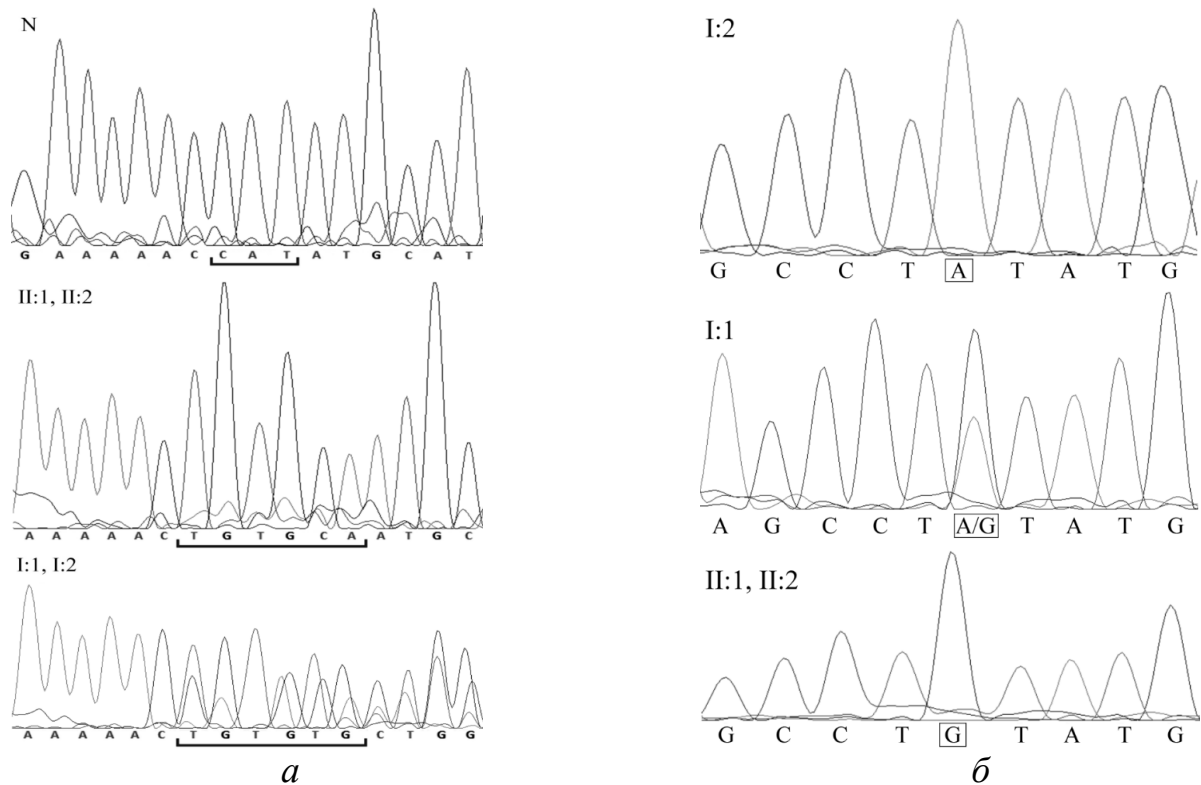


Рис. 3. Хроматограма секвенування продукту ПЛР локусу гена *ZNF527* з заміною с.806_808 CAT>TGTGCA (*a*) та локусу гена *PUS3* з заміною с.212A>G (*б*) зразків ДНК членів родини № 094: II:1 – пацієнт 1, II:2 – пацієнт 2, I:1 – батько, I:2 – матір, N – індивід, гомозиготний за алелем CAT

І на останок, обоє пацієнтів ідентифіковані як компаунди за двома несинонімічними мононуклеотидними замінами с.1475G>A та с.1891G>A в гені *ERHA1*, в той час як батько є гетерозиготним носієм заміни с.1891G>A, а мати – гетерозиготним носієм заміни с.1475G>A (рис. 4.). Таким чином, за результатами сегрегаційного аналізу визначено, що гени *ZNF527*, *PUS3* та *ERHA1* можуть бути генами-кандидатами ІН у сибсів з родини № 094.

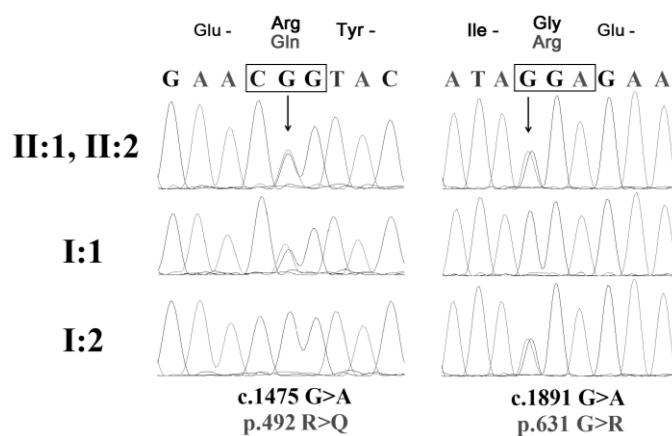


Рис. 4. Хроматограма секвенування продуктів ПЛР гена *ERHA1* зразків ДНК членів родини № 094. II:1 – пацієнт 1, II:2 – пацієнт 2, I:1 – батько, I:2 – мати

Аналіз мутації *rs386809049*, виявленої в ході повноекзомного секвенування в гені *ZNF527*. Ген *ZNF527* (Gene ID: 84503) людини розташований на довгому плечі 19-ої хромосоми в хромосомному локусі 19q13.1, складається з 4 екзонів та 3 інтронів (Nagase et al., 2001). Для гена *ZNF527* на даний момент підтверджено наявність двох альтернативних сплайс-форм, проте найпоширенішим продуктом гена є білок (609 а.к.з.) *ZNF527*, що містить цинкові пальці (zinc finger protein 527). Мутація *c.806_808CAT>TGTGCA* (*rs386809049*) локалізована у кодуючій послідовності гена *ZNF527* і призводить до заміни двох а.к.з на три – *p.Pro269_Tyr270delinsLeuCysAsn* в міждоменній ділянці білка *ZNF527*.

Аналіз результатів проведеного нами множинного вирівнювання амінокислотних послідовностей білків-ортологів *ZNF527* показав, що а.к.з *Pro269* та *Tyr270* є еволюційно консервативними. Це вказує на їх можливу функціональну значимість. Таким чином, *ZNF527* може бути потенційним геном-кандидатом, мутації в якому асоційовані з ІН. *ZNF527* є членом великої родини транскрипційних факторів – Krueppel C2H2-type zinc-finger protein family, – що містять Krueppel-асоційований бокс (Krueppel associated box – KRAB) і представлені тільки у тетрапод (Urrutia R., 2003). Міссенс і нонсенс-мутації в ряді генів білків, що містять цинкові пальці залучені у патогенез Х-зчеплених (*ZNF41*, *ZNF81*, *ZNF674*, *ZNF711*) і аутосомних (*ZNF592*, *ZNF589*) форм несиндромальної інтелектуальної недостатності (Nicolas et al., 2010; Shoichet et al., 2003; Lugtenberg et al., 2006; Tarpey 2009; Kleefstra et al., 2004; Agha et al., 2014) Порушення функціонування генів підроддини протеїнів *ZNF-KRAB* (*ZNF302*, *ZNF181*, *ZNF599* і *ZNF30*) може також бути причиною порушення інтелектуального розвитку при мікрodelеційному синдромі 19q13.11 (Gana et al., 2012).

Тому наступний етап наших досліджень був присвячений порівняльному аналізу розповсюдження алельних варіантів гена *ZNF527* за *rs386809049* у групі 65 пацієнтів з легкою ІН, у яких не виявлено цитогенетичних і відомих молекулярно-генетичних чинників патогенезу ІН та у контрольній групі. Детекцію алельних варіантів гена *ZNF527* проводили за допомогою ПДРФ-аналізу продуктів ампліфікації *in vitro* послідовностей ДНК з використанням ендонуклеази рестрикції *NdeI*. При порівнянні дослідної та контрольної груп у розподілі генотипів та алелів за заміною *rs386809049* гена *ZNF527* достовірної різниці ($p > 0,05$) виявлено не було (табл. 1.).

Таблиця 1

Розподіл генотипів та алелів за заміною *c.806_808CAT>TGTGCA* гена *ZNF527* у пацієнтів з легкою інтелектуальною недостатністю та в контрольній групі

	Контрольна група, n = 250	Група пацієнтів з легкою ІН, n = 64
Генотип, n (%)		
CAT/CAT	169 (67,6)	46(70,8)
CAT/TGTGCA	78 (31,2)	18(27,7)
TGTGCA/TGTGCA	3 (1,3)	1(1,5)
Алель, n (частота)		
CAT	416 (0,832)	110(0,846)
TGTGCA	84 (0,168)	20(0,154)

В групі пацієнтів з легкою ІН було виявлено одного гомозиготного носія мінорного алеля TGTGCA гена *ZNF527*. Проте і в контрольній групі були виявлені три здорових індивіда з подібним генотипом. Оскільки саме цей генотип раніше ідентифіковано у сибсів з несиндромальною ІН, а частота мінорного алеля сягала 16,8% в контрольній групі, заміну с.806_808CAT>TGTGCA в гені *ZNF527* було виключено зі списку потенційних кандидатів, залучених в патогенез ІН.

Аналіз мутації с.212A>G, виявленої в ході повноекзомного секвенування в гені PUS3. Ген *PUS3* (Gene ID: 83480) людини локалізований на довгому плечі 11-ї хромосоми в хромосомному локусі 11q24.2, складається з 3 екзонів та 2 інтронів і кодує псевдоуридинсинтазу 3, до складу якої входить 481 а.к.з (Sato et al., 2013).

Вперше ідентифікована нами мононуклеотидна заміна с.212A>G в 1-му екзоні гена *PUS3* призводить до амінокислотної заміни р.71Y>C, що локалізована в каталітичному домені псевдоуридинсинтази 3 людини. Псевдоуридинсинтаза 3 є членом високо-консервативної родини тРНК-псевдоуридинсинтаз truA і каталізує перетворення уридину на псевдоуридин (Ψ) в 39 положенні антикодонових стебла та петлі (anticodon stem and loop - ASL) більшості ядерних тРНК (Chen and Patton, 2000; Zhao et al., 2007). Псевдоуридин у положеннях 38-40 в ASL тРНК відіграє важливу роль у забезпеченні ефективності і точності трансляції (Agris P., 2004). Було показано, що ці модифікації призводять до підвищення термічної стабільності ASL, і є необхідними для формування коректної антикодон-кодової взаємодії або конформаційних змін тРНК в процесах трансляції (Perret et al., 1990; Zerfass and Beier, 1992; Tomita et al., 1999). Отже, ген *PUS3* є геном «домашнього господарства». Таким чином, мутації в гені *PUS3*, які призводять до втрати псевдоуридинсинтазної активності продукту гена і, як наслідок, відсутність псевдоуридину-39 у багатьох типах тРНК, можуть мати глобальний вплив на ефективність синтезу білка в організмі.

Аналіз результатів проведеного нами множинного вирівнювання амінокислотних послідовностей білків-ортологів показав, що тирозин у позиції 71 є високо-консервативним від *Escherichia coli* до людини (рис. 5.). Це вказує на можливу функціональну значимість цього залишку. Використавши кристалічну структуру каталітичного домену мітохондріальної PUS1 людини як шаблон, ми провели комп'ютерне моделювання за гомологією просторових структур нормального та мутантного каталітичних доменів білка PUS3 людини. Це дало змогу з'ясувати, що заміна високо-консервативного ароматичного тирозину в 71 положенні на сірковмісний цистеїн локалізована в безпосередній близькості від активного центру білка PUS3 (~10Å від каталітичного аспартату) і може призвести до змін в конформаційній гнучкості та просторовій організації каталітичного домену (рис.6.).

Проте, ми не можемо стверджувати, що мутація р.71Y>C призводить до зміни псевдоуридилатсинтазної активності білка PUS3 людини. Особливо враховуючи той факт, що Sibert et al. в експериментах із сайт-направленого мутагенезу PUS1 людини показали, що навіть залишки, безпосередньо залучені до утворення активного центру разом із критичним аспартатом 118 – Y173 (відповідає Y195 в PUS3) та R267 (відповідає R280 в PUS3), – не відіграють істотної ролі у каталізі (Sibert et al., 2008). В той же час, було показано, що відсутність псевдоуридинсинтазної активності

Mutated	A F D F S A H G R R H V A L R I A	C M G W G Y Q G F A S Q E
<i>Homo sapiens</i>	A F D F S A H G R R H V A L R I A	Y M G W G Y Q G F A S Q E
<i>Pan troglodytes</i>	A F D F S A H G R R H V A L R I A	Y M G W G Y Q G F A S Q E
<i>Macaca mulatta</i>	A F D F S A H G R R H V A L R I A	Y M G W G Y Q G F A S Q E
<i>Mus musculus</i>	A F D F S A H G R R H V A L K I A	Y L G W G Y Q G F A S Q E
<i>Rattus norvegicus</i>	A F D F S A H G R R H V A L K I A	Y L G W G Y Q G F A S Q E
<i>Canis lupus familiaris</i>	A F D F S A H G R R H V A L K I A	Y L G W G Y Q G F A S Q E
<i>Bos Taurus</i>	A F D F S A H G Q R H V A L K I A	Y L G W G Y Q G F A S Q E
<i>Gallus gallus</i>	P F D F S A H G R R H V A L R I A	Y L G W G Y Q G F A S Q E
<i>Danio rerio</i>	P F D F S A H P R R H V A L R L A	Y L G W G Y Q G F A V Q E
<i>Xenopus tropicalis</i>	A F D F S A H P K Q H V A L R L A	Y L G W G Y Q G F A S Q E
<i>Drosophila melanogaster</i>	K F D W S S A H K R H V L L K I T	Y F G W D Y Q G F A C Q E
<i>Anopheles gambiae str.</i>	P F D F A K C F K R H I L L R F Y	Y L G W G Y Q G F A A Q E
<i>Caenorhabditis elegans</i>	T L D F L A H P R R K I A I Q F F	Y L G W E H D G L V Q Q P
<i>Saccharomyces cerevisiae</i>	K F D F S K H N T R F I A L R F A	Y L G W N Y N G L A V Q K
<i>Kluyveromyces lactis</i>	E F D F S K Y N T R F V A F K F A	Y L G W N Y N G L A I Q K
<i>Escherichia coli</i>	M S D Q Q Q P P V Y K I A L G I E	Y D G S K Y Y G W Q R Q N
<i>Homo sapiens</i>	A F D F S A H G R R H V A L R I A	Y M G W G Y Q G F A S Q E
<i>Homo sapiens PUS1</i>	-----SKR K I V L L M A	Y S G K G Y H G M Q R N V

Рис. 5. Аналіз результатів множинного вирівнювання амінокислотних послідовностей білків-ортологів *PUS3* людини: консервативні амінокислотні положення виділені сірим кольором, амінокислотне положення Y71 виділене чорним

myopathy and sideroblastic anemia - мітохондріальна міопатія та сидеробластна анемія) (Mitchell et al., 1999; Bykhovskaya et al., 2004; Patton et al., 2005; Fernandez-Vizarra et al., 2006). Місценс-мутації в гені *PUS1* (мітохондріальний аналог *PUS3*) людини, які призводять до заміни консервативної амінокислоти Arg 144 в активному центрі ферменту, та мутації амінокислоти Glu220 були ідентифіковані в якості детермінаторів рідкісного мульти-системного аутосомно-рецесивного синдрому MLASA (Bykhovskaya et al., 2004; Patton et al., 2005; Fernandez-Vizarra et al., 2006). Згадані мутації в гені *PUS1* призводять до втрати псевдоуридину в деяких тРНК, що знижує ефективність синтезу цілого ряду білків. Мутації в гені *DKC1* є причиною дуже рідкісного мульти-системного синдрому Hoeyraal-Hreidarsson та X-зчепленого вродженого дискератозу (Mitchell et al., 1999; Lim et al., 2014).

призводить до суттєвого зниження показників росту колоній прокаріотів (*E. coli* та *S. typhimurium*) (Tsui et al., 1991; Winkler M., 1996).

Також продемонстровано, що нокаутні за геном *PUS3* лінії дріжджів є життєздатними, але темпи росту відповідних колоній значно знижуються, особливо при температурі 37°C (Lecointe et al., 1998). Якщо відсутність псевдоуридинсинтазної активності має такий ефект на ріст прокаріот і дріжджів, яким може бути вплив важливих для каталітичної активності ферменту місценс-мутацій в гені *PUS3* на організм людини?

Фізіологічне значення дисфункції псевдоуридинсинтаз можна оцінити на прикладі таких розладів, як DKC (dyskeratosis congenita - вроджений дискератоз) / синдром Hoeyraal-Hreidarsson та MLASA (mitochondrial

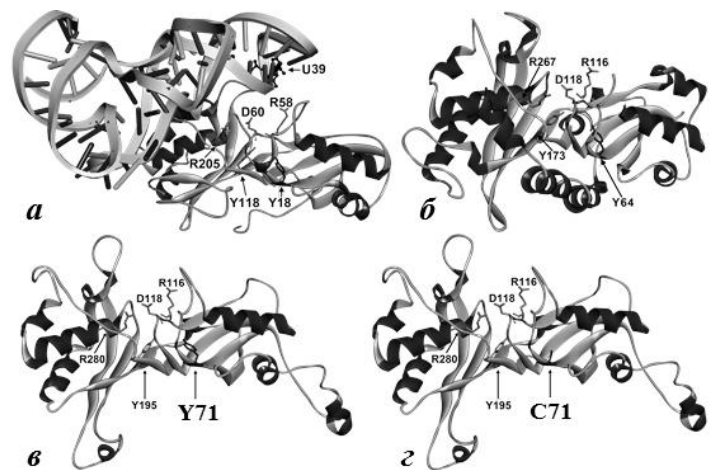


Рис. 6. Кристалічна структура бактеріальної TruA в комплексі з лейцил-тРНК (а, PDB ID: 2NQP) та каталітичного домену *PUS1* людини (б, PDB ID: 4NZ6) і моделі за гомологією каталітичного домену *PUS3* людини дикого типу (в) та мутантного p.71Y>C (г)

Таким чином, очевидно, що мутації в генах псевдоуридинсинтаз, як і в інших генах «домашнього господарства», мають плейотропний ефект, а спричинені ними патології мають мультисистемний синдромальний характер. Подібний плейотропний ефект мав спостерігатися і у випадку ідентифікованої в ході даного дослідження гомозиготної мутації с.212A>G в гені *PUS3*, якщо б ця мутація впливала на каталітичну активність ферменту. Проте, фенотип обох обстежених нами сибсів з ІН має несиндромальний характер. Крім того, ні в анамнезі пацієнтів, ні в ході детального біохімічного аналізу для основних груп білків плазми крові, амінокислот та ацилкарнітинів, проаналізованих за допомогою тандемної мас-спектрометрії, не виявлено відхилень від норми. Тобто ми не отримали свідчень на користь гіпопротеїнемії, яка є загальним симптомом для MLASA та ДКС і могла б бути свідченням порушення синтезу білка. Тому ми схилиємося до думки, що мутація с.212A>G в гені *PUS3* є нейтральною і не може бути основною причиною ІН.

Для проведення порівняльного аналізу розповсюдження алельних мутацій с.212A>G гена *PUS3* в контрольній групі та групі пацієнтів з легкою ІН була розроблена методика ампліфікації *in vitro* послідовностей ДНК з використанням сайт-специфічної ПЛР. За результатами молекулярно-генетичного дослідження вперше отримано дані про розподіл частот генотипів та алельних варіантів за мутацією с.212A>G гена *PUS3* у популяції України. Серед 250 проаналізованих в даному дослідженні зразків ми виявили лише одного гетерозиготного носія мутації с.212A>G, а частота мінорного алеля 212G склала 0,002. Ми не виявили заміну с.212A>G у жодного з 65 індивідів в групі пацієнтів з легкою інтелектуальною недостатністю. Таким чином, асоціації заміни с.212A>G в гені *PUS3* з патогенезом інтелектуальної недостатності доведено не було. Водночас, враховуючи низьку частоту алеля 212G (0,002) в популяції України і той факт, що амінокислотна заміна р.71Y>C знаходиться в безпосередній близькості від активного центру білка *PUS3*, ми не виключаємо, що мутація с.212A>G в гені *PUS3* може бути модифікатором фенотипу при деяких патологіях, включаючи синдромальну інтелектуальну недостатність.

Аналіз мутацій с.1475G>A та с.1891G>A, виявлених в ході повноекзомного секвенування в гені EPHA1. У обох сибсів з несиндромальною ІН нами ідентифікована мононуклеотидна заміна с.1475G>A (rs11768549) в 7 екзоні у компаунді з заміною с.1891G>A (раніше не описана) в 11 екзоні гена *EPHA1*. Ці заміни призводять до амінокислотних замін: р.492R>Q – в фібронектиновому повторі III типу ектодомену та р.631G>R – в гліцин-багатому мотиві (P-loop) тирозинкіназного домену білка *EPHA1*. Ген *EPHA1* (Gene ID: 2041) людини локалізований на довгому плечі 7-ої хромосоми в хромосомному локусі 7q34 і складається з 18 екзонів та 17 інтронів (Maru et al., 1990; Owshalimpur and Kelley, 1999). Білковий продукт гена *EPHA1* молекулярною масою близько 108 кДа складається з 976 амінокислотних залишків (Kullander and Klein, 2002).

За результатами проведеного нами множинного вирівнювання амінокислотних послідовностей білків-ортологів встановлено еволюційну консервативність амінокислот R492 та G631, що вказує на їх можливу функціональну значимість. Комп'ютерне моделювання за гомологією просторових структур нормального та мутантного тирозинкіназного домену білка *EPHA1* людини дало змогу з'ясувати, що

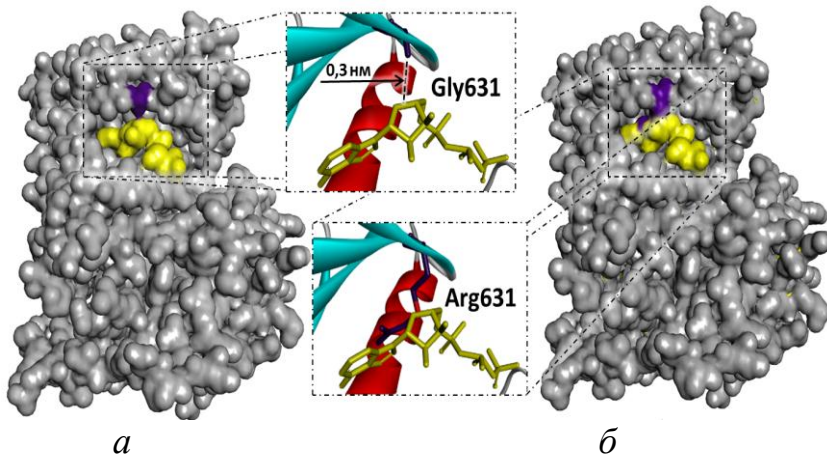


Рис. 7. Модель за гомологією тирозинкіназного домену білка *ERHA1* людини дикого типу (а) та мутантного р.631G>R (б), що були отримані з використання ресурсу Swiss Model server. Амінокислотні залишки Gly631 та Arg631 показані в фіолетовому кольорі, молекули АТФ представлені в жовтому колір

може призвести до зниження ефективності зв'язування АТФ і протеїнкіназної активності загалом.

Заміна позитивно зарядженої амінокислоти аргініну на глутамін в 492 положенні локалізована в фібронектиновому повторі III типу ектодомену ERHA1 (рис. 9.), що виконує роль структурного спейсера і відіграє важливу роль у формуванні інтерфейсів протеїн-протеїнової взаємодії з білками-партнерами та лігандом. Комп'ютерне моделювання за гомологією просторових структур нормального та мутантного ектодомену білка ERHA1 людини дало змогу з'ясувати, що аргінін 492 в послідовності білка ERHA1 людини відповідає аргініну в 486 положенні ERHA4. Аргінін 486 в послідовності білка ERHA4 залучений у формування інтерфейсів протеїн-протеїнової взаємодії з білками-партнерами та лігандом при кластеризації ефринних рецепторів в процесах ефрин/ЕРН-сигналіну (рис. 10.) (Seiradake et al., 2010; Xu et al., 2013). Тому заміна позитивно зарядженої амінокислоти аргініну в 492 положенні на глутамін в ERHA1 може знизити ефективність зв'язування з лігандами або

амінокислотна заміна гліцину в 631 положенні на аргінін локалізована в АТФ-зв'язуючому сайті тирозинкіназного домену білка ERHA1 (рис. 7) і може призвести до зниження ефективності зв'язування АТФ. Як можна бачити з рис. 7б., бічна група аргініну в 631 положенні відстоїть в простір АТФ-зв'язуючого сайту ERHA1 і просторово ускладнює взаємодію з АТФ. Гліцин в 631 положенні є високо-консервативним і входить до числа амінокислотних залишків, що є критично важливими для формування АТФ-зв'язуючого сайту тирозинкіназ (рис. 8.). Таким чином, ми можемо припустити, що мутація р.631 G>R

ERHA1	DPAWLMVDTVI	G	E	G	E	F	G	E	V	Y	R	G	T	L	R	L	P	S
ERHA2	HPSCVTRQKVI	G	A	G	E	F	G	E	V	Y	K	G	M	L	K	T	S	S
ERHA3	DATNISIDKVV	G	A	G	E	F	G	E	V	C	S	G	R	L	K	L	P	S
ERHA4	DASCIKIEKVI	G	V	G	E	F	G	E	V	C	S	G	R	L	K	V	P	G
ERHA5	EASCITIERVI	G	A	G	E	F	G	E	V	C	S	G	R	L	K	L	P	G
ERHA6	DPSRIRIERVI	G	A	G	E	F	G	E	V	C	S	G	R	L	K	T	P	G
ERHA7	DASCIKIERVI	G	A	G	E	F	G	E	V	C	S	G	R	L	K	L	P	G
ERHB1	DVSFVKIEEVI	G	A	G	E	F	G	E	V	Y	K	G	R	L	K	L	P	G
ERHB2	DISCVKIEQVI	G	A	G	E	F	G	E	V	C	S	G	H	L	K	L	P	G
ERHB3	DVSCVKIEEVI	G	A	G	E	F	G	E	V	C	R	G	R	L	K	Q	P	G
FGFR1	PRDRLVLGKPL	G	E	G	C	F	G	Q	V	V	L	A	E	A	I	G	L	D
RET	PRKNLVLGKTL	G	E	G	E	F	G	K	V	Y	K	A	T	A	F	H	L	K
DDR1	PRSRIRFKEKL	G	E	G	Q	F	G	E	V	H	L	C	E	V	D	S	P	Q
SRC	PRESLRLEVKL	G	Q	G	C	F	G	E	V	W	M	G	T	W	N	G	T	T
ABL1	ERTDITMKHKL	G	G	Q	Y	G	E	V	Y	E	G	V	W	K	K	Y	S	
LYN	PRESIKLVKRL	G	A	G	Q	F	G	E	V	W	M	G	Y	N	N	S	T	
BTK	DPKDLTFLKEL	G	T	G	Q	F	G	V	V	K	Y	G	K	W	R	G	Q	Y

Рис. 8. Аналіз результатів множинного вирівнювання амінокислотних послідовностей представників рецепторних та нерцепторних тирозинкіназ людини: консервативні амінокислоти представлені в сірому кольорі, G631 (номер у послідовності ERHA1) – в чорному

білками - партнерами і ефрин / Eph-сигналіngu загалом.

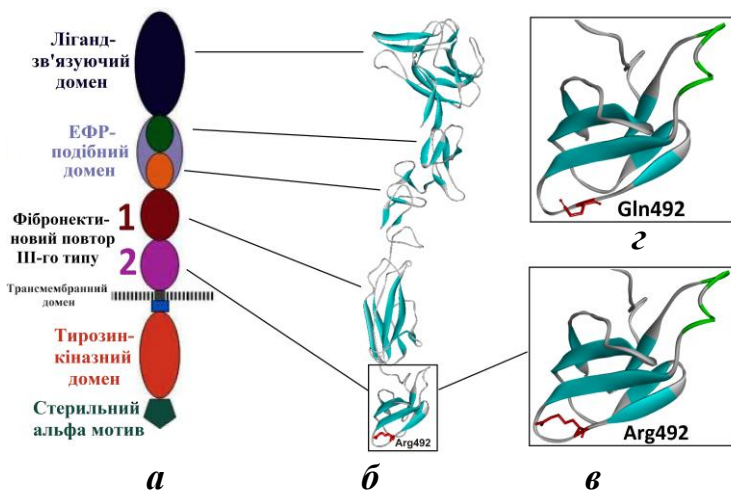


Рис. 9. Схематичне зображення структури ектодомену білка ЕРНА1 людини (а). Модель за гомологією ектодомену білка ЕРНА1 людини дикого типу (б, в) та мутантного р.492R>Q (з), що були отримані з використання ресурсу Swiss Model server: амінокислотні залишки Arg492 і Gln492 показані в червоному кольорі

експресії, проведених на ембріонах миші, які показали, що профіль експресії ЕРНА1 динамічно змінюється в ході гастрюляції, нейруляції та ранніх періодів сомітогенезу (Duffy et al., 2006). Також було показано, що рівень експресії ЕРНА1 зростає в клітинах Пуркінє миші під час їх диференціації (Saywell et al., 2014). Окрім того, ЕРНА1 на високому рівні експресується в головному мозку дорослої людини (Karch et al., 2012).

ЕРНА1 бере участь у прямій сигналізації в клітинах, що експонують рецептор, і зворотній сигналізації в клітинах, що експонують ліганд, через взаємодію з глікозилфосфатидилінозитол-зв'язаними А-ефринами (GPI-linked A Ephrins), що загалом забезпечує зв'язок між сусідніми популяціями клітин, регуляцію процесів міграції клітин та аксонального наведення (Wilkinson D., 2000; Kullander and Klein, 2002; Himanen and Nikolov, 2003; Poliakov et al., 2004; Pasquale E., 2005). Варто відзначити, що основним сигнальним шляхом, через який реалізується ЕРНА1-сигналінг, є Rho-

ЕРНА1 (Ephrin type-A receptor 1) є першим ідентифікованим членом найбільшої підродини рецепторних тирозинкіназ (РТК) – ефринових рецепторів або Eph-рецепторів (erythropoietin-producing hepatocellular receptors family), що відіграє ключову роль в процесах розвитку нервової системи (Hirai et al., 1987; Wilkinson D., 2000; Kullander and Klein, 2002; Noren and Pasquale, 2004). Ефринові рецептори та їх заякорені у мембрану ліганди, відомі як ефрини, беруть участь у регуляції таких біологічних процесів як міграція та адгезія клітин, наведення аксонів, синаптична пластичність, ангиогенез та онкогенез (Wilkinson D., 2000; Kullander and Klein, 2002; Himanen and Nikolov, 2003; Poliakov et al., 2004; Pasquale E., 2005). Роль ЕРНА1 в процесах ембріогенезу та морфогенезу тканин висвітлена в дослідженнях його

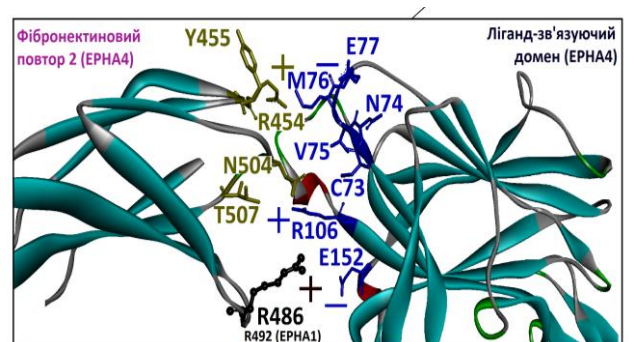


Рис. 10. Інтерфейс взаємодії лігандзв'язуючого домену білка ЕРНА4 з фібронектиновим повтором III-го типу 2 іншого білка ЕРНА4: залучені амінокислотні залишки виділені синім та зеленим кольором, R486 – чорним

ГТФазний сигнальний каскад (Noren and Pasquale, 2004; Yamazaki et al., 2009).

Попередні дослідження показали, що ефрін-А1-індукована активація *EPHA1* призводить до активації RhoA-кінази та інактивації ГТФази Rac1 (Yamazaki et al., 2009). Ці дослідження показують, що *EPHA1*-індуковане зниження активності кінази ІЛК може призвести до гальмування активності білка альфа-PIX (*ARHGEF6*), що активує ГТФазу Rac1, і, як результат, – зсуву рівноваги між Rho-ГТФазами у бік підвищення активності RhoA-кінази (Yamazaki et al., 2009). Rho-ГТФазний сигнальний шлях грає ключову роль у реалізації синаптичної передачі (Ba et al., 2013). Значна кількість генів, які кодують білки, що беруть участь в клітинній сигналізації через Rho-ГТФази або низхідні ефектори цього сигнального каскаду (*OPHN1*, *PAK3*, *ARHGEF6*, *LIMK1*, *FGD1*, *MEGAP* та *FMR1*), є детермінаторами деяких форм інтелектуальної недостатності та/або асоційовані з хворобою Альцгеймера (Nadif Kasri and Van Aelst, 2008; Ba et al., 2013). Отже, є підстави розглядати *EPHA1* як перспективний ген-кандидат, асоційований з ІН, білковий продукт якого залучений до розвитку та функціонування ЦНС.

Для оцінки розповсюдження МНЗ с.1475G>A та с.1891G>A гена *EPHA1*, виявлених у хворих сибсів з родини №094, проводився порівняльний аналіз розподілу генотипів та алельних варіантів в групі пацієнтів з легкою ІН та контрольній групі. Для проведення молекулярно-генетичного аналізу алельних варіантів с.1475G>A та с.1891G>A гена *EPHA1* нами було розроблено методики з використанням ПЛР з наступним ПДРФ-аналізом та сайт-специфічної ПЛР, відповідно.

За результатами даного дослідження заміна с.1891G>A не була виявлена у жодного з 315 індивідів, включених в дане дослідження, як в популяційній вибірці неспоріднених донорів з різних регіонів України (n=250), так і в групі пацієнтів з легкою інтелектуальною недостатністю (n=65). Крім того, МНЗ с.1891G>A не було ідентифіковано в жодній з популяцій, досліджених в ході таких проектів, як Exome Aggregation Consortium (URL: <http://exac.broadinstitute.org>), International HarMap Project (URL: <http://hapmap.ncbi.nlm.nih.gov/>) та 1000 Genomes Project (URL: <http://browser.1000genomes.org/>). Таким чином, ми припускаємо, що транзиція с.1891G>A в 11-му екзоні гена *EPHA1* є патогенною міссенс-мутацією, яка зустрічається з украй низькою частотою. Що стосується нуклеотидної заміни с.1475G>A в гені *EPHA1*, у контрольній групі нами було виявлено лише 5 гетерозиготних носіїв мінорного алеля 1475A – його частота в популяції України становить 0,01 (табл. 2.). Порівняння з частотами цього алеля в різних популяціях світу, отриманими в рамках проекту “1000 геномів”, дозволяє зробити висновок про відсутність суттєвих відмінностей у розповсюдженні даного варіанта в популяціях з населенням європейського походження та популяції України. Частота мінорного алеля 1475A гена *EPHA1* в групі пацієнтів з легкою ІН (0,038) була достовірно (p<0,05) вищою порівняно з контрольною групою (0,01). Більше того, за розрахунками показника відношення шансів встановлено, що індивіди-носії мінорного алеля 1475A гена *EPHA1* мають в 4 рази вищий відносний ризик розвитку ІН (OR= 4,05; ДІ 95%: 1,15 – 14,56; p=0,02). Таким чином, нами отримано дані на користь того, що МНЗ с.1475G>A (rs11768549) гена *EPHA1* може бути залученою до патогенезу ІН у хворих з родини №094 і її можна розглядати як імовірний фактор спадкової схильності до розвитку ІН.

Розподіл генотипів та алелів за заміною с.1475G>А, гена *EPHA1* у пацієнтів з легкою інтелектуальною недостатністю та в контрольній групі

	Контрольна група, n = 250	Група пацієнтів з легкою ІН, n = 65
с.1475G>А		
Генотип, n (%)		
GG	245 (98)	60 (92,3)
GA	5 (2)	5 (7,7)
AA	0	0
Алель, n (частота)		
G	495 (0,99)	125 (0,962)
A	5 (0,01)	5 (0,038)*

Примітки: n – кількість індивідів; * – різниця між значенням показника в досліджуваних групах є статистично достовірною ($p < 0,05$)

На користь цього також можуть свідчити дані отримані в ході кількох GWAS (genome-wide association study) досліджень, в яких було виявлено асоціацію ідентифікованої нами заміни с.1475G>А (rs11768549), а також мутації rs202178565 та поліморфізмів rs11767557 і rs11771145 гена *EPHA1* з хворобою Альцгеймера – нейродегенеративним захворюванням, що супроводжується стійким порушенням когнітивних функцій (деменцією) в результаті органічного ураження мозку (Seshadri et al., 2010; Naj et al., 2011; Hollingworth et al., 2011; Wang et al., 2015; Vardarajan et al., 2015). Встановлено, що мінорні алельні варіанти п.3022С (rs11767557, синонім п.75-3022Т>С) та п.1399А (rs11771145, синонім п.75-1399Г>А) гена *EPHA1* асоційовані зі знизеним відносним ризиком розвитку хвороби Альцгеймера. Також було показано, що варіант с.1475G>А гена *EPHA1* асоційований зі швидкою прогресією деменції при хворобі Альцгеймера (Wang et al., 2015).

Ми вважаємо, що отримані дані про розподіл генотипів за МНЗ с.1475G>А та с.1891G>А в гені *EPHA1*, а також локалізація відповідних амінокислотних замінів р.492R>Q та р.631G>R в консервативних мотивах функціонально важливих доменів залученого до розвитку та функціонування ЦНС рецептора *EPHA1*, є вагомими аргументами на користь того, що вперше ідентифікована нами мутація с.1891G>А в компаунді з мононуклеотидною заміною с.1475G>А (rs11768549) гена *EPHA1* є детермінуючим генетичним чинником інтелектуальної недостатності у пацієнтів з обстеженої родини №094. Оскільки обоє пацієнтів мають компаундний генотип за двома несинонімічними МНЗ с.1475G>А та с.1891G>А в гені *EPHA1*, в той час, як здоровий батько є гетерозиготним носієм заміни с.1891G>А, а здорова мати – гетерозиготним носієм заміни с.1475G>А, ми констатуємо, що несиндромальна ІН у пацієнтів з обстеженої родини успадковується за аутосомно-рецесивним типом.

EPHA1-сигналінг реалізується через Rho-ГТФазний сигнальний каскад, який є однією з небагатьох відомих на сьогодні генних мереж, асоційованих з ІН (Noren and Pasquale, 2004; Yamazaki et al., 2009). Гени *OPHN1*, *PAK3*, *ARHGEF6*, *LIMK1*,

FGD1, *MEGAP* та *FMRI*, що кодують білки Rho-ГТФазного шляху та його ефектори, були ідентифіковані в якості детермінаторів інтелектуальної недостатності та/або асоційовані з хворобою Альцгеймера (Nadif Kasri and Van Aelst, 2008; Ba et al., 2013). Ми пропонуємо ген *EPHA1* як новий, асоційований з Rho-ГТФазним сигнальним шляхом, ген-детермінатор інтелектуальної недостатності, білковий продукт якого залучений до розвитку та функціонування ЦНС.

Дослідження поліморфізму генів EPHA1 та LIF як факторів спадкової схильності розвитку легкої інтелектуальної недостатності. Переважна більшість останніх досліджень, метою яких була ідентифікація генів-детермінаторів або генів-кандидатів ІН, були зосереджені на вивченні випадків помірної та важкої ІН. В той же час, легку інтелектуальну недостатність, частка якої становить 85% всіх випадків, розглядають як мультифакторне захворювання із спадковою схильністю.

На сьогодні, для пошуку генів, ймовірно залучених до розвитку мультифакторних патологій, існують два основні підходи. Перший з них (так звані GWAS) полягає в рандомізованому дослідженні варіювання на рівні усього геному з наступним аналізом асоціації поліморфних варіантів з патологією. Альтернативним є підхід вивчення генів-кандидатів. Він полягає в підборі вже відомих за своєю функцією генів та дослідженні асоціації їх поліморфізмів з певною патологією у порівнянні з контролем.

Виділений нами в ході даного дослідження як ген-детермінатор *EPHA1* був обраний для аналізу в якості гена схильності до мультифакторних форм інтелектуальної недостатності. За результатами проведеного нами біоінформатичного аналізу послідовності регуляторних ділянок гена *EPHA1*, на додачу до ідентифікованої в ході секвенування екзому двох хворих сибсів з несиндромальною ІН МНЗ с.1475G>A, нами були обрані для дослідження МНЗ п.3022Т>С (rs11767557) та п.1399G>А (rs11771145) локалізовані у регуляторній некодуючій послідовності *EPHA1-AS1* (*EPHA1 antisense RNA 1*) поряд з промотором гена як можливі чинники регуляції експресії *EPHA1*. На користь нашого вибору вказують дані кількох GWAS-досліджень, в яких було виявлено асоціацію поліморфізмів п.3022Т>С та п.1399G>А гена *EPHA1* з нейродегенеративним захворюванням – хворобою Альцгеймера (Seshadri et al., 2010; Naj et al., 2011; Hollingworth et al., 2011).

В якості ще одного генетичного чинника схильності до мультифакторних форм інтелектуальної недостатності може розглядатися ген *LIF*, що кодує нейроцитокін лейкемія-інгібуючий фактор (LIF), який, як і ген *EPHA1*, є ключовим регулятором процесів ембріогенезу та нейрогенезу у ссавців (Barres et al., 1993; Sugiura et al., 2000; Holmberg and Patterson, 2006). Ген *LIF*, як і *EPHA1*, є частиною однієї з небагатьох відомих на сьогодні генних мереж, асоційованих з ІН, – ERK/MAP-кіназного сигнального шляху, який відіграє ключову роль у регуляції росту клітин, їх поділу та проліферації (Niwa et al., 2009; Hirai et al., 2011). Гени (наприклад, *RPS6KA3* та *SYNGAP1*), що кодують білки ERK/MAPK-шляху та його ефектори, також були ідентифіковані в якості детермінаторів несиндромальної інтелектуальної недостатності (Field et al., 2006; Krepischi 2010; et al., Hamdan 2011). Біоінформатичний аналіз послідовності регуляторних ділянок гена *LIF* дозволив нам обрати для дослідження поліморфізм с.4524Т>G (rs929271). Показано, що цей поліморфізм, який розташований в 3'-нетрансльованій області (3'-НТО) гена *LIF*, може спричинювати зменшення

стабільності мРНК і, тим самим, впливати на рівень секреції LIF (Ishida et al., 2001). В попередніх дослідженнях було показано асоціацію поліморфізму с.4524Т>G з шизофренією та погіршенням функції робочої пам'яті (Okahisa et al., 2010).

Отже, на наступних етапах наших досліджень, з метою пошуку факторів спадкової схильності розвитку інтелектуальної недостатності, ми провели порівняльний аналіз розподілу генотипів та алельних варіантів за поліморфними локусами п.3022Т>С (rs11767557) і п.1399G>А (rs11771145) гена *EPHA1* (табл. 3) та с.4524 Т>G гена *LIF* (табл. 4) в групі пацієнтів з легкою ІН та контрольній групі.

Таблиця 3

Розподіл генотипів та алелів за мононуклеотидними замінами п.3022Т>С та п.1399G>А гена *EPHA1* у пацієнтів з легкою інтелектуальною недостатністю та в контрольній групі.

	Контрольна група, n = 250	Група пацієнтів з легкою ІН, n = 65
п.3022Т>С		
Генотип, n (%)		
ТТ	141 (56,4)	47 (72,3)
ТС	91 (36,4)	17 (26,2)
СС	18 (7,2)	1 (1,5)
ТС+СС	109(43,6)	18(27,7)*
Алель, n (частота)		
Т	373 (0,746)	111 (0,854)
С	127 (0,254)	19 (0,146)*
п.1399G>А		
Генотип, n (%)		
GG	96 (38,4)	31 (47,7)
GA	117 (46,8)	31 (47,7)
AA	37 (14,8)	3 (4,6)*
GG+AG	213 (85,2)	62 (95,4)*
Алель, n (частота)		
G	309 (0,618)	93 (0,715)
A	191 (0,382)	37 (0,285)*

Примітки: n – кількість індивідів; * – різниця між значенням показника в досліджуваних групах є статистично достовірною (p<0,05)

Молекулярно-генетичний аналіз поліморфізмів с.4524 Т>G в гені *LIF* та п.3022Т>С і п.1399G>А в гені *EPHA1* проводився на основі розроблених нами методів ПЛР з наступним аналізом методом ПДРФ (рис. 11.-13.).

При порівнянні розподілу генотипів за алельними варіантами п.3022Т>С та п.1399G>А в дослідній та контрольній групах виявили статистично достовірне (p=0,02) перевищення частки носіїв мінорного алеля 3022С гена *EPHA1* в контрольній групі (0,436) порівняно з групою пацієнтів з легкою ІН (0,277).

Таблиця 4

Розподіл генотипів та алелів за мононуклеотидною заміною с.4524 Т>G гена *LIF* у пацієнтів з легкою інтелектуальною недостатністю та в контрольній групі

	Контрольна група, n = 250	Пацієнти з легкою ІН, n = 64
Генотип, n (%)		
ТТ	113 (45,2)	18 (30,8)
TG	111 (44,4)	35 (49,4)
GG	26 (10,4)	11 (19,8)
TG+GG	137(54,8)	46(71,9)*
Алель, n (частота)		
Т	336 (0,672)	71 (0,555)
G	164 (0,328)	57 (0,445)*

Примітки: n – кількість індивідів; * – різниця між значенням показника в досліджуваних групах є статистично достовірною ($p < 0,05$)

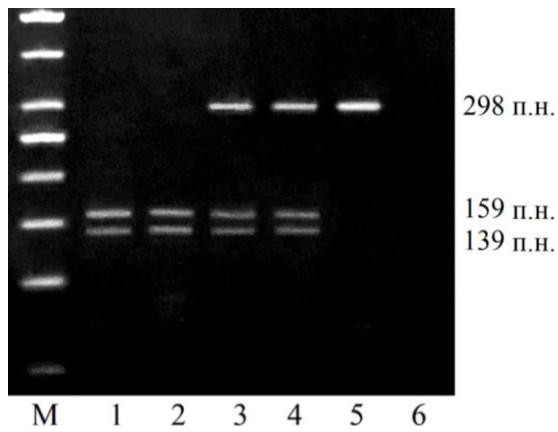


Рис. 11. Електрофореграма розділення фрагментів ПЛР-продукту гена *LIF*, що утворились після гідролізу ендонуклеазою рестрикції *HinfI*, в 2% агарозному гелі: М – маркер молекулярної маси (з кроком 50 п.н.); 1, 2 – індивіди з гомозиготним генотипом ТТ, 3, 4, – індивіди з гетерозиготним генотипом TG, 5 – індивід з гомозиготним генотипом GG, 6 – негативний контроль (H₂O)

Частота гомозиготного за мінорним алелем генотипу 1399AA гена *EPHA1* в контрольній групі (0,148) була достовірно ($p=0,03$) вищою порівняно з групою пацієнтів з легкою ІН (0,046). Більше того, за розрахунками показника відношення шансів

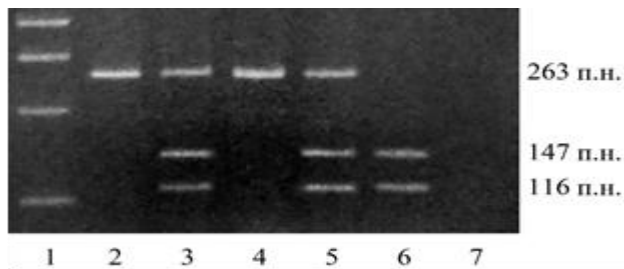


Рис. 12. Електрофореграма розділення фрагментів ПЛР-продукту гена *EPHA1* (п.3022Т>С), що утворились після його гідролізу специфічною ендонуклеазою рестрикції *RsaI*: 1 – маркери молекулярної маси (з кроком 100 п.н.); 2, 4 – індивіди, що мають генотип ТТ; 3, 5 – індивіди з генотипом ТС; 6 – індивід з генотипом СС; 7 – негативний контроль (H₂O)

встановлено, що: індивіди-носії мінорного алеля 3022С гена *EPHA1* мають вдвічі нижчий відносний ризик розвитку ІН (OR=0,5; ДІ 95%: 0,27 – 0,9; $p=0,02$); індивіди-носії мінорного алеля 1399А гена *EPHA1* мають приблизно втричі нижчий відносний ризик розвитку ІН (OR=0,28; ДІ 95%: 0,08 – 0,93; $p=0,03$). Також було встановлено, що частота мінорного алеля 4524G гена *LIF* в групі пацієнтів з легкою ІН (0,445) є достовірно ($p=0,01$) вищою порівняно з контрольною групою (0,328). За розрахунками показника відношення шансів встановлено, що: індивіди-носії мінорного алеля 4524G гена *LIF* мають вдвічі вищий відносний ризик розвитку ІН.

Підсумовуючи вищенаведене, можемо стверджувати, що за результатами проведеного наукового дослідження вдалося отримати свідчення на користь залучення мутантних варіантів с.1475G>A та с.1891G>A гена *EPHA1* в патогенез інтелектуальної недостатності, що дозволяє нам виділити ген *EPHA1* в якості нового, асоційованого з Rho-ГТФазним сигнальним шляхом, гена-детермінатора цього захворювання. Встановлено, що алельні варіанти с.1475G>A гена *EPHA1* та с.4524T>G гена *LIF* є генетичними маркерами схильності до розвитку мультифакторних форм інтелектуальної недостатності. Також показано, що носійство мінорних алельних варіантів 3022C та 1399A гена *EPHA1* має протекторний ефект, що зумовлює знижений відносний ризик розвитку інтелектуальної недостатності.

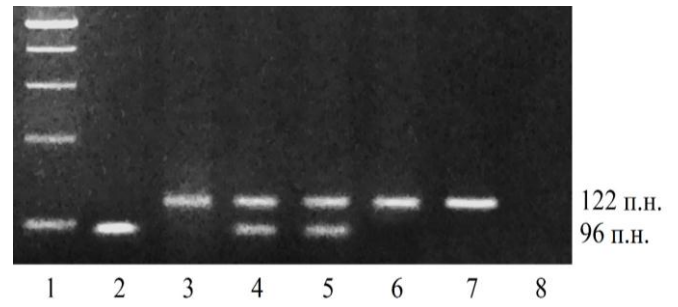


Рис. 13. Електрофореграма розділення фрагментів ПЛР-продукту гена *EPHA1* (n.1399G>A), що утворились після його гідролізу специфічною ендонуклеазою рестрикції *DraI*: 1 – маркери молекулярної маси (з кроком 100 п.н.); 2,– індивід, що має генотип AA; 4, 5 – індивіди з генотипом GA; 3, 6, 7 – індивіди з генотипом GG; 8 – негативний контроль (H₂O)

ВИСНОВКИ

Вирішення поставлених в даній роботі завдань дозволило ідентифікувати *EPHA1* як новий потенційний ген-детермінатор інтелектуальної недостатності та встановити асоціацію поліморфізмів с.1475G>A (rs11768549), н.3022T>C (rs11767557), н.1399G>A (rs11771145) гена *EPHA1* та с.4524T>G (rs929271) гена *LIF* із спадковою схильністю до розвитку інтелектуальної недостатності.

1. Вперше ідентифіковано міссенс мутації с.1891G>A в 11-му екзоні гена *EPHA1* та с.212A>G в 1-му екзоні гена *PUS3* та оцінено їх розповсюдженість в популяції здорового населення України.
2. Не виявлено асоціації мутацій с.524C>T (rs141416270) та с.1135_1136insA в гені *SCEL*, с.585G>T (rs146391418) та с.586C>T (rs138503303) в гені *Cborf223*, с.806_808CAT>TGTGCA (rs386809049) в гені *ZNF527* та с.212A>G в гені *PUS3* з патогенезом інтелектуальної недостатності.
3. Показано, що амінокислотні заміни р.492R>Q та р.631G>R локалізовані в високо-консервативних мотивах функціонально важливих доменів *EPHA1* і можуть призвести до зниження ефективності зв'язування з лігандами або білками-партнерами (р.492R>Q) та зв'язування АТФ з тирозинкіназним доменом (р.631G>R) і ефективності ефрин/*EPHA1*-сигналіngu загалом.

4. Доведено, що носійство алельних варіантів 1475A (rs11768549) гена *EPHA1* та 4524G (rs929271) гена *LIF* є фактором спадкової схильності до розвитку інтелектуальної недостатності.
5. Показано, що носійство алельних варіантів 3022C (rs11767557) та 1399A (rs11771145) гена *EPHA1* має протекторний ефект, що зумовлює знижений відносний ризик розвитку інтелектуальної недостатності.

ПЕРЕЛІК НАУКОВИХ ПРАЦЬ, ОПУБЛІКОВАНИХ ЗА ТЕМОЮ ДИСЕРТАЦІЇ:

1. *EPHA1* gene SNPs analysis in population of Ukraine / Gulkovskiy R.V., Chernushyn S.Y., Kravchenko S.A., Bychkova G.M., Livshits L.A. // *Biopolymers & cell.* – 2013. – Vol. 29, №.5. – P. 506–510. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алельними варіантами, біоінформатичний аналіз, статистична обробка даних.*
2. Association of the leukemia inhibitory factor gene polymorphism rs929271 with idiopathic mild intellectual disability / Gulkovskiy R. V., Volkova L. S., Livshits L. A. // *Biopolymers & cell.* – 2015. – Vol. 31, №.1. – P. 34-37. *Особистий внесок здобувача – узагальнення отриманих клініко-генеалогічних даних, генотипування індивідів за досліджуваними алельними варіантами, статистична обробка даних.*
3. Novel gene *PUS3* c.A212G mutation in Ukrainian family with intellectual disability / Gulkovskiy R.V., Chernushyn S.Y., Livshits L.A. // *Biopolymers and Cell.* – 2015. – Vol. 31, №.2. – P. 123-130. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алельними варіантами, біоінформатичний аналіз, статистична обробка даних.*
4. Association of the *EPHA1* gene polymorphism with idiopathic mild intellectual disability / Gulkovskiy R. V., Sivolob A.V., Livshits L. A. // *Biopolymers & cell.* – 2015. – Vol. 31, №.4. – P. 271-276. *Особистий внесок здобувача – узагальнення отриманих клініко-генеалогічних даних, генотипування індивідів за досліджуваними поліморфними варіантами, статистична обробка даних.*
5. *ZNF527* Gene rs386809049 Analysis in Population of Ukraine / Gulkovskiy RV, Chernushyn S.Y., Kravchenko S.A., Livshits L.A. // *Cytology and Genetics.* – 2015. – Vol. 49, №.4. – P. 240–244. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алельними варіантами, біоінформатичний аналіз, статистична обробка даних.*
6. *EPHA1* C1475T and C1891T polymorphisms in Ukrainian patients with idiopathic intellectual disability / R. V. Gulkovskiy, G.M. Bychkova, S.A. Kravchenko, N.V.Hryshchenko, S.Y. Chernushyn, M.Nõukas, M.Sauk, L. Milani, T. Pippucci, F. Balombo, C. Graziano, L. A. Livshits // *European Journal of Human Genetics*, Vol. 20 Suppl. 1, European human genetics conference 2013, June 8 - 11, 2013, Paris, France, P. 175. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алельними варіантами, біоінформатичний аналіз, статистична обробка даних.*
7. *EPHA1* as a new candidate gene for autosomal recessive non-syndromic intellectual disability / R. Gulkovskiy G. Volynets, S. Chernushyn, S. Kravchenko, L. A. Livshits // *European Journal of Human Genetics*, Vol. 22 Suppl. 1, European human genetics conference 2014, May 31 - June 3, 2014, Milan, Italy, P. 153. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алельними варіантами, біоінформатичний аналіз, статистична обробка даних.*
8. *EPHA1* as a new candidate gene for non-syndromic intellectual disability / R. Gulkovskiy G. Volynets, S. Chernushyn, S. Kravchenko, L. A. Livshits // *АСТА*

Biochimica Polonica, Vol. 61 Suppl. 1, Abstracts of the BIO 2014 Congress, September 9-12, 2014, Warsaw, Poland, P. 120. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алейними варіантами, біоінформатичний аналіз, статистична обробка даних.*

9. Аналіз rs929271 гена *LIF* в популяції України / Гулковський Р.В., Волкова Л. С. // Матеріали ІХ Міжнародної конференції молодих науковців “Біологія: від молекули до біосфери”, 18-20 листопада, 2014, Харків, Україна, С. 48-49. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алейними варіантами, статистична обробка даних.*

10. *LIF* gene polymorphism rs929271 and association with idiopathic mild intellectual disability / R. Gulkovskiy, L. Volkova, L. Livshits // Abstracts of XI International Scientific Conference for Students and PhD Students “Youth and Progress of Biology”, April 20-23, 2015. Lviv, Ukraine, P. 194-195. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алейними варіантами, статистична обробка даних.*

11. Leukemia inhibitory factor gene polymorphism rs929271 is associated with mild intellectual disability / Gulkovskiy R. V., Volkova L. S., Livshits L. A. // European Journal of Human Genetics, Vol. 23 Suppl. 1, European human genetics conference 2015, June 6 - 9, 2015, Glasgow, Scotland, United Kingdom, P. 153. *Особистий внесок здобувача – генотипування індивідів за досліджуваними алейними варіантами, статистична обробка даних.*

АНОТАЦІЯ

Гулковський Р.В. Генні мутації у дітей з інтелектуальною недостатністю. - Рукопис.

Дисертація на здобуття наукового ступеня кандидата біологічних наук за спеціальністю 03.00.22 - молекулярна генетика. - Інститут молекулярної біології і генетики НАН України, Київ, 2016.

Дисертаційна робота здобувача присвячена дослідженню ролі мутацій в генах *EPHA1*, *PUS3*, *ZNF527*, *SCEL* та *Sborf223*, виявлених в ході повноекзомного секвенування у двох сибсів з інтелектуальною недостатністю, а також поліморфізму нуклеотидної послідовності гена *LIF* як генетичних факторів патогенезу інтелектуальної недостатності.

Вперше ідентифіковано міссенс мутації с.1891G>A в 11-му екзоні гена *EPHA1* та с.212A>G в 1-му екзоні гена *PUS3*. Вперше отримано дані про розповсюдження виявлених нами у пацієнтів з інтелектуальною недостатністю алейних варіантів с.1475G>A (rs11768549) та с.1891G>A (нова мутація) гена *EPHA1*, с.212A>G (нова мутація) гена *PUS3* та с.806_808CAT>TGTGCA (rs386809049) гена *ZNF527* у популяційній вибірці з України. Отримано моделі просторових структур нормальних та мутантних білкових продуктів генів *EPHA1* та *PUS3*. Вперше отримано докази на користь того, що мутації с.1475G>A та с.1891G>A в гені *EPHA1* є генетичними чинниками патогенезу інтелектуальної недостатності моногенної природи. Розроблено методики детекції мутантних варіантів с.1475G>A та с.1891G>A гена *EPHA1*, а також мононуклеотидних замінів rs11767557 і rs11771145 в гені *EPHA1* та rs929271 в гені *LIF*, придатні для генетичного тестування в групах з інтелектуальною недостатністю. Встановлено, що носійство алейних варіантів 1475A гена *EPHA1* та 4524G (rs929271) гена *LIF* є фактором спадкової схильності до розвитку інтелектуальної недостатності. Показано, що носійство алейних варіантів

3022C (rs11767557) та 1399A (rs11771145) гена *EPHA1* зумовлює зниження відносного ризику розвитку інтелектуальної недостатності.

Ключові слова: мутація, алельний поліморфізм, мононуклеотидна заміна, ген *EPHA1*, ген *LIF*, інтелектуальна недостатність.

АННОТАЦИЯ

Гулковский Р.В. Генные мутации у детей с интеллектуальной недостаточностью. - Рукопись.

Диссертация на соискание ученой степени кандидата биологических наук по специальности 03.00.22 - молекулярная генетика. - Институт молекулярной биологии и генетики НАН Украины, Киев, 2016.

Диссертационная работа соискателя посвящена исследованию роли мутаций в генах *EPHA1*, *PUS3*, *ZNF527*, *SCEL* и *Sborf223*, выявленных в ходе полноэкзомного секвенирования у двух сиблингов с интеллектуальной недостаточностью, а также полиморфизма нуклеотидной последовательности гена *LIF* как генетических факторов патогенеза интеллектуальной недостаточности.

Впервые идентифицированы миссенс мутации с.1891G>A в 11-м экзоне гена *EPHA1* и с.212A>G в 1-м экзоне гена *PUS3*. Впервые получены данные о распространении выявленных нами у пациентов с интеллектуальной недостаточностью аллельных вариантов с.1475G>A (rs11768549) и с.1891G>A (новая мутация) гена *EPHA1*, с.212A>G (новая мутация) гена *PUS3* и с.806_808CAT>TGTGCA (rs386809049) гена *ZNF527* в популяционной выборке из Украины. Получены модели по гомологии пространственных структур нормальных и мутантных белковых продуктов генов *EPHA1* и *PUS3*. Впервые получены доказательства в пользу того, что мутации с.1475G>A и с.1891G>A в гене *EPHA1* являются генетическими факторами патогенеза интеллектуальной недостаточности моногенной природы. Разработаны методики детекции мутантных вариантов с.1475G>A и с.1891G>A гена *EPHA1*, а также мононуклеотидных замен rs11767557 и rs11771145 в гене *EPHA1* и rs929271 в гене *LIF*, пригодные для генетического тестирования в группах с интеллектуальной недостаточностью. Установлено, что носительство аллельных вариантов 1475A гена *EPHA1* и 4524G (rs929271) гена *LIF* является фактором наследственной предрасположенности к развитию интеллектуальной недостаточности. Показано, что носительство аллельных вариантов 3022C (rs11767557) и 1399A (rs11771145) гена *EPHA1* обуславливает снижение относительного риска развития интеллектуальной недостаточности.

Ключевые слова: мутація, алельний поліморфізм, мононуклеотидна заміна, ген *EPHA1*, ген *LIF*, інтелектуальна недостаточность.

SUMMARY

Gulkovskyi R.V. Gene mutations in children with intellectual disability. – Manuscript.

Thesis for Philosophy Doctor (PhD) degree in Biology, specialty 03.00.22 – molecular genetics. – Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine, Kyiv, 2016.

The study presented in the thesis aimed to clarify the role of *EPHA1*, *PUS3*, *ZNF527*, *SCEL*, *C6orf223* genes mutations, identified by exome sequencing of two affected siblings with intellectual disability (ID), and *LIF* gene polymorphism as genetic factors of ID pathogenesis.

The results of Sanger sequencing conducted for two affected brothers and their healthy non-consanguineous parents and following family segregation analysis of identified mutations with the pathogenesis of ID revealed possible candidate targets for the study. Substitution c.524 C> T (rs141416270) and insertion c.1135_1136insA (novel) in *SCEL* gene as well as SNPs c.585G>T (rs146391418) and c.586C>T (rs138503303) in *C6orf223* gene were excluded from the list of potential candidates involved in the pathogenesis of ID, because it turned out that one of the healthy parents has the same genotype on mutations in genes *SCEL* and *C6orf223* as affected brothers.

The results of site orthologs analyses for ZNF527, PUS3 and EPHA1 proteins revealed that amino acid positions that are substituted due to the investigated mutations are conserved, indicating that there may be an evolutionarily important function.

As the next step in the evaluation of the *ZNF527* (c.806_808CAT>TGTGCA), *PUS3* (c.212A>G) and *EPHA1* (c.1475G>A and c.1891G>A) genes involvement in ID pathogenesis the comparative analysis of genotypes distribution was performed in the group of 65 patients with mild ID (IQ 50-70) and group of healthy volunteers (n=250) from Ukraine.

The *ZNF527* TGTGCA allele frequencies in the group of healthy volunteers (0,168) and in ID patients group (0,154) were no significantly different. Moreover, such a high minor allele frequency allows us to suggest that analyzed polymorphism rs386809049 in *ZNF527* gene cannot be the major cause of ID.

The mutation c.212A>G in *PUS3* gene results in substitution of Tyrosine 71 to Cysteine located in the catalytic domain of human PUS3. The PUS3 catalytic domain 3D structure was modeled using crystal structures of the human PUS1 catalytic domain as a template. It allowed us to discover that substitution p.71Y>C is located in close proximity (~10Å) to active site of PUS3. The possible role of this substitution in changes in catalytic domain conformational flexibility and spatial organization may be supposed. Nevertheless, we still cannot suppose that p.71Y>C mutation can result in change of pseudouridylate synthase efficiency. We did not identified *PUS3* gene c.212A>G substitution in any of 65 investigated ID patients. In turn, the only one heterozygous carrier for the c.212A>G substitution was identified in group of healthy volunteers from Ukraine. Low frequency of 212G allele (0,002) was shown in the population of Ukraine. Thus, the association of *PUS3* c.212A>G with ID was no revealed.

To define if the nonsynonymous mutations c.1475G>A (rs11768549) and c.1891G>A (novel) in *EPHA1* gene had an influence on the 3D structure of EPHA1 protein we modeled the structure of the corresponding EPHA1 domains using Swiss Model server. The SNP c.1475G>A in *EPHA1* gene results in substitution of positive charged Arg492 to uncharged Gln492 located in the fibronectin type III repeat of EPHA1 ectodomain involved in signal transduction and binding with ligands or protein-partners. The SNP c.1891G>A in *EPHA1* gene results in Gly631 to Arg631 substitution located in the glycine-rich region of EPHA1 tyrosine kinase domain, responsible for ATP binding. Supposedly c.1475G>A and c.1891G>A mutations may directly cause changes in EPHA1 domains conformational flexibility and functional efficiency resulting in impaired EPH-

signaling. Importantly, that 1891G>A substitution was not detected in any of investigated 250 persons from control group as well as among 65 ID patients. Moreover, 1891G>A substitution was not identified in any population investigated in frame of Exome Aggregation Consortium project (URL: <http://exac.broadinstitute.org>), International HapMap Project (URL: <http://hapmap.ncbi.nlm.nih.gov/>) and 1000 Genomes Project (URL: <http://browser.1000genomes.org/>). Allelic frequency of another c.1475G>A substitution was: G – 0,99 and C – 0,01 in control group; G – 0,962 and C – 0,038 in ID patients. The significant differences ($p < 0,05$) for allele frequencies of *EPHA1* c.1475G>A SNP between ID patients and the controls were identified. Statistical analysis showed association of minor 1475-A allele (OR=3,96; 95% CI=1,13-13,89; $p=0,02$) with higher risk of mild ID development. Based on genetic analysis, bioinformatics and protein modeling results we propose *EPHA1* gene mutations c.1475G>A and c.1891G>A as a cause of an autosomal recessive non-syndromic form of ID in investigated family.

The next task of the study was to examine the role of *EPHA1* (rs11767557, rs11771145) and *LIF* (rs929271) gene polymorphic variants role as factors of genetic susceptibility to mild ID development. Assays for the detection of *EPHA1* gene rs11767557, rs11771145 and *LIF* gene rs929271 SNPs based on PCR followed by RFLP analysis were developed. The comparative analysis of genotypes distribution was performed in the group of 65 patients with mild ID and group of 250 healthy volunteers from Ukraine (as a control group).

Our results suggest that *LIF* gene polymorphism rs929271 is associated with mild ID. It was determined that total frequency of hetero- and homozygous carriers of *LIF* gene rs929271 minor allele, in this case that of guanine (G), is reliably higher ($p=0,01$) in the ID patients (71,9%) compared to the control group (55%). The minor G-allele occurred less frequently 0,328 – in the control group than 0,445 – in the ID patients ($p=0,02$). It was shown that the risk of mild ID development is increased for both hetero- and homozygous carriers of minor rs929271-G allele and the odd ratio was 2,09 (95% CI: 1.14-3,81). Our results also suggest that SNPs rs11767557 and rs11771145 in the *EPHA1* gene are associated with mild ID. Statistical analysis showed significant association ($p < 0,05$) of wild-type rs11767557-T (OR=1.99; 95% CI=1,18-3,37) and rs11771145-G (OR=1,55; 95% CI=1,02-2,37) alleles with higher risk of mild ID development. Therefore, we propose *LIF* gene rs929271 and *EPHA1* gene rs11768549 (c.1475G>A), rs11767557, rs11771145 polymorphisms as a new markers of genetic susceptibility for ID.

Keywords: mutation, allelic polymorphism, SNP, *EPHA1* gene, *LIF* gene, intellectual disability.