

Міністерство освіти і науки України
Департамент науки і освіти Харківської обласної державної адміністрації
Комунальний заклад «Харківська обласна Мала академія наук
Харківської обласної ради»

Відділення хімії та біології

Секція: загальна біологія

***IN SILICO* АНАЛІЗ ПАТОГЕННИХ ТА ЙМОВІРНО
ПАТОГЕННИХ ОДНОНУКЛЕОТИДНИХ ВАРІАНТІВ
ГОМЕОЗИСНИХ ГЕНІВ**

Роботу виконала:

Розуменко Софія Олексіївна, учениця
10 класу Комунального закладу
«Харківський ліцей № 47 Харківської
міської ради», вихованка
Комунального закладу «Харківська
обласна Мала академія наук
Харківської обласної ради»

Науковий керівник:

Садовниченко Юрій Олександрович,
доцент кафедри медичної біології
Харківського національного медичного
університету, кандидат біологічних
наук, Заслужений працівник освіти
України

IN SILICO АНАЛІЗ ПАТОГЕННИХ ТА ЙМОВІРНО ПАТОГЕННИХ ОДНОНУКЛЕОТИДНИХ ВАРІАНТІВ ГОМЕОЗИСНИХ ГЕНІВ

Розуменко Софія Олексіївна; Комунальний заклад «Харківська обласна Мала академія наук Харківської обласної ради»; Комунальний заклад «Харківський ліцей № 47 Харківської міської ради»; 10 клас; м. Харків;

Садовниченко Юрій Олександрович; доцент кафедри медичної біології Харківського національного медичного університету, кандидат біологічних наук, Заслужений працівник освіти України

У процесі ембріогенезу тварин одну з провідних ролей відіграють гомеозисні гени, кодують ключові транскрипційні фактори. Мутації цих генів спричиняють різні форми раку, порушення опорно-рухової та нервової систем, органів чуття тощо. Відомі сотні однонуклеотидних варіантів гомеозисних генів, однак їх систематичному аналізу та встановленню клінічної значущості досі приділялося мало уваги.

Метою роботи був пошук патогенних та ймовірно патогенних однонуклеотидних варіантів гомеозисних генів *HOXA13*, *HOXB13* та *HOXD13 in silico*.

В роботі використано анотовані послідовності генів *HOXA13*, *HOXB13* та *HOXD13*, кодованих ними транскрипційних факторів НохА-13, НохВ-13 та НохD-13, депонованих у базах NCBI, дані про варіанти генів з баз NCBI та Uniprot, а також біоінформатичні інструменти fathmm, PANTHER, PhD-SNP, PolyPhen-2, SIFT та SNPs&Go.

Встановлено, що серед досліджених гомеозисних генів у біоінформатичних базах dbSNP та Uniprot депоновано найбільше однонуклеотидних варіантів *HOXB13*, 82,4 % з яких віднесені до патогенних.

Визначено, що найбільше SNV гомеобокс-генів за допомогою біоінформатичних інструментів також визначається для гена *HOXB13*. Максимальну кількість варіантів дозволяють проаналізувати web-сервіси PANTHER та PhD-SNP, а найменшу – SIFT.

Найбільше патогенних та умовно патогенних варіантів *in silico* виявлено в гені *HOXB13* – 107 та 206 відповідно.

В генах *HOXA13*, *HOXB13* та *HOXD13* встановлено 162 патогенних SNV, які не зазначені в біоінформатичних базах, найбільше їх було в гені *HOXD13*.

Густина SNV була найвищою в гені *HOXB13*, при цьому розподіл однонуклеотидних варіантів по гомеозисних генах був нерівномірним.

Ключові слова: ембріональний розвиток, гомеозисні гени, біоінформатичний аналіз, однонуклеотидні варіанти

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. ОГЛЯД ЛІТЕРАТУРИ	6
1.1. Генетичний контроль ембріонального розвитку людини.	6
1.2. Роль гомеозисних генів у ембріональному розвитку людини.	13
1.3. Роль <i>HOX</i> -генів у патології людини.	15
РОЗДІЛ 2. ОБ'ЄКТИ ТА МЕТОДИ ДОСЛІДЖЕНЬ	17
2.1. Матеріали дослідження.	17
2.2. Біоінформатичний аналіз.	18
РОЗДІЛ 3. АНАЛІЗ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ	20
3.1. Аналіз патогенних та умовно патогенних варіантів гена <i>HOXA13</i> <i>in silico</i>	20
3.1. Аналіз патогенних та умовно патогенних варіантів гена <i>HOXB13</i> <i>in silico</i>	24
3.1. Аналіз патогенних та умовно патогенних варіантів гена <i>HOXD13</i> <i>in silico</i>	27
ВИСНОВКИ	31
СПИСОК ВИКОРИСТАНИХ ЛІТЕРАТУРНИХ ДЖЕРЕЛ	32

ВСТУП

Актуальність теми. Ембріональний розвиток є ретельно узгодженим процесом, який потребує максимально точного генетичного контролю [2, 20]. Одну з провідних ролей в процесі ембріогенезу тварин відіграють гомеозисні гени, або гомеобокс-гени, які кодують ключові транскрипційні фактори [11]. В людини гомеозисні гени зібрані у чотири кластери – *HOXA*, *HOXB*, *HOXC* та *HOXD* [11]. Мутації цих генів спричинюють різні форми раку, порушення опорно-рухової та нервової систем, органів чуття тощо, особливо це стосується генів *HOXA13*, *HOXB13* та *HOXD13* [18, 19]. Однак хоча біоінформатичні ресурси містять сотні варіантів цих генів, до патогенних віднесено небагато з них, а систематичному аналізу генів приділялося мало уваги. Отже, передбачення клінічної значущості однонуклеотидних варіантів (SNV) генів *HOXA13*, *HOXB13* та *HOXD13* потребує подальшого дослідження.

Метою роботи був пошук патогенних та ймовірно патогенних однонуклеотидних варіантів гомеозисних генів *HOXA13*, *HOXB13* та *HOXD13* *in silico*.

Передбачалося виконання таких завдань:

1. Вивчити різноманітність однонуклеотидних варіантів гомеозисних генів у біоінформатичних базах.

2. Проаналізувати SNV генів *HOXA13*, *HOXB13* та *HOXD13* за допомогою біоінформатичних інструментів fathmm, PANTHER, PhD-SNP, PolyPhen-2, SIFT та SNPs&Go.

3. Оцінити патогенність та умовну патогенність досліджених варіантів гомеозисних генів.

4. Визначити SNV генів *HOXA13*, *HOXB13* та *HOXD13*, патогенність яких не зазначена в біоінформатичних базах.

5. Дослідити густину SNV у гомеозисних генах.

Об'єкт дослідження: гомеозисні гени *HOXA13*, *HOXB13* та *HOXD13*.

Предмет дослідження: патогенні та умовно патогенні однонуклеотидні варіанти гомеозисних генів.

Методи дослідження: у роботі використано біоінформатичні методи визначення патогенності однонуклеотидних варіантів генів, статистичні методи аналізу отриманих даних.

Наукова новизна одержаних результатів. Уперше проведено комплексний біоінформатичний аналіз однонуклеотидних варіантів генів *HOXA13* та *HOXB13*. Встановлено 162 SNV, які не зазначені в біоінформатичних базах як патогенні.

Рекомендації щодо використання одержаних результатів. Доцільно проводити прицільне секвенування гомеозисних генів у хворих на різні форми раку, долонно-підшовний генітальний синдром, синдром Гуттмахера, синполідактилії, брахідактилії тощо для встановлення асоціацій. Матеріали науково-дослідницької роботи можуть бути використані для одержання навчального матеріалу на уроках біології, заняттях гуртків МАН України.

Особистий внесок автора полягає у постановці завдань, збиранні матеріалу для досліджень, проведенні біоінформатичного аналізу, аналізі й узагальненні результатів, формулюванні висновків та написанні роботи.

РОЗДІЛ 1

ОГЛЯД ЛІТЕРАТУРИ

1.1. Генетичний контроль ембріонального розвитку людини

Ембріональний розвиток є ретельно узгодженим процесом, який потребує максимально точного генетичного контролю. Незважаючи на існування певних особливостей ембріогенезу в різних видів, для людини він був вивчений здебільшого на тваринних моделях [2, 20].

Перші декілька діб розвитку зародку людини процес його дроблення контролюється продуктами генів материнського ефекту [20]. З третьої-четвертої діб після запліднення процес ущільнення бластомерів у морулу вже забезпечується продуктами генів зародку, зокрема кадгеринами (*Cdh1*). Підтримка плюрипотентності дочірніх клітин забезпечується експресією генів гомеодомених транскрипційних факторів *POU5F1/OCT4* та *NANOG*, однак якщо в ранніх бластомерах вона фіксується в усіх клітинах, то згодом – лише в клітинах внутрішньої клітинної маси (ВКМ). Диференціювання бластомерів у клітини трофобласту потребує спочатку синтезу транскрипційних факторів *Cdx2* та *Tead4*, а потім – *Eomes* та *Elf5*. Вказані транскрипційні фактори є компонентами складної генної мережі регуляції, зокрема *Yap1* є кофактором *Tead4*, який за відсутності фосфорилування локалізується в ядрах зовнішніх клітин, а зв'язування *Yap1* з *Tead4* в цих клітинах спричинює експресію гена *CDX2* у зиготі, а пізніше – й інших генів трофобласту, що, в свою чергу, спричинює диференціювання цих клітин в клітини трофоектодерми. Транскрипційний фактор *Oct4* є необхідним для детермінації клітин ВКМ, *Nanog* – для підтримки плюрипотентності і утворення епібласта, а *Sall4* – для диференціювання клітин ВКМ в епібласт. З *Oct4* взаємодіє інший транскрипційний фактор – *Sox2*, який сприяє зв'язуванню першого з його генами-мішенями [20]. Гени *OCT4* і *CDX2* є супресорами один одного, тому зв'язування продукту одного з промотором іншого спричинює

диференціювання клітин у ВКМ або трофобласт [20].

До моменту диференціювання ВКМ на епібласт та гіпобласт *NANOG* експресується виключно у першому, а ген транскрипційного фактора *GATA6* – в гіпобласті. Вважається, що експресія гена *GATA6* індукується активацією шляху трансдукції сигналу RTK–MAPK через фактор росту фібробластів (FGF) та зв'язувальний білок 2 рецептора фактору росту Grb2, що може пригнічувати експресію гена *NANOG*. Експресії генів *GATA6* та *GATA4* сприяє транскрипційний фактор Sox17, що є необхідною умовою для диференціювання ВКМ в гіпобласт. Транскрипція генів раннього ембріогенезу та диференціювання ВКМ в цілому підлягає епігенетичному контролю [20].

Активація сигнального шляху Wnt зумовлює формування в задній частині ембріона первинної смужки. Процес гастрюляції стимулюється членами надродини трансформувальних факторів росту TGF- β , зокрема Nodal. Вплив Nodal та деяких FGF, зокрема FGF8, зумовлює утворення первинної смужки завдяки опусканню клітин епібласту та їх перетворення на клітини мезо- й ентодерми [20].

В процесі гастрюляції також задіяні транскрипційні фактори SNAI1, SNAI2, SNAI3, MESP1, MESP2, EOMES тощо. Експресія генів родини *SNAIL* контролюється факторами росту TGF- β та FGF [20]. SNAI1 репресує ген *CDH1*, що призводить до припинення процесів компактизації в зародку. SNAI2 стимулює епітеліально-мезенхімальний перехід та забезпечує міграцію й підтримання функцій меланобластів [20].

Трансдукція сигналів шляхами Wnt та Notch регулюється за типом негативного зворотного зв'язку, що зумовлює осциляції експресії генів *HES7*, *LNFG* та *DLL3* у клітинах пресомітової мезодерми, та, у певних фазах коливання концентрацій їхніх продуктів, надає цим клітинам компетентності закладати соміти. Транскрипція гена *FGF8* спочатку спостерігається тільки в клітинах хвостової бруньки, але надалі вона поширюється й на інші частини зародку, щоправда при цьому виникає задньо-передій градієнт концентрації FGF8. Крім того, в сомітах фіксується експресія гена синтезу одного з факторів

росту – ретиноєвої кислоти (*ALDH1A2/RALDH2*). Збільшення концентрації ретиноєвої кислоти активує специфічні гени сомітогенезу, зокрема *MESP1* та *MESP2*, тоді як *FGF8* є інгібітором цього процесу [20].

Встановлено, що цей концентраційний градієнт забезпечує транспорт везикул з морфогенами, зокрема з Sonic hedgehog (Shh) та ретиноєвою кислотою в лівий бік. Цей потік морфогенів сприймається нерухомими сенсорними війками і спричинює підвищення внутрішньоклітинного кальцію в клітинах з лівого боку бруньки. Секреція везикул, вивільнення Shh та ретиноєвої кислоти, а також надходження йонів кальцію в клітини з лівого боку опосередковані сигнальним шляхом Fgf. Гени, які експресувалися тільки в лівій частині бруньки, асиметрично експресуються і в мезодермі лівої латеральної пластинки, що призводить до асиметричної експресії генів, зумовлюючи ліво-правосторонню асиметрію [20]. На ранніх стадіях розвитку ембріонів експресія гена *Gsc* (*Gooseoid*) була зафіксована у гензенівському вузлику та краніальній частині первинної смужки, на пізніх стадіях вона була обмежена прехордальною пластинкою. Аналогічно, експресія *Nodal* у первинному вузлику була тимчасовою і зникла до формування хордальної пластинки. Однак згодом експресія у вузлику була виявлена по обидва боки задньої частини хордальної пластинки, і коли ембріон перебував на стадії 2-3 соміту, експресія у вузлику ставала асиметричною і концентрувалася в мезодермі лівої латеральної пластинки. Ймовірно, що в задній частині хорди відбувається опосередкований війками рух морфогенів від вузлика для встановлення ліво-правосторонньої асиметрії у ссавців. Крім того, не виключено, що поздовжня вісь тіла визначаються за допомогою «активності організатора» первинного вузлика [20].

Динеїнові ручки єдиних війок клітин первинного вузлика людини містять ліво-правий білок динеїн, що кодується геном *DNAH11*, який є гомологом *Lrd* мишей (ліво-правий динеїн), та інших динеїнових білків. Для правильної збірки та функціонування динеїну потрібні білки, що кодуються генами *KTU*, *CCDC39* та *CCDC40* [20].

В мишей та деяких інших тваринних моделях потік від вузлика призводить до того, що *Nodal* асиметрично експресується в мезодермі лівої бічної пластинки. Потім *Nodal* індукує експресію *Lefty2* та *Pitx2* в мезодермі лівої латеральної пластинки, а також *Lefty1* в лівій нижній частині нервової трубки. У цих тваринних моделях *Lefty1* обмежує експресію *Nodal* обмеженою областю, тоді як *Lefty2* запобігає поширенню сигналу *Nodal* у праву частину зародку; *Pitx2* бере участь у підтримці ліво-правосторонньої асиметрії під час органогенезу [20].

Багато генів, які експресуються на ранніх стадіях ембріонального розвитку, беруть участь і в інших процесах розвитку на пізніших стадіях ембріонального і внутрішньоутробного розвитку. Крім того, між деякими з цих генів існує певна функціональна надлишковість, яка запобігає загибелі зародків внаслідок мутацій в одному чи більше з цих генів. Наприклад, хоча однією з головних функцій *SOX2* є індукція плюрипотентності, він також бере участь у підтримці нейрональних стовбурових клітин і, отже, нейрогенезі, розвитку зорового келиха, генерації периферичних сенсорних та вегетативних нейронів через взаємодію з сигнальним шляхом Notch [20]. Нещодавно було показано, що *Sall1* підвищує експресію *Nanog* в ембріональних стовбурових клітинах мишей і фізично взаємодіє як з *Nanog*, так і з *Sox2*, а також пригнічує диференціювання ектодерми і мезодерми. Ці дані свідчать про те, що *Sall1* бере безпосередню участь у регуляції плюрипотентності та диференціюванні тканин на ранніх етапах ембріонального розвитку [20].

У міру дозрівання бластоцисти та формування в неї усіх необхідних для імплантації факторів, на їхніх поверхнях утворюється білок перлекан (HSPG2). Функції HSPG2 включають адгезію клітин, підтримання базальної мембрани та хряща, а також закріплення ацетилхолінестерази в нервово-м'язових синапсах, втрата якого може бути причиною міотонії [20].

Експресія гена сімейства *TGF- β* – *BMP-4*, який кодує кістковий морфогенетичний білок-4, у ектодермі та мезодермі ембріона під час гастрюляції спричинює перетворення ектодерми на епідерміс, а також

диференціювання мезодерми на проміжну мезодерму та мезодерму бічної пластинки. Секреція білків ногіну, хордину та фолістатину в первинному вузлику, хорді та прехордальній мезодермі інактивує BMP-4 та забезпечує нейтралізацію ектодерми й викликає утворення хорди та приосьової мезодерми з вихідної мезодерми [2].

Однак ці нейральні індуктори здатні викликати формування лише тканини переднього і середнього мозку. Індукція структур каудальної частини нервової пластинки (задній і середній мозок) залежить від секреції двох білків – Wnt-3a та FGF. Окрім того, у становленні краніо-каудальної осі певну роль відіграє ретиноева кислота, яка викликає перетворення краніальних сегментів на каудальні шляхом регуляції експресії генів гомеобоксу [2].

Сигнали стосовно диференціювання сомітів надходять від прилеглих структур, у тому числі хорди, нервової трубки, епідермісу і меходерми бічної пластинки. Білок Shh виробляється хордою та нижньою частиною нервової трубки та індуктує перетворення вентромедіальної частини соміту на склеротом. Утворившись, клітини склеротома виділяють фактор транскрипції PAX1, який ініціює каскад хрящово- та кістковотвірних генів до формування хребта. Експресія фактора PAX3, який регулюється білками Wnt з дорзальної частини нервової трубки, детермінує дермоміотропний регіон соміту. Білки Wnt того самого походження діють також на дорзомедіальну частину соміту, стимулюючи її до експресії м'язово-специфічного гена *Myf5* і до перетворення на епаксіальну мускулатуру. Взаємодія між білком-інгібітором BMP-4 та, можливо, FGF, а також мезодермою бічної пластинки й активуючими факторами Wnt з епідермісу стимулює дорзолатеральну частину соміту до експресії іншого м'язово-специфічного гена, *MyoD*, та формування гіпаксіальних м'язів кінцівок і стінки тіла. Середній фрагмент дорзальної частини соміту індукується до утворення дерми під впливом нефротрофіну 3 (NT-3), який продукується дорзальною ділянкою нервової трубки [2].

Зачатки кінцівок з'являються наприкінці четвертого тижня розвитку у формі виступів вентролатеральних відділів тіла ембріона. Вони утворені

мезенхімою, що походить із соматичної мезодерми бічної пластинки, і вкриті шаром кубічних клітин ектодерми [2]. Значну роль у розвитку кінцівок відіграють гомеозисні HOX-гени [2, 11, 18].

По завершенні детермінації розміщення кінцівок уздовж краніокаудальної осі постає необхідність регуляції їхнього росту вздовж проксимодистально, передньо-задньої і дорзовентральної осей. Початок росту кінцівки ініціюється *FGF8*. З початком росту ген *RFNG* (гомолог гена *Drosophila fringe*) індукує утворення в дорзальній ділянці ектодерми кінцівки апікального ектодермального гребеня. Цей ген індукує експресію *Ser-2* гомолога *Drosophila serrate*, на межі між клітинами, у яких здійснюється експресія *RFNG*, і клітинами, в яких цей ген не експресується. Це місце відповідає ділянці закладки апікального ектодермального гребеня. Регуляція утворення апікального ектодермального гребеня відбувається за участю гена *EN1*, експресія якого здійснюється в ектодермальних клітинах вентральної ділянки, оскільки цей ген пригнічує експресію *RFNG*. Після формування гребеня в ньому починається експресія *FGF4* та *FGF8* — факторів, необхідних для виникнення зони розвитку, що утворена популяцією клітин мезенхіми, що інтенсивно діляться, суміжних з апікальним ектодермальним гребенем. Ріст кінцівки в дистальному напрямку є наслідком інтенсивної проліферації цих клітин під впливом факторів росту фібробластів. Упродовж росту мезенхімні клітини проксимального кінця зони розвитку щоразу більше віддаляються від гребеня і його впливу, сповільнюють інтенсивність поділу та вступають у процес диференціювання [2].

Характер структур, розміщених навколо передньо-задньої осі кінцівки, регулюється зоною поляризуючої активності, що являє собою кластер клітин заднього краю кінцівки поблизу від тулуба. Ці клітини продукують ретиноеву кислоту, що ініціює продукцію білків *shh* (*sonic hedgehog*) — секреторного фактора, який регулює становлення передньо-задньої осі. Зокрема, під його впливом пальці розміщуються у правильному порядку, а великий палець займає крайнє положення з променевої (передньої) сторони. В процесі росту кінцівки

зона поляризуючої активності просувається в дистальному напрямку, залишаючись поблизу заднього краю апікального ектодермального гребеня. Порушення утворення ретиноївої кислоти або *shh* на передньому краю кінцівки при нормальному розвитку зони поляризуючої активності на задньому її краю зумовлює дзеркальне подвоєння структур кінцівки [2].

Розвиток уздовж дорзовентральної осі відбувається під впливом клітин ектодерми кінцівки. Сигнали від цих клітин передаються через *Wnt7a* — секреторний фактор, що індукує експресію в дорзальній мезенхімі транскрипційного фактора *Lmx1*, який містить гомеодомен. *Lmx1* забезпечує дорзальну локалізацію клітин, забезпечуючи становлення дорзовентральної осі. Окрім того, *Wnt7a* сприяє експресії *shh* у зоні поляризуючої активності і цим опосередковано впливає на становлення передньозадньої осі. Загалом усі гени, що визначають розвиток кінцівок, регулюються за принципом зворотного зв'язку. Так, *FGF* апікального епітеліального гребеня активізують експресію *SHH* в зоні поляризуючої активності, тоді як *Wnt7a* підтримує дію *shh* [2].

Інші чинники також відіграють певну роль у розвитку кінцівок. Одна з груп таких речовин — це ізоформи білків формінів, що в нормі виявляються в мезенхімі і апікальному епітеліальному гребені. Без цих білків апікальний епітеліальний гребінь дезорганізується внаслідок відсутності *FGF4*, що, в свою чергу, призводить до втрати регуляторного впливу *Shh*. Внаслідок цього припиняється розвиток кінцівки вздовж передньо-задньої осі. Отже, для нормального росту і диференціювання кінцівок істотним є цілий каскад молекулярних механізмів, що взаємодіють [2].

Провідну роль в формуванні міотому та м'язів відіграють гени *BMP4*, *MyoD*, *Myf5*, і, можливо, *FGF*, у взаємодії з Wnt-білками [2].

Генетична регуляція формування внутрішніх органів вивчена ще не повністю. Розвиток серця пов'язаний з експресією генів *Nkx2-5/CSX*, *TBX5*, *BMP2*, *WNT* тощо. Формування нирки потребує скоординованої транскрипції генів *WT1*, *GDNF*, *RET*, *MET*, *FGF2*, *BMP7*, *PAX2*, *WNT4* тощо. Утворення яєчок потребує експресії гена *SRY* [2].

Регуляція розвитку ока пов'язана з транскрипцією генів *PAX6*, *SHH*, *PAX2*, *BMP7* тощо. У диференціювання спинного мозку щонайменше залучені гени *PAX3*, *PAX*, *MSX1*, *MSX2*, *SHH*, *BMP4* та *BMP7*. Формування головного мозку потребує експресії гомеозисних *HOX*-генів, *LIM1*, *OLX2*, *FGF8*, *BF1*, *SHH*, *BMP4* та *BMP7* [2].

1.2. Роль гомеозисних генів у ембріональному розвитку людини

Одну з провідних ролей в процесі ембріогенезу тварин відіграють гомеозисні гени, або гомеобокс-гени, відмітною ознакою яких є наявність консервативної послідовності з 180 пар нуклеотидів – гомеобоксу, яка кодує особливу ділянку білка – гомеобокс, що відповідає за зв'язування з певними послідовностями ДНК [11]. Вони кодують транскрипційні фактори, що активують каскади генетичних регуляторних ферментів, таких як сегментація і формування осей тіла. Багато генів гомеобоксу зібрано в гомеотичні кластери, хоча інші гени також містять гомеодомени. Важливим генетичним кластером, відповідальним за становлення краніокаудальної осі, є гомеотичний генний комплекс дрозофіли *Hom-C*. Його гени, що належать до класів гомеотичних генів *Antennapedia* та *Bithorax*, організовані у вигляді функціональної одиниці на одній хромосомі. Так, гени, що зумовлюють розвиток краніальніших структур, локалізовані на 3'-кінці ДНК і експресуються першими, у той час як гени, що контролюють розвиток каудальніших структур, проявляються пізніше і лежать ближче до 5'-кінця. Ці гени, збережені в людини, існують у вигляді чотирьох копій – *HOXA*, *HOXB*, *HOXC* та *HOXD*, організація і експресія яких подібна до такої в дрозофіли. Так, кожен кластер лежить на окремій хромосомі, а гени у кожній групі нумеруються від 1 до 13. Гени, що мають той самий номер, але належать до різних груп, утворюють паралогусні групи, такі як *HOXA4*, *HOXB4*, *HOXC4* та *HOXD4*. Характер експресії цих генів, судячи з результатів досліджень мишей, у яких вилучено один або кілька цих генів, дозволяє підтримати гіпотезу про їхню роль у краніокаудальній організації

похідних усіх трьох зародкових листків. Наприклад, характер транскрипції гена *HOXB* у задньому мозку відповідає її сегментарній природі та її поділу на ромбомери. Так, за винятком *HOXB1*, гени *HOXB*-групи переднішої (3') локалізації експресуються з частковим перекриванням у передніших ромбомерах (сегментах). Крім того, такого типу кодування дотримуються мігруючі клітини нервового гребеня, які несуть його до горлових дуг. Подібний *HOX*-код із частковим перекриванням існує в сомітах і хребцях, де гени кожного кластеру розташовані ближче до 3'-кінця, експресуються у краніальніших сегментах. Достеменно невідомо, яким чином код передається для контролю диференціювання дуг, але в його реалізації можуть бути залучені *Shh* та ретиноїди [2, 11].

Позиції зачатків кінцівок уздовж на поверхні тіла зародку також визначаються родиною *HOX*-генів — регуляторів ембріогенезу, які діють через транскрипційні фактори *Tbx5* для верхньої кінцівки та *Tbx4* — для нижньої [2]. Експресія *HOX*-генів здійснюється вздовж цієї осі і перекривається від голови до хвоста [2, 11, 18]. При цьому експресія деяких з них обмежена в краніальному напрямку більше, ніж інших. Наприклад, краніальна межа експресії гена *HOXB8* сягає краніального краю верхньої кінцівки і неправильна експресія цього гена призводить до зміщення цих кінцівок [2, 18].

Типи і розміри кісток кінцівки значною мірою регулюються гомеозисні *HOX*-генами. Так, експресія *HOX*-гена, що відбувається внаслідок комбінованої експресії *shh*, *FGF* та *Wnt7a*, реалізується у фазовій послідовності, що відповідає формуванню трьох частин кінцівки: проксимальної, що містить плечову або стегнову кістку; середньої, що містить променево і ліктьову кістки або кістки гомілки; дистальної, що містить кістки кисті або стопи. Гени кластерів *HOXA* і *HOXD* є основними детермінантами розвитку кінцівок, і зміни в їхній коекспресії можуть бути причиною відмінностей в будові верхньої та нижньої кінцівок. Подібно до того, як це відбувається при становленні краніокаудальної осі зародка, *HOX*-гени, перекриваючи один одного, здійснюють регуляцію розміщення кінцівок [2].

Значну роль *HOX*-гени відіграють й у процесах формування головного мозку [2].

1.3. Роль *HOX*-генів у патології людини

Встановлено, що мутації щонайменше 10 *HOX*-генів у людини можуть спричинювати патології (*HOXA1*, *HOXA2*, *HOXA11*, *HOXA13*, *HOXB1*, *HOXB13*, *HOXC13*, *HOXD4*, *HOXD10* та *HOXD13*) [19].

Мутації гена *HOXA1* спричинюють важкі розлади нервової системи, зокрема Bosley–Salih–Alorainy syndrome та Athabaskan brainstem dysgenesis syndrome [19].

В декількох осіб з варіантами гена *HOXA2* було діагностовано мікродію [19].

Делеція в гені *HOXA11* асоційована з радіоульнарним синостозом та тромбоцитопенією [19].

Гетерозиготність за мутацією гена *HOXA13* зумовлювала долонно-підшовний генітальний синдром та синдром Гуттмахера [19]. Внаслідок надмірної експресії цього гена може розвинутися стравохід Баррета, а пізніше – рак стравохода [14].

Рецесивні мутації гена *HOXB1* спричинюють вроджений параліч лицьового нерва, зниження слуху, косоокість, випинання середньої частини обличчя, закинутий ніс, ускладнення живлення, затримку мови, деформацію вушних раковин, порушення формування лицьового нерва та завитки [19].

Варіанти гена *HOXB13* асоційовані з раннім раком передміхурової залози, грудей та колоректальним раком [6, 19].

Мутації гена *HOXC13* спричинюють ектодермальну дисплазію типу 9 з переважним ураженням волосся та нігтів, а також рак шийки матки [12, 19].

Заміна нуклеотидів у гені *HOXD4*, ймовірно, зумовлює гостру лімфобластну лейкемію та порушення опорно-рухової системи [19].

Міссенс-мутації гена *HOXD10* асоційовані з порушеннями стоп та

хворобою Шарко-Марі-Тута [19].

Варіанти гена *HOXD13* зумовлюють синполідактилію типу II, брахідактилію типу D, синдактилію типу V, рак шийки матки, та, ймовірно, синдром VACTERL (дефекти хребців, анальна атрезія, вади серця, трахео-стравохідний свищ, ниркові аномалії та аномалії кінцівок) [10, 12, 19].

Більшість з наведених вище асоціацій були встановлені лише для декількох пацієнтів, тоді як виявлені для генів *HOXA13*, *HOXB13*, *HOXD13* є поширенішими і становлять значний клінічний інтерес.

РОЗДІЛ 2

ОБ'ЄКТИ ТА МЕТОДИ ДОСЛІДЖЕНЬ

Об'єктами дослідження були гомеозисні гени людини *HOXA13*, *HOXB13*, *HOXD13*. Дослідження проводили *in silico* у вересні-листопаді 2023 р. Проводили біоінформатичне дослідження патогенних та ймовірно патогенних однонуклеотидних варіантів цих генів.

2.1. Матеріали дослідження

Матеріалами дослідження слугували анотовані послідовності гомеобоксних білків людини НохА-13, НохВ-13 та НохD-13, а також їхні однонуклеотидні варіанти, депоновані у базах Nucleotide Національного центру біотехнологічної інформації США (<https://www.ncbi.nlm.nih.gov/nucscore>) та Uniprot (<https://www.uniprot.org>).

Ген *HOXA13* розташований в локусі 7q15.2. Він містить 12743 пар нуклеотидів і включає два екзони. Білок НохА-13 складається з 388 амінокислотних залишків [13, 15]. У роботі проаналізовано 355 варіантів гена *HOXD13*, депонованих у базі Uniprot.

Ген *HOXB13* розташований в локусі 17q21.32. Він містить 10987 пар нуклеотидів і включає два екзони. Білок НохВ-13 складається з 284 амінокислотних залишків [7]. У роботі проаналізовано 873 варіанти гена *HOXD13*, депонованих у базі Uniprot.

Ген *HOXD13* розташований в локусі 2q31.1. Він містить 10135 пар нуклеотидів і включає два екзони. Білок НохВ-13 складається з 343 амінокислотних залишків [9]. У роботі проаналізовано 365 варіантів гена *HOXD13*, депонованих у базі Uniprot.

2.2. Біоінформатичний аналіз

Вплив заміни амінокислот у поліпептидному ланцюгу на функціональність білкового продукту гена досліджено за допомогою шести біоінформатичних онлайн-інструментів з різними алгоритмами: fathmm, PANTHER, PhD-SNP, PolyPhen-2, SIFT та SNPs&Go.

Онлайн-сервіс Fathmm передбачає вплив місенс-мутацій на функції білка на основі порівняння збереження консервативних послідовностей у рамках прихованих марковських моделей (ПММ) з «ваговими коефіцієнтами патогенності», що представляють загальну толерантність білка/домену до мутацій. Він визначає варіанти як «damaging» або «tolerated» [21].

Біоінформатичний інструмент PANTHER визначає ймовірний вплив несинонімічних заміни на функції білка з використанням алгоритму PANTHER-PSEP (позиційно-специфічне еволюційне збереження), який ґрунтується на тривалості збереження певної амінокислоти в еволюційній лінії. Він визначає варіанти як «probably damaging» або «probably benign» [22].

Алгоритм визначення ефекту несинонімічних заміни на структуру та функції білка, що лежить в основі онлайн-сервісу PhD-SNP, базується на класифікаторі на основі опорних векторів, навченому та протестованому на послідовності білка та інформації профілю. Він визначає варіанти як «neutral» або «disease» [5].

Біоінформатичний інструмент PolyPhen-2 (Polymorphism Phenotyping v2) оцінює вплив заміни амінокислот на структуру та функції білків людини через прямі фізичні та еволюційні порівняння. Його алгоритм дозволяє визначити варіанти як «probably damaging», «possibly damaging» та «benign» [3].

Онлайн-сервіс SIFT прогнозує зміни структури та функцій білка за несинонімічних заміни на основі гомології послідовностей і фізичних властивостей амінокислот. Він позначає варіанти як «affect protein function» та «tolerated» [16].

Алгоритм визначення ефекту несинонімічних заміни на структуру та

функції білка, що використаний у біоінформатичному інструменті SNPs&GO, аналогічний такому сервісу PhD-SNP, однак враховує також і функціональну інформацію. Він визначає варіанти як «neutral» або «disease» [4, 5, 22].

Висновок про ймовірну патогенність варіантів зроблено за визначення їх як таких трьома або чотирма ресурсами, а про високу ймовірність патогенність – щонайменше п'ятьма.

Густину однонуклеотидних варіантів визначали як їх кількість на 60 нуклеотидів.

РОЗДІЛ 3

АНАЛІЗ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

3.1. Аналіз патогенних та умовно патогенних варіантів гена *HOXA13* *in silico*

У базі dbSNP Національного центру біотехнологічної інформації США (<https://www.ncbi.nlm.nih.gov/snp>) на тепер зареєстровано 3350 однонуклеотидних варіантів (SNV) гена *HOXA13*, 720 з яких припадають на екзони (21,5 %). 473 SNV гена *HOXA13* (65,7 %) належать до міссенс-мутацій. До патогенних та ймовірно патогенних віднесено тільки чотири з них (0,6 %).

У базі Uniprot, яка підтримується Європейським та Швейцарським інститутами біоінформатики (EBI та SIB відповідно), а також ресурсом інформації про білки (PIR), міститься інформація про 344 SNV гена *HOXA13* (<https://www.uniprot.org>), шість з яких віднесено до патогенних (1,7 %).

Найменше варіантів гена *HOXA13* – 244 – було виявлено сервісом SIFT, а найбільше – 343 – сервісом PANTHER, при цьому найменше патогенних та умовно патогенних серед них було визначено сервісом PhD-SNP – 104 (31,1 %), а найбільше – сервісом PolyPhen-2 – 228 (66,9 %) (табл. 3.1), що може бути зумовлене особливостями їхніх алгоритмів [3-5, 16, 21-23].

Таблиця 3.1

Визначення патогенних та умовно патогенних варіантів гена *HOXA13* за допомогою різних біоінформатичних інструментів

SNV	Біоінформатичний інструмент					
	fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
Загальна кількість	334	343	334	341	244	334
Кількість патогенних та умовно патогенних	131	142	104	228	118	214
% патогенних та умовно патогенних	39,2	41,4	31,1	66,9	48,4	64,1

Аналіз результатів дослідження варіантів гена *HOXA13* за допомогою біоінформатичних інструментів показало, що найменше SNV – 22 – були визначені як патогенні чи умовно патогенні усіма шістьма web-сервісами (6,4 %), а найбільше – 66 – тільки чотирма (19,2 %) (рис. 3.1). Таким чином, як патогенні можна оцінити 60 варіантів (17,4 %), як умовно патогенні – 122 (35,5 %) (рис. 3.1, табл. 3.2). При цьому, тільки SNV rs121912542 зазначений як патогенний варіант у базах Uniprot та dbSNP, а 59 SNV не зазначені як такі в жодній з них.

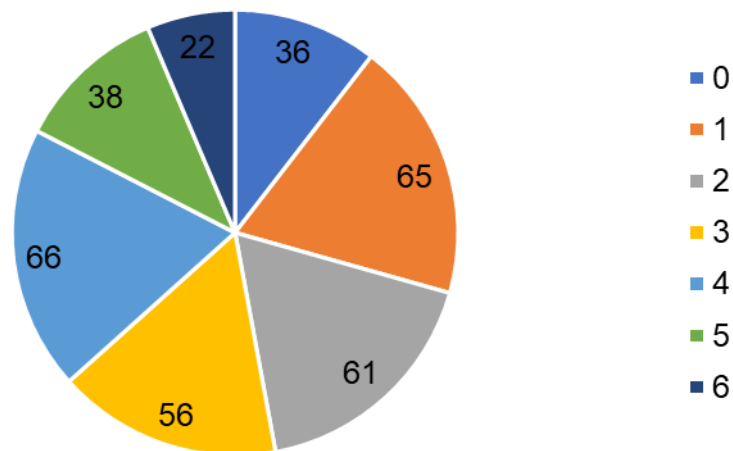


Рис. 3.1. Розподіл варіантів гена *HOXA13* за ступенем патогенності.

Таблиця 3.2

**Варіанти гена *HOXA13*,
не зареєстровані у біоінформатичних базах як патогенні**

1	SNV	Заміна	Біоінформатичний інструмент					
			fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
2	3	4	5	6	7	8	9	
1.	rs1051241147	T2R						
2.	rs1248621784	H8Y						
3.	rs1243248573	R10L						
4.	rs758509164	M17L						
5.	rs1311186012	Y20F						
6.	rs1358459859	D21N						
7.	rs1459803001	G25R						
8.	rs1243158347	M35R						
9.	TCGA novel	G67E						
10.	rs1488684272	G68R						
11.	rs1267410502	N69T						

Продовження таблиці 3.2

	2	3	4	5	6	7	8	9
12.	rs1354582897	A95E						
13.	rs751230129	S164L						
14.	rs920593760	Y171F						
15.	rs1400633388	G175R						
16.	rs1472409629	Y178H						
17.	rs911090483	C181R						
18.	rs761927877	R183H						
19.	rs1296082976	K209E						
20.	rs767443800	R224S						
21.	rs764293809	A225S						
22.	rs763227546	K226E						
23.	rs1423585132	A229D						
24.	rs1245700489	P239H						
25.	TCGA novel	Q244H						
26.	rs879249716	D251G						
27.	rs765319714	H266R						
28.	rs766860410	Q277R						
29.	rs772451529	G284S						
30.	rs745440294	Q288R						
31.	rs1381446788	Y290D						
32.	rs1304497031	C291Y						
33.	rs1393058385	E294K						
34.	rs557028335	Q295P						
35.	TCGA novel	H300Y						
36.	rs1190905313	R320S						
37.	rs1254158162	R321S						
38.	rs1364830866	G322R						
39.	TCGA novel	R323S						
40.	rs1207176698	K324E						
41.	rs776302177	V327M						
42.	rs1404022159	Q333L						
43.	rs867556606	E336K						
44.	rs1157618621	A342T						
45.	rs1365198545	K349E						
46.	rs749486073	R352G						
47.	rs746016880	R354G						
48.	rs746016880	R354W						
49.	rs1249661007	S356P						
50.	rs149615551	T359M						
51.	rs772829383	V366G						
52.	VAR_075341	I368F						
53.	rs1398380500	I368T						
54.	rs1583447648	F370V						
55.	VAR_017775	Q371L						
56.	COSM3880422	R374G						
57.	rs1046132538	R374K						
58.	VAR_075342	V375F						
59.	rs1583447631	V375G						

Вивчення густини замін амінокислот у білку НохА-13 показало, що цей

показник в середньому по гену складав 0,30, був найнижчим у позиціях 81-100, а найвищим – у позиціях 201-220. Максимальна густина замін була у понад 1,7 рази нижчою, ніж у гені *LDLR* [8]. Патогенні заміни були відсутні в позиціях 41-60 та 101-160, а найбільше їх було в позиціях 361-380, тоді як найменше умовно патогенних варіантів було в позиціях 101-120, а найбільше – у позиціях 21-60 та 161-180. Таким чином, як заміни амінокислот в білку в цілому, так і патогенні й умовно патогенні розподілені по гену нерівномірно, що може бути зумовлене функціональними особливостями відповідних ділянок білка. Це саме стосується й SNV, що зумовлюють ці заміни.

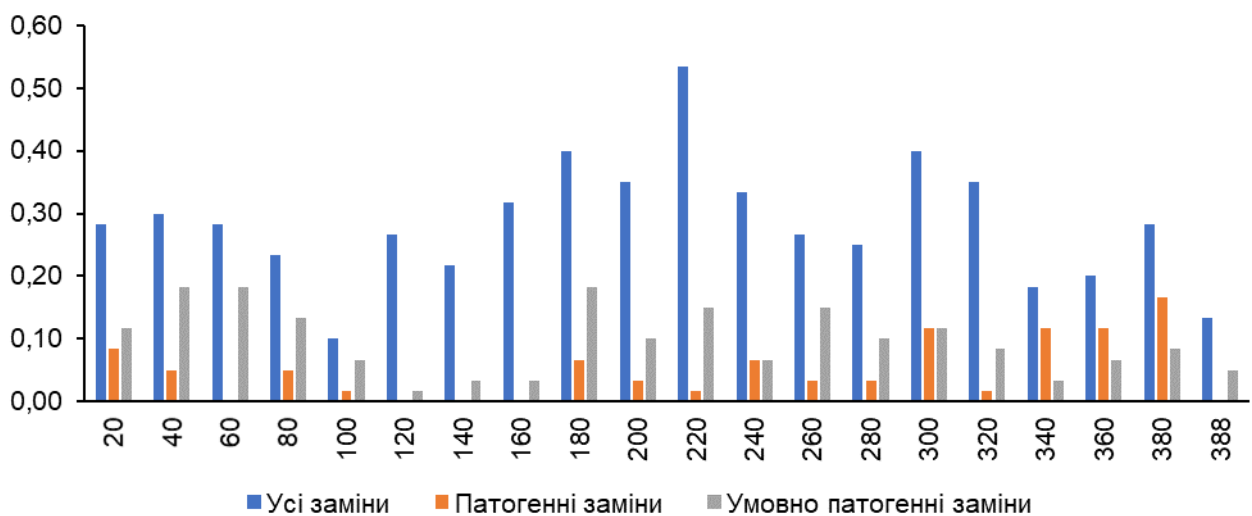


Рис. 3.2. Густина замін амінокислот у білку NoxA-13.

Отже, щонайменше за долонно-підшовного генітального синдрому, синдрому Гуттмахера, стравоходу Баррета та раку стравоходу [14, 19] варто проводити прицільне секвенування гена *NOXA13* на предмет асоціацій з ними 59 додатково встановлених SNV гена.

3.2. Аналіз патогенних та умовно патогенних варіантів гена *HOXB13* *in silico*

У базі dbSNP (<https://www.ncbi.nlm.nih.gov/snp>) натеper зареєстровано 3522 одонуклеотидних варіантів (SNV) гена *HOXB13*, 902 з яких припадають на екзони (26,1 %). 628 SNV гена *HOXB13* (69,6 %) належать до міссенс-мутацій. До патогенних та ймовірно патогенних віднесено тільки 1 з них (0,1 %).

У базі Uniprot (<https://www.uniprot.org>) міститься інформація про 870 SNV гена *HOXB13*, 717 з яких віднесено до патогенних (82,4 %).

Найменше варіантів гена *HOXB13* – 277 – було виявлено сервісом SIFT, а найбільше – 650 – сервісом PolyPhen-2, при цьому найменше патогенних та умовно патогенних серед них було визначено сервісом SIFT – 145 (52,4 %), а найбільше – сервісом fathmm – 410 (66,9 %) (табл. 3.3), що може бути зумовлене особливостями алгоритмів, використаних у цих сервісах [3-5, 16, 21-23].

Таблиця 3.3

Визначення патогенних та умовно патогенних варіантів гена *HOXB13* за допомогою різних біоінформатичних інструментів

SNV	Біоінформатичний інструмент					
	fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
Загальна кількість	623	619	650	624	277	267
Кількість патогенних та умовно патогенних	410	269	294	399	145	160
% патогенних та умовно патогенних	65,8	43,5	45,2	63,9	52,4	59,9

Аналіз результатів дослідження варіантів гена *HOXB13* за допомогою біоінформатичних інструментів показало, що найменше SNV – 8 – були визначені як патогенні чи умовно патогенні усіма шістьма web-сервісами (0,9 %), а найбільше – 266 – жодним (30,6 %) (рис. 3.3). Таким чином, як

патогенні можна оцінити 107 варіантів (17,4 %), як умовно патогенні – 206 (23,7 %) (рис. 3.3). При цьому, 27 з зазначених SNV не вказані як патогенні варіанти ані в базі UniProt, ані в базі dbSNP (табл. 3.4).

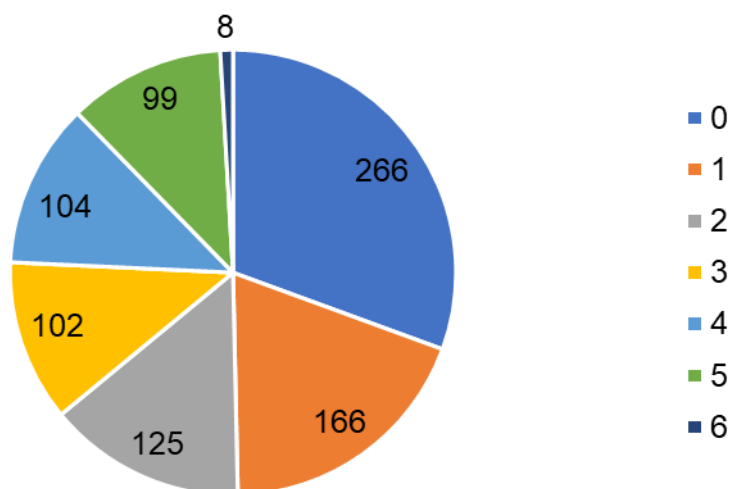


Рис. 3.3. Розподіл варіантів гена *HOXB13* за ступенем патогенності.

Таблиця 3.4

**Варіанти гена *HOXB13*,
не зареєстровані у біоінформатичних базах як патогенні**

1	SNV	Заміна	Біоінформатичний інструмент					
			fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
2	3	4	5	6	7	8	9	
1.	rs775273363	A128V						
2.	rs770891609	Y137S						
3.	rs575899185	Q138H						
4.	rs1597933974	Y169H						
5.	rs1242831503	Q188K						
6.	COSM1130240	K218E						
7.	COSM4067402	R220C						
8.	rs1597932884	R220H						
9.	COSM3518751	P222S						
10.	TCGA novel	K225N						
11.	rs757776260	Q227R						
12.	rs2038218044	L228F						
13.	rs1370849824	E232K						
14.	rs778481913	R233W						
15.	rs2038217786	E234G						
16.	TCGA novel	E234K						
17.	rs1272588905	K245Q						
18.	rs754598240	R246M						
19.	rs751610575	S250L						
20.	rs775731553	T253P						

Продовження таблиці 3.4

	2	3	4	5	6	7	8	9
21.	rs949512119	T253S						
22.	rs2038216342	L255Q						
23.	rs759838136	R258S						
24.	rs1207429780	I262V						
25.	rs747483887	W263R						
26.	rs1163383938	K273N						
27.	TCGA novel	K273T						

Вивчення густини замін амінокислот у білку НохВ-13 показало, що вона в середньому по гену становила 1,02. Цей показник був найнижчим у позиціях 121-140, а найвищим – у позиціях 41-60. Максимальна густина замін майже у 2,5 рази перевищувала таку в гені *HOXA13* (рис. 3.2) та у 1,5 рази – вказану для гена *LDLR* [8]. Патогенні заміни були відсутні в позиціях 21-120, фа найбільше їх було в позиціях 241-260. Найменше умовно патогенних варіантів було в позиціях 101-120, а найбільше – у позиціях 161-180. Таким чином, заміни амінокислот в білку більш поширені ближче до обох його кінців, а патогенні зосереджені у другій його половині.

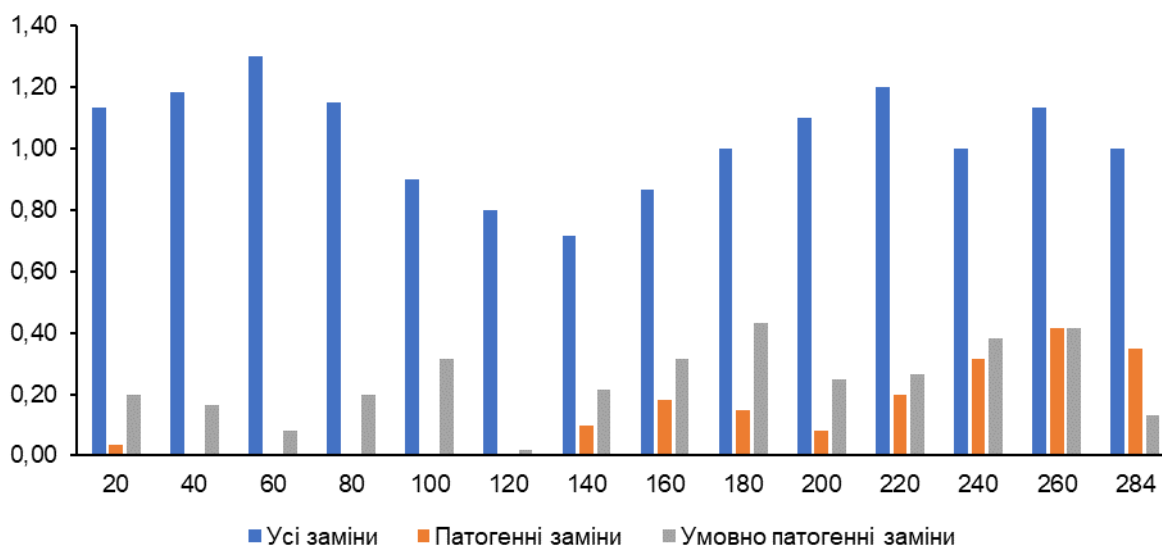


Рис. 3.4. Густина замін амінокислот у білку НохВ-13.

Отже, щонайменше за раннього раку передміхурової залози, грудей та колоректального раку [6, 19] варто проводити прицільне секвенування гена *НОХВ13* на предмет асоціацій з ними 27 додатково встановлених SNV гена.

3.3. Аналіз патогенних та умовно патогенних варіантів гена *HOXD13* *in silico*

У базі dbSNP (<https://www.ncbi.nlm.nih.gov/snp>) натеper зареєстровано 4495 однонуклеотидних варіантів (SNV) гена *HOXD13*, 936 з яких припадають на екзони (20,8 %). 654 SNV гена *HOXD13* (69,9 %) належать до місенс-мутацій. До патогенних та ймовірно патогенних віднесено тільки 11 з них (1,2 %).

У базі Uniprot (<https://www.uniprot.org>) міститься інформація про 352 SNV гена *HOXD13*, 19 з яких віднесено до патогенних (82,4 %).

Найменше варіантів гена *HOXD13* – 245 – було виявлено сервісом SIFT, а найбільше – 352 – сервісом PANTHER, при цьому найменше патогенних та умовно патогенних серед них було визначено сервісом PhD-SNP – 74 (24,0 %), а найбільше – сервісом fathmm – 256 (74,4 %) (табл. 3.5), що може бути зумовлене особливостями алгоритмів, використаних у цих сервісах [3-5, 16, 21-23]. Ці показники перевищують встановлені А. Іпатовою та М. Бражник для цього гена [1].

Таблиця 3.5

Визначення патогенних та умовно патогенних варіантів гена *HOXD13* за допомогою різних біоінформатичних інструментів

SNV	Біоінформатичний інструмент					
	fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
Загальна кількість	344	352	308	340	245	341
Кількість патогенних та умовно патогенних	256	140	74	193	187	220
% патогенних та умовно патогенних	74,4	39,8	24,0	56,8	76,3	64,5

Аналіз результатів дослідження варіантів гена *HOXD13* за допомогою біоінформатичних інструментів показало, що найменше SNV – 13 – були визначені як патогенні чи умовно патогенні жодним web-сервісом (3,7 %), а найбільше – 75 – двома (21,3 %) (рис. 3.5). Таким чином, як патогенні можна

оцінити 89 варіантів (25,3 %), як умовно патогенні – 119 (33,8 %) (рис. 3.8). При цьому, 76 з зазначених SNV не вказані як патогенні варіанти ані в базі UniProt, ані в базі dbSNP (табл. 3.6), тоді як в дослідженні А. Іпатової та М. Бражник таких варіантів було 56 [1].

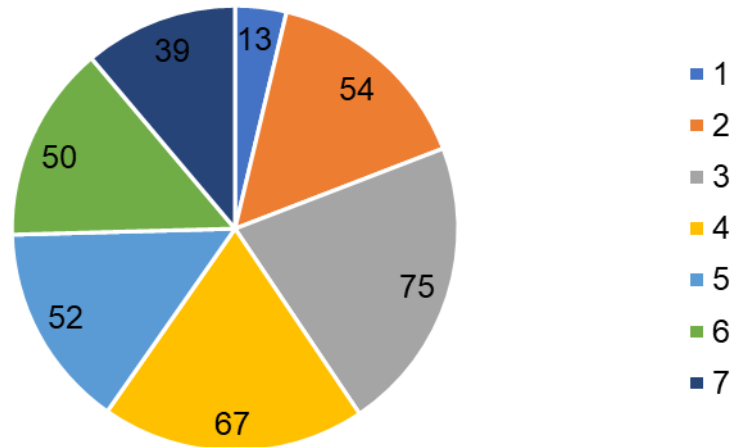


Рис. 3.5. Розподіл варіантів гена *HOXD13* за ступенем патогенності.

Вивчення густини заміни амінокислот у білку НохD-13 показав, що вона в середньому по гену становила 0,35, що в 3 рази менше цього показника в гені *HOXB13* та відповідає такому в гені *HOXA13*. Цей показник був найнижчим у позиціях 321-340, а найвищим – у позиціях 61-80. Максимальна густина заміни була в 2,5 рази нижчою, ніж в гені *HOXB13* та відповідала такій у гені *HOXA13* (рис. 3.2, 3.4). Патогенні заміни були відсутні в позиціях 21-40 та 81-120, тоді як найбільше їх було в позиціях 301-320. Найвищий показник був у 1,6 рази вищим, ніж у гені *HOXA13*, і в 1,6 рази нижчим, ніж в гені *HOXB13*. Найменше умовно патогенних варіантів було в позиціях 321-340, а найбільше – у позиціях 61-80. Таким чином, заміни амінокислот в білку розподілені нерівномірно, а патогенні зосереджені у другій його половині.

Отже, щонайменше за деяких форм синполідактилії та брахідактилії [9, 10, 19] варто проводити прицільне секвенування гена *HOXD13* на предмет асоціацій з ними 76 додатково встановлених SNV гена.

Таблиця 3.6

**Варіанти гена *HOXD13*,
не зареєстровані у біоінформатичних базах як патогенні**

1	SNV	Заміна	Біоінформатичний інструмент					
			fathmm	PANTHER	PhD-SNP	PolyPhen-2	SIFT	SNPs&GO
2	3	4	5	6	7	8	9	
1.	TCGA novel	W7S						
2.	rs1290585670	R41G						
3.	rs1356700815	G51R						
4.	rs1356700815	G51W						
5.	rs987876622	S72F						
6.	rs1334930092	Y76N						
7.	rs1384132694	G128R						
8.	rs847195	F131L						
9.	rs1204577598	Y135H						
10.	rs770945263	Y136H						
11.	rs745735651	C138R						
12.	rs1467183582	R139S						
13.	rs771851753	S141L						
14.	rs1224734678	G145D						
15.	COSM5058275	K166R						
16.	rs200067774	Y167C						
17.	rs922763007	R186G						
18.	rs763149315	R186L						
19.	rs763149315	R186Q						
20.	rs1431472260	A187G						
21.	rs1431472260	A187V						
22.	rs536676052	K188E						
23.	rs768084107	K188R						
24.	rs753157172	E189K						
25.	rs1323144880	F192C						
26.	COSM1009835	Y193C						
27.	rs554977181	Y200S						
28.	rs1245608993	Q201K						
29.	rs758363080	Q201R						
30.	rs147720746	Y206C						
31.	COSM1305871	E217K						
32.	COSM4086915	R219Q						
33.	rs1211490634	M226T						
34.	rs1359873193	G228R						
35.	rs577016423	Y229C						
36.	rs1574943380	Y229D						
37.	rs142600246	Q230H						
38.	rs1574943387	W238G						
39.	rs764371902	Q241H						
40.	rs761192563	Q241R						
41.	rs575180284	V242G						
42.	COSM4086916	Y243H						
43.	rs1162865491	C244R						
44.	rs544065459	K246N						
45.	rs1318043833	Q248P						

	2	3	4	5	6	7	8	9
46.	rs374920154	H253Y						
47.	rs779349115	V263M						
48.	rs1233152035	C271Y						
49.	rs770519205	R274Q						
50.	COSM6155482	R275I						
51.	rs1201943345	G276R						
52.	TCGA novel	R277E						
53.	rs1266860217	R277W						
54.	rs773877947	P282L						
55.	TCGA novel	Y283C						
56.	rs759071152	L288R						
57.	rs1304647913	E292Q						
58.	rs758726777	E294A						
59.	COSM1401419	Y295C						
60.	rs925188102	K299E						
61.	TCGA novel	F300S						
62.	TCGA novel	F300V						
63.	rs755071995	K303E						
64.	rs376366804	D304G						
65.	rs767313108	R307W						
66.	rs1205745904	R308C						
67.	rs770226826	R308H						
68.	rs1272438744	A311G						
69.	rs1301927745	T313A						
70.	rs1432063993	T313M						
71.	rs1301927745	T313P						
72.	COSM4086919	E317D						
73.	rs1457102355	N326K						
74.	COSM1482204	N326T						
75.	COSM3570176	R327Q						
76.	rs1436274970	S336P						

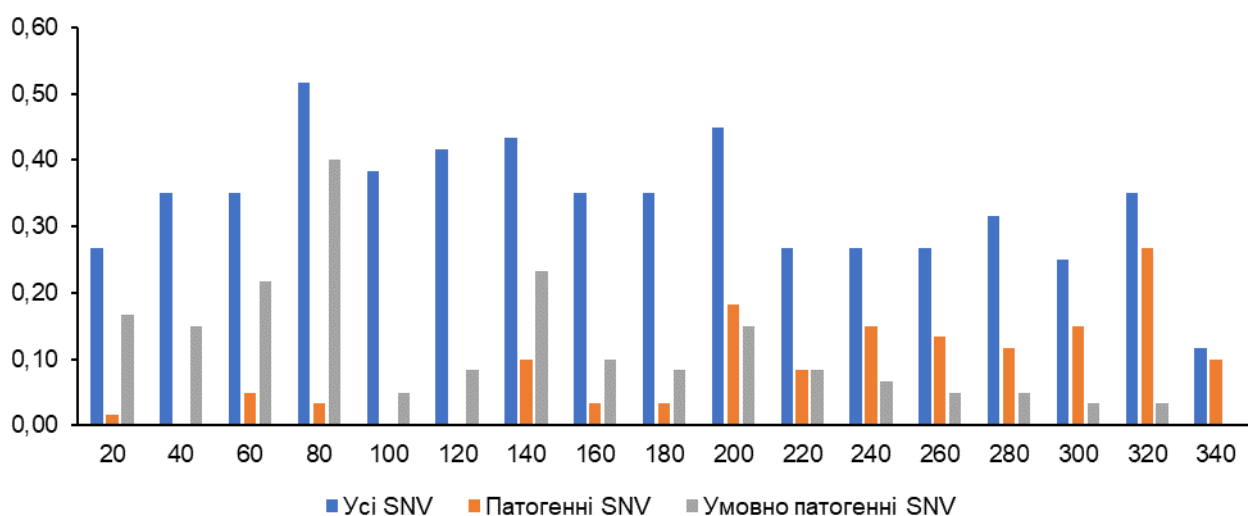


Рис. 3.6. Густина замін амінокислот у білку NoxD-13.

ВИСНОВКИ

На підставі наведених даних ми дійшли наступних висновків:

1. Встановлено, що серед досліджених гомеозисних генів у біоінформатичних базах dbSNP та Uniprot депоновано найбільше однонуклеотидних варіантів *HOXB13*, 82,4 % з яких віднесені до патогенних.

2. Визначено, що найбільше SNV гомеобокс-генів за допомогою біоінформатичних інструментів також визначається для гена *HOXB13*. Максимальну кількість варіантів дозволяють проаналізувати web-сервіси PANTHER та PhD-SNP, а найменшу – SIFT.

3. Найбільше патогенних та умовно патогенних варіантів *in silico* виявлено в гені *HOXB13* – 107 та 206 відповідно.

4. В генах *HOXA13*, *HOXB13* та *HOXD13* встановлено 162 патогенних SNV, які не зазначені в біоінформатичних базах, найбільше їх було в гені *HOXD13*.

5. Густина SNV була найвищою в гені *HOXB13*, при цьому розподіл однонуклеотидних варіантів по гомеозисних генах був нерівномірним.

СПИСОК ВИКОРИСТАНИХ ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1. Іпатова АВ, Бражник МО. Біоінформатичний аналіз патогенних та ймовірно патогенних однонуклеотидних варіантів гомеозисного гена HOXD13. *Шоста міжнародна конференція молодих учених «Харківський природничий форум»* (18-19 травня 2023 р.), збірник тез. Харків: ХНПУ імені Г. С. Сковороди, 2023. С. 403–404.
2. Садлер Т.В. Медична ембріологія за Лангманом. Львів: НАУТІЛУС, 2001. 550 с.
3. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. *Nat Methods*. 2010;7(4):248–249.
4. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat*. 2009. 30(8):1237–1244.
5. Capriotti E., Calabrese R., Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics*. 2006;22(22): 2729–2734.
6. Chandrasekaran G, Hwang E, Kang T, Kwon DD, Kwangsung P, Lee J-J, Lakshmanan V-K. Computational Modeling of complete HOXB13 protein for predicting the functional effect of SNPs and the associated role in hereditary prostate cancer. *Sci Rep*. 2017;7:43830.
7. Chen Y, Dufour CR, Han L, Li T, Xia H, Giguere V. Hierarchical Phosphorylation of *HOXB13* by mTOR Dictates Its Activity and Oncogenic Function in Prostate Cancer. *Mol Cancer Res*. 2023;21(10):1050–1063.
8. Chora JT, Medeiros AM, Alves AC, Bourbon M. Analysis of publicly available *LDLR*, *APOB*, and *PCSK9* variants associated with familial hypercholesterolemia: application of ACMG guidelines and implications for familial hypercholesterolemia diagnosis. *Genet Med*. 2018;20(6):591–598.

9. Gottschalk A, Sczakiel HL, Hulsemann W, Schwartzmann S, Abad-Perez AT, Grunhagen J, Ott CE, Spielmann M, Horn D, Mundlos S, Jamsheer A, Mensah MA. HOXD13-associated synpolydactyly: Extending and validating the genotypic and phenotypic spectrum with 38 new and 49 published families. *Genet Med.* 2023;25(11):100928.
10. Guo R, Fang X, Mao H, Sun B, Zhou J, An Y, Wang B. A Novel Missense Variant of *HOXD13* Caused Atypical Synpolydactyly by Impairing the Downstream Gene Expression and Literature Review for Genotype–Phenotype Correlations. *Front Genet.* 2021;12:731278.
11. Hubert KA, Wellik DM. Hox genes in development and beyond. *Development.* 2023;150(1): dev192476.
12. Juárez-Rendón KJ, Castro-García MA, Prada-Ortega DG, Rivera G, Ruíz-Godoy LM, Enríquez-Cárcamo VI, Reyes-Lopez MA. Variants Identified in the *HOXC13* and *HOXD13* Genes Suggest Association with Cervical Cancer in a Cohort of Mexican Women. *Genes.* 2023;14(2):358.
13. Innis JW. Hand-Foot-Genital Syndrome. 2006 Jul 11 [updated 2019 Aug 8]. In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Amemiya A, editors. GeneReviews(R) [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2024. Available from <http://www.ncbi.nlm.nih.gov/books/NBK1423/>
14. Janmaat VT, Nesteruk K, Spaander MC, Verhaar AP, Yu B, Silva RA, Phillips WA, Magierowski M, van de Winkel A, Stadler HS, Sandoval-Guzmán T, van der Laan LJ, Kuipers EJ, Smits R, Bruno MJ, Fuhler GM, Clemons NJ, Peppelenbosch MP. HOXA13 in etiology and oncogenic potential of Barrett’s esophagus. *Nat Commun.* 2023;12(1): 3354.
15. Mazziotta C, Iaquina MR, Tramarin ML, Badiale G, Cervellera CF, Tonnini G, Patergnani S, Pinton P, Lanza G, Gafa R, Tognon M, Martini F, De Mattei M, Rotondo JC. Hsa-microRNA-1249-3p/Homeobox A13 axis modulates the expression of beta-catenin gene in human epithelial cells. *Sci Rep.* 2023;13(1):22872.

16. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812–3814.
17. Niakan K, Eggan K. Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Dev Biol.* 2013;375(1):45–54.
18. Pineault KM, Welli DM. Hox Genes and Limb Musculoskeletal Development. *Curr Osteoporos Rep.* 2014;12(4):420–427.
19. Quinonez SC, Innis JW. Human *HOX* gene disorders. *Mol Genet Metab.* 2014;111(1):4–15.
20. Shahbazi MN. Mechanisms of human embryo development: from cell fate to tissue shape and back. *Development.* 2020;147(14):dev190629.
21. Shihab HA, Gough J, Mort M, Cooper DN, Day INM, Gaunt TR. Ranking Non-Synonymous Single Nucleotide Polymorphisms based on Disease Concepts. *Hum Genomics.* 2014;8(1):11.
22. Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics.* 2016;32(14):2230–2232.
23. Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat.* 2011;32(4): 358–368.