

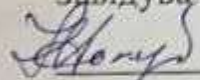
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Факультет біотехнології і біотехніки

Кафедра біоенергетики, біоінформатики та екобіотехнології

До захисту допущено:

Завідувач кафедри

 Наталія ГОЛУБ

«05» квітня 2023 р.

Дипломна робота

на здобуття ступеня бакалавра

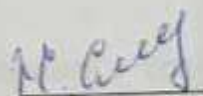
за освітньо-професійною програмою «Біотехнології»

спеціальності 162 «Біотехнології та біоінженерія»

на тему: «Первинний аналіз диференційно експресованих генів у
плаценті людини між першим і другим триместрами фізіологічної
вагітності»

Виконала:

студентка IV курсу, групи БМ-91
Жалілова Євгенія Олександрівна



Керівник:

Доцент кафедри біоенергетики,
біоінформатики та екобіотехнології,
к.т.н., с.н.с.

Маринченко Лоліта Вікторівна

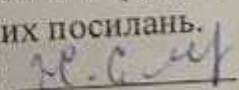


Рецензент:

Доцент кафедри біоенергетики,
біоінформатики та екобіотехнології,
к.б.н., с.н.с.

Ключко Віталій Вікторович



Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.
Студент (-ка) 

Київ – 2023 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет біотехнології і біотехніки

Кафедра біоенергетики, біоінформатики та екобіотехнології

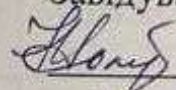
Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 162 «Біотехнології та біоінженерія»

Освітньо-професійна програма «Біотехнології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

 Наталія ГОЛУБ

« 17 » квітня 20__ р.

ЗАВДАННЯ

на дипломну роботу студенту

Жалілова Євгенія Олександрівна

1. Тема роботи «Первинний аналіз диференційно експресованих генів у плаценті людини між першим і другим триместрами фізіологічної вагітності», керівник роботи Маринченко Лоліта Вікторівна, доцент кафедри КББЕ, к.т.н., с.н.с., затверджені наказом по університету від « 22 » жовтня 2023 р. № 1888-с

2. Термін подання студентом роботи 05.06.2023

3. Вихідні дані до роботи: список диференційно експресованих генів між 1 і другим триместрами фізіологічної вагітності: <https://github.com/Sashkow/r-affymetrix-integration-pipeline/blob/main/data/difexp.tsv>

4. Зміст роботи: Розкриття змісту понять: плацента, прееклампсія, експресія генів, наведення методів аналізу експресії генів, визначення терміну volcano plot, його ефективності порівняно з іншими методами, побудова volcano plot, визначення найбільш значущих диференційно експресованих генів між 1 та 2 триместрами вагітності, аналіз функцій отриманих 20 генів.

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)

Презентація на захист

6. Консультанти розділів роботи*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання 17.04.2023

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Літературний пошук матеріалів	17.04.2023 30.04.2023	– викон.
2	Розробка алгоритму побудови графіку volcano plot	30.04.2023 07.05.2023	– викон.
3	Функціональний аналіз отриманих генів	08.05.2023 21.05.2023	– викон.
4	Оформлення результатів роботи	22.05.2023 04.06.2023	– викон.

Студент

С.С.

Євгенія ЖАЛІЛОВА

Керівник

М.М.

Лоліта МАРИНЧЕНКО

ЗМІСТ

ВСТУП.....	5
1 ЛІТЕРАТУРНИЙ ОГЛЯД.....	9
1.1 Плацента.....	9
1.2 Патологічний стан вагітних жінок – прееклампсія.....	9
1.3 Методи для аналізу експресії генів: мікроарей-технології і РНК-секвенування.....	14
1.3.1 Характеристика мікроарей-технології.....	17
1.3.2 Отримання даних РНК-секвенування для визначення експресії генів	21
1.4 Функціональний аналіз.....	23
1.4.1 Диференційно експресовані гени	23
1.4.2 Функціональний аналіз експресованих генів. Використання бази даних String-db	24
1.4.3 Візуалізація даних. Побудова volcano plot для диференційно експресованих генів	26
2 МАТЕРІАЛИ І МЕТОДИ.....	29
2.1 Матеріали досліджень.....	29
2.2 Методи досліджень	29
3 РЕЗУЛЬТАТИ І ОБГОВОРЕННЯ	36
4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ	59
ВИСНОВКИ	68
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	69

ВСТУП

У сучасній молекулярній біології досягнуто значних успіхів, завдяки яким отримано багато нових знань про функції генів, механізми регуляції генної експресії, процеси розвитку та патологічні стани організму. Розвиток методів секвенування ДНК, зокрема масового паралельного секвенування (NGS), дає змогу отримати великі обсяги геномних даних за короткий проміжок часу. Це відкрило нові можливості для вивчення генетичної основи хвороб, ідентифікації генетичних маркерів та розвитку персоналізованої медицини.

Для аналізу та інтерпретації великих обсягів біологічних даних, що отримані з високотехнологічних методів, таких як секвенування геному, мас-спектрометрія та мікрочіп-технології, використовують біоінформатичні методи. Ці методи дають змогу виявляти, наприклад, складні зв'язки між генетичними варіаціями, виразністю генів та фенотипом, що, зокрема, сприяє розумінню молекулярних механізмів життєдіяльності організмів та розвитку нових методів діагностики та лікування хвороб.

Прееклампсія – це серйозне захворювання, яке може виникнути під час вагітності у деяких жінок. Ця хвороба проявляється поєднанням високого артеріального тиску (гіпертензії) і наявністю білка в сечі (протеїнурії). Прееклампсія зазвичай розвивається після 20-го тижня вагітності і може бути небезпечною як для матері, так і для плода. Для матері прееклампсія може призвести до ураження органів, таких як нирки, печінка, легені та мозок, через обмеження притоку крові та пошкодження судин. Прееклампсія також може впливати на розвиток плода і функціонування плаценти, що може призвести до обмеженого постачання кисню та поживних речовин до плода. Це може спричинити різноманітні проблеми з розвитком плода, включаючи недостатній приріст ваги, передчасні пологи або навіть смерть плоду.

Саме тому **актуальним** є пошук способів, які дадуть змогу виявити преєклампсію до перших її проявів. Одним із таких способів є виявлення маркерів, які б свідчили про загрозу, для чого необхідно проаналізувати молекулярні механізми, які призводять до цього. Оскільки преєклампсія діагностується після 20 тижня вагітності, потрібно знайти ці маркери на більш ранніх етапах розвитку плода, наприклад, проаналізувати зміни експресії генів між першим та другим триместрами вагітності, які є критичними для розвитку дисфункцій плаценти.

Для цієї мети незамінними є методи аналізу експресії генів, які дають змогу швидко і наглядно визначити відмінності експресії генів між двома станами.

Раніше було проведено аналіз експресії генів плаценти людини у стані преєклампсії у роботах Брю О. і Саллівана М., Йонг Х. і Чана С. [1,2].

Проблемою виявлення найбільш значущих диференційно експресованих генів є те, що існує розбіжність між статистичним і біологічним значенням диференційної експресії. У деяких біологічних застосуваннях гени вибирали за зміною їх рівня експресії (наприклад, використання 2-кратної зміни (Fold change) для відсічення незначущого рівня), а варіабельність експресії генів не враховували. Одним з варіантів вирішення цієї проблеми є використання методу volcano plot, який поєднує статистичний показник значущості p-value з Fold change.

Тому **метою** нашої роботи є виявлення топ-20 диференційно експресованих генів у плаценті людини між другим та першим триместрами вагітності за допомогою методу volcano plot.

Для досягнення цієї мети було поставлено такі **завдання**:

- виконати науковий літературний пошук за темою роботи, навести методи аналізу експресії генів, обґрунтувати вибір застосування методу volcano plot;
- за даними про експресію генів між першим та другим триместрами фізіологічної вагітності, побудувати діаграму volcano plot;

- визначити 20 найбільш значущих диференційно експресованих генів за фізіологічної вагітності між 1 та 2 триместрами;
- проаналізувати функції отриманих 20 генів, рівень експресії яких може бути маркером прееклампсії.

Наукова новизна: вперше за допомогою методу volcano plot було визначено гени з найбільшою різницею в експресії між другим і першим триместрами за фізіологічної вагітності.

Практична значимість: розроблений метод може бути застосований для надійного виявлення найбільш значущих експресованих генів між 1 і 2 триместрами фізіологічної вагітності, а отриманні в ході аналізу дані щодо диференційно експресованих генів можуть бути використані для подальших досліджень у визначенні надійних ранніх маркерів прееклампсії.

Об'єкт дослідження: диференційно експресовані гени між 1 та 2 триместрами фізіологічної вагітності.

Предмет дослідження: застосування біоінформатичного методу volcano plot для виявлення найбільш значущих диференційно експресованих генів.

Методи досліджень: біоінформатичні методи (volcano plot).

Апробація. Результати роботи було обговорено на семінарах відділу системної біології Інституту молекулярної біології і генетичної інженерії НАН України, опубліковано тези у збірнику робіт XVII Міжнародної науково-практичної конференції студентів, аспірантів і молодих вчених «Біотехнологія XXI століття»:

Жалілова Є.О., Лихенко О.К., Оболенська М.Ю. Топ-двадцять диференційно експресованих генів у плаценті людини між другим і першим триместрами фізіологічної вагітності. 19.05.5023

Робота виконана на базі Інституту молекулярної біології та генетичної інженерії НАН України, відділу системної біології в рамках виконання теми «Первинний аналіз диференційно експресованих генів у плаценті людини

між першим і другим триместрами фізіологічної вагітності», номер держреєстрації 05417101.

Висловлюю щирі подяки завідувачій відділу, професору, доктору біологічних наук Оболенській Марії Юріївні за наукове консультування в процесі виконання роботи.

1 ЛІТЕРАТУРНИЙ ОГЛЯД

1.1 Плацента

Плацента – це тимчасовий орган, який розвивається під час вагітності та поєднує організм матері з плодом у триєдину систему – мати – плацента – плод.

Плацента відіграє вирішальну роль у здоров'ї та благополуччі як плоду, так і його матері. Функціонально плацента підтримує ріст плоду, оскільки вона сприяє доставці кисню та поживних речовин і видаленню відходів [3].

Розвиток плаценти людини вимагає скоординованої взаємодії між лініями трофобласта плаценти та ендометрієм матері.

Порушення функціонування плаценти є основною причиною загальних розладів вагітності, включаючи викидні, затримку розвитку плода, мертвонародження та смертельно небезпечну для матері та плода преєклампсію. Незважаючи на критичну важливість для репродуктивного успіху та здоров'я протягом усього життя, наше розуміння розвитку плаценти є невеликим, головним чином через етичні обмеження вивчення раннього або хронологічного розвитку плаценти, відсутність довгострокових моделей *in vitro* або порівняльних моделей на тваринах [4].

1.2 Патологічний стан вагітних жінок – преєклампсія

Більшість науковців схилиються до думки, що дисфункція плаценти провокує розвиток преєклампсії.

Було показано, що з народженням плаценти зникають основні клінічні симптоми преєклампсії, а у разі пухирного занеску (трофобластичної хвороби, який характеризується розростанням ворсин хоріона у стінці матки за відсутності нормальної ембріональної тканини або часткової присутності елементів нежиттєздатного плода [5]) за наявності у порожнині матки клітин трофобласту і відсутності самого плоду теж можуть спостерігатися клінічні ознаки преєклампсії. На молекулярному і клітинному рівнях у вагітних з преєклампсією плацента має змінений склад клітин та експресію генів.

Прееклампсія визначається як наявність у вагітної систолічного артеріального тиску (САТ), що перевищує або дорівнює 140 мм рт. ст., або діастолічного артеріального тиску (ДАТ), що перевищує або дорівнює 90 мм рт. ст. або вище, у разі двох вимірювань з інтервалом принаймні 4 години у пацієнтки із зазвичай нормальним тиском.

Крім того, спостерігається протеїнурія, що перевищує або дорівнює 0,3 г/л білка в 24-годинному зразку сечі, або білок у тест-смужці сечі (якщо кількісне вимірювання недоступне).

Прееклампсія з тяжкими ознаками визначається як наявність одного з таких симптомів або ознак:

- САТ 160 мм рт. ст. або вище або ДАТ 110 мм рт. ст. або вище, у разі двох вимірювань з інтервалом щонайменше 4 години, поки пацієнт дотримується лежачого режиму (якщо раніше не було розпочато антигіпертензивну терапію);
- порушення функції печінки, про що свідчить аномально підвищена концентрація печінкових ферментів у крові (вдвічі, порівняно з нормальною концентрацією), сильний постійний біль у верхньому квадранті чи епігастральній ділянці, який не реагує на фармакотерапію та не пояснюється альтернативними діагнозами;
- прогресуюча ниркова недостатність (концентрація сироваткового креатиніну $>1,1$ г/л або подвоєння концентрації сироваткового креатиніну за відсутності інших захворювань нирок);
- вперше виниклі церебральні розлади або порушення зору;
- набряк легенів;
- тромбоцитопенія (кількість тромбоцитів $< 100\,000/\text{мкл}$).

У пацієнтки з гіпертензією, що виникла вперше і без протеїнурії, ознаками прееклампсії можуть бути будь-які з такого переліку:

- кількість тромбоцитів нижче $100\,000/\text{мкл}$;

- рівень сироваткового креатиніну вище 1,1 г/л або подвоєння сироваткового креатиніну за відсутності інших захворювань нирок;
- рівні печінкових трансаміназ принаймні вдвічі перевищують нормальні концентрації;
- набряк легенів;
- церебральні або зорові симптоми.

Фактори ризику прееклампсії такі:

- багатоплідна вагітність;
- прееклампсія при попередній вагітності;
- хронічна гіпертонія;
- діабет;
- тромбофілія;
- індекс маси тіла до вагітності більше 30;
- вік матері 35 років і старше;
- захворювання нирок;
- допоміжні репродуктивні технології.

Ознаки та симптоми

Оскільки клінічні прояви прееклампсії можуть бути неоднорідними, діагностика прееклампсії може бути не простою. Прееклампсія без тяжких ознак може протікати безсимптомно. Багато випадків виявляються під час планового пренатального скринінгу.

Пацієнти з прееклампсією з важкими ознаками виявляють ураження кінцевих органів і можуть скаржитися на таке:

- головний біль;
- порушення зору: нечіткі, мерехтливі скотоми;
- змінений психічний стан;
- сліпота: може бути кіркова або сітківка;
- задишка;
- набряк: раптове збільшення набряку тіла або набряк обличчя;

- біль в епігастрії або правому верхньому квадранті живота;
- слабкість або нездужання: може бути ознакою гемолітичної анемії [4].

На цей час методи лікування прееклампсії включають невідкладну антигіпертензивну терапію [6], лікування сульфатом магнію [7-8] та аспірином [9]. Однак, у разі тяжкого перебігу вагітності і досягнення 34-ти тижнів, найбільш рекомендованим методом є розродження.

Біомаркери для діагностики, передбачення та прогнозу прееклампсії

Висновок комітету ACOG (American College of Obstetricians and Gynecologists), виданий у 2015 році та підтверджений у 2017 році, не рекомендує проводити скринінг для прогнозування прееклампсії, окрім отримання відповідної історії хвороби. Через відсутність адекватних методів скринінгу та важкі наслідки захворювання, усі жінки з підозрою на прееклампсію проходять обстеження, яке потребує інтенсивних ресурсів, що часто потребує кількадечної госпіталізації для подальшого обстеження. Були досліджені й інші методи скринінгу, включаючи метаболомічні шляхи та комбіновані метаболомічні та протеомні підходи до даних. Пряме порівняння методу скринінгу на основі алгоритму Fetal Medicine Foundation у 2017 році (поєднання факторів матері, середнього артеріального тиску, індексу пульсації маткової артерії та PlGF (Placental Growth Factor)) продемонстрували перевагу над методами скринінгу, рекомендованими на даний момент Національним інститутом здоров'я та догляду та ACOG.

В дослідженні ASPRE (аспірин для профілактики прееклампсії, заснованого на доказах) було використано алгоритм скринінгу, який включав додавання сироватки матері, що містить PAPP-A (асоційований з вагітністю протеїн А плазми крові). Цей алгоритм був застосований для скринінгу жінок у першому триместрі вагітності з метою виявлення тих, у кого існує високий ризик розвитку прееклампсії.

Жінок із високим ризиком рандомізували для отримання 150 мг аспірину або плацебо щодня до 36 тижнів вагітності. Щоденне застосування низьких доз аспірину жінками з високим ризиком було пов'язане зі значно нижчою частотою передчасної прееклампсії, ніж плацебо, а рівень виявлення передчасної прееклампсії становив 76,7% — 43,1% для доношеної прееклампсії — з рівнем хибнопозитивних тестів 9,1% [10].

Ці результати свідчать про те, що ранній скринінг із застосуванням біомаркерів плазми та візуалізаційні дослідження роблять доцільним раннє втручання та можливу профілактику захворювання.

У майбутньому аналізи біомаркерів є економічно ефективним та надійним методом скринінгу, який зможе врятувати життя, навіть у сферах обмеженого досвіду та ресурсів.

Натомість, аналіз диференційно експресованих генів у плаценті людини є важливим для пошуку причин прееклампсії з кількох причин:

- Плацента відіграє ключову роль у розвитку прееклампсії: Прееклампсія є станом, пов'язаним з дисфункцією плаценти. Аналізуючи гени, які проявляють відмінну експресію в плаценті, можна отримати важливі відомості про процеси, що відбуваються в плаценті під час розвитку прееклампсії.

- Виявлення молекулярних механізмів: Диференційна експресія генів може допомогти ідентифікувати молекулярні механізми, які стоять в основі прееклампсії. Виявлення генів, які виявляють відмінну експресію, може розкрити дисрегуляцію у генетичній та молекулярній рівновазі, яка може спричиняти розвиток цього стану.

- Потенційні біомаркери та рання діагностика: Аналізуючи диференційно експресовані гени, можна виявити потенційні біомаркери прееклампсії. Це може сприяти ранньому виявленню та діагностиці прееклампсії, що важливо для покращення результатів вагітності та здоров'я матері та плода.

- Розуміння патогенезу: дослідження диференційної експресії генів дає змогу глибше зрозуміти патогенез прееклампсії, тобто механізми, що призводять до її розвитку. Це може сприяти розробці нових підходів до лікування та профілактики прееклампсії [10].

Перший і другий триместри вагітності є критичними для розвитку дисфункцій плаценти і для своєчасного встановлення діагнозу. Тому слід звернутись до аналізу диференційно експресованих генів у плаценті людини між першим і другим триместрами фізіологічної вагітності з метою виявлення найбільш значущої різниці в експресії генів з метою порівняння отриманих результатів з результатами подальшого дослідження за умов патології.

1.3 Методи для аналізу експресії генів: мікроарей-технології і РНК-секвенування

Гени кодують білки, а білки визначають функцію клітини. Таким чином, тисячі генів, що експресуються в клітині, встановлюють, як ця клітина буде функціонувати. Крім того, кожен етап інформаційного потоку в клітині, від ДНК до РНК і білка, регулює функцію клітини, регулюючи кількість і тип білка, що виробляється.

Кількість певного білка в клітині відображає баланс між біохімічними шляхами синтезу і деградації цього білка. З боку синтезу виробництво білка починається з транскрипції (перетворення ДНК на РНК) і продовжується трансляцією (перетворення РНК на білок). Тому контроль цих процесів відіграє важливу роль у визначенні того, які білки присутні в клітині і в яких кількостях. Крім того, те, як клітини обробляють транскрипти РНК і новоутворені білки, також сильно впливає на рівень білків [11].

Експресія генів – це процес перетворення інформації, закодованої в генах, у некодуючі і кодуючі РНК. Останні транлюються у білки. Експресія генів відбувається в кілька етапів [11].

Транскрипція: це процес створення РНК-копії гена з ДНК. Транскрипція відбувається за участю РНК-полімерази та інших факторів.

Посттранскрипційні зміни. Після транскрипції РНК потребує додаткової обробки, яка включає видалення інтронів (некодуючих участків гена) та з'єднання екзонів (кодуючих участків гена). Цей процес називається сплайсинг. Транспортування РНК: після обробки РНК може бути транспортована з ядра клітини до цитоплазми. Транспортування РНК відбувається за допомогою білкових комплексів.

Трансляція: це процес синтезу білка на матриці інформаційної (кодуючої) РНК. Трансляція відбувається на рибосомах, де послідовність РНК перекладається в послідовність амінокислот, які утворюють білок.

Посттрансляційна модифікація: після трансляції білок може бути підданий різноманітним модифікаціям, таким як глікозилювання, фосфорилування, ацетилювання тощо, які впливають на його структуру та функції.

Також відбувається сплайсинг (механізм, завдяки якому вирізаються інтрони – некодуючі ділянки пре-мРНК, тобто послідовності, які потім не будуть використані як матриця для біосинтезу білків).

Процесинг мРНК відбувається безпосередньо під час транскрипції в складі РНК-полімеразного комплексу, але за умов *in vivo*, у живих клітинах.

Щоб утворювати білки, ДНК транскрибується в матричну РНК, або мРНК, яка транслюється рибосомою в білок. Однак деякі гени кодують РНК, яка не перетворюється на білок; ці РНК називаються некодуючими РНК, або нкРНК.

Ці РНК можуть виконувати різні функції: рибосомні РНК (рРНК), транспортні (тРНК), малі інтерферуючі РНК (siRNA), мікроРНК, довгі некодуючі РНК та інші. Деякі нкRNA можуть мати одразу декілька функцій, взаємодіючи з різними молекулярними компонентами клітинного апарату.

Усі РНК, транскрибовані з генів, називаються транскриптами.

Транскриптом визначається як сукупність усіх зчитувань транскриптів, присутніх у клітині.

Аналіз РНК має декілька переваг порівняно з аналізом білків:

- РНК відображає активність генів: мРНК є проміжним продуктом генетичної інформації, який використовується для синтезу білків. Аналіз РНК дає змогу вимірювати рівень експресії генів.
- Ширший діапазон інтенсивності експресії: рівень експресії РНК може відрізнятися від рівня експресії білків, оскільки РНК є стабільнішою і дає змогу виявити менші зміни в експресії генів. Також при аналізі слід враховувати некодуючі РНК, які не транскрибуються.
- Вища роздільна здатність: аналіз РНК може забезпечити вищу роздільну здатність порівняно з аналізом білків, особливо коли розглядається велика кількість генів.

Відносна легкість визначення генотипу: аналіз РНК може допомогти визначити генетичний профіль організму, тоді як аналіз білків не дає змоги визначити точний генетичний склад організму [11].

Існує два підходи для аналізу експресії генів:

- Мікроарей-технології, засновані на гібридизації РНК з олігонуклеотидами ДНК: цей підхід використовується для вимірювання рівня експресії генів шляхом визначення кількості РНК, що гібридизується з заздалегідь відомими олігонуклеотидами, нанесеними на чіп. Цей метод дає можливість аналізувати експресію тисяч генів одночасно.
- Секвенування РНК: цей підхід використовується для вимірювання рівня експресії генів шляхом прямого секвенування РНК. Він забезпечує можливість детально досліджувати транскриптом.

Обидва методи, РНК-секвенування і використання мікрочіпів з олігонуклеотидами, використовуються для вивчення експресії генів та визначення того, які гени активні в певних умовах або зразках тканин.

1.3.1 Характеристика мікроарей-технології

Завершення визначення послідовності геному людини у 2001 році відкрило шлях для багатьох експериментів і досліджень. Одним з них є визначення ділянок ДНК, які контролюють нормальний і патологічний стан людини. Функціональна геноміка вивчає функцію генів шляхом паралельного вимірювання експресії геному [12].

Аналіз мікрочіпів можна розділити на два основні етапи: виробництво зонда та виробництво мішені (кДНК) з флуоресцентною міткою.

Невідомі молекули ДНК розрізають на фрагменти ендонуклеазами рестрикції; до цих фрагментів ДНК приєднуються флуоресцентні маркери. Мічена кДНК-мішень гібридує з іммобілізованим зондом, утворюючи дволанцюговий комплекс, а фрагменти ДНК, що залишились не зв'язаними, змиваються.

Зв'язані фрагменти ДНК можна ідентифікувати за зміною інтенсивності їх флуоресцентного випромінювання за допомогою лазерного променя. Якщо послідовності мішені та зонда комплементарні, виникає сигнал, який використовується для визначення наявності певних РНК у зразку, і визначає, які саме РНК присутні в невідомому зразку мішені, оскільки є змога ідентифікації кількох фрагментів ДНК одночасно [12]. Для запису картини випромінювання флуоресценції та ідентифікації ДНК використовують відповідне програмне забезпечення. Цей метод використання ДНК-чіпів є дуже швидким, крім того, він чутливий і специфічний.

Мікрочіп забезпечує основу для генотипування тисяч різних локусів одночасно, що корисно для досліджень асоціацій і зчеплення для виділення хромосомних ділянок, пов'язаних із певним захворюванням. Чіп чи чип?

Мікроарей-технологія ґрунтується на можливості розміщення великої кількості (десятки тисяч) різних послідовностей ДНК на невеликій поверхні, чіпі. Різні фрагменти ДНК розташовуються так, що ідентичність кожного

фрагмента відома за його позицією в масиві. Існує два типи мікрочипів: мікрочипи для експресії генів і тканинні мікрочипи.

Такий метод, як зворотна транскриптаза-полімеразна ланцюгова реакція (ЗТ-ПЛР), дають змогу проаналізувати лише кілька генів в одному експерименті. Однак мікрочипи не тільки дають змогу аналізувати набагато більше генів, ніж це було можливо раніше, але й мають ту перевагу, що на них не впливає попередній відбір генів.

мРНК – це проміжна молекула, яка переносить генетичну інформацію з ядра клітини в цитоплазму для синтезу білка. Коли ген експресується або перебуває в активному стані, багато копій мРНК, що відповідають певному гену, виробляються в процесі, який називається транскрипцією [12].

Оцінюючи різні мРНК, ми можемо оцінити генетичну інформацію або експресію генів. Це допомагає зрозуміти різні процеси, що обумовлюють кожну зміну генетичної експресії.

Для аналізу експресії генів за допомогою мікрочипів відбирається мРНК зразку, з якої отримують кДНК реверсивною транскриптазою.

Microarray Affymetrix і Microarray Illumina – це два види мікрочипів, які використовують різні принципи організації зондів. Обидва методи мають свої переваги та недоліки, і вибір між ними залежить від конкретного завдання та доступних ресурсів.

Технологія мікрочипів Affymetrix використовує олігонуклеотидні масиви для виявлення та аналізу експресії генів, а також варіацій послідовності ДНК і змін кількості копій. Ось спрощений огляд того, як працюють мікрочипи Affymetrix:

Конструкція мікроматриці: мікроматриці Affymetrix розроблено за допомогою комп'ютерних алгоритмів, які вибирають конкретні короткі послідовності ДНК, які називаються зондами. Ці зонди вибираються відповідно до конкретних генів або варіантів, що представляють інтерес [13].

Переваги методу Microarray Affymetrix:

- Цей метод широко використовується в дослідженнях експресії генів, що означає, що є великий досвід використання цього методу і його результати добре зрозумілі.
- Чіпи Affymetrix мають високу повторюваність та надійність, що означає, що результати можна повторити у різних лабораторіях.
- Різні види чіпів Affymetrix використовуються для аналізу експресії генів для різних організмів.
- Недоліки методу Microarray Affymetrix:
- Чіпи Affymetrix дорожчі, ніж чіпи Illumina, що може бути проблемою для досліджень з обмеженим бюджетом.
- Чіпи Affymetrix не так легко налаштовувати для дослідження нових генів чи нових організмів, як чіпи Illumina.

Bead-based microarray technology Illumina (технологія мікрочіпів на основі кульок) використовує мікроскопічні кульки (beads), які містять олігонуклеотиди, що відповідають певним генам.

Кожна кулька має унікальний ідентифікатор, який дає змогу відрізнити її від інших кульок на мікрочіпі. Кульки розташовуються на мікрочіпі відповідно до ідентифікаторів, що дає змогу одночасно визначати експресію багатьох генів.

Для проведення аналізу експресії генів, як і для Microarray Affymetrix, кДНК, отримана зворотньою транскриптазою, позначається флуоресцентною міткою і наноситься на поверхню мікрочіпа, де взаємодіє з олігонуклеотидами.

Переваги методу Microarray Illumina:

- Чіпи Illumina доступніші за вартістю, ніж чіпи Affymetrix.
- Цей метод можна використовувати для аналізу більшої кількості генів, ніж метод Affymetrix.
- Чіпи Illumina можуть бути легко налаштовані для вивчення нових генів або нових організмів.

Недоліки методу Microarray Illumina:

– Чіпи Illumina можуть мати більш високу варіабельність та помилки, що може знизити точність результатів.

– Цей метод є менш поширеним і відповідно менш досвіду у його використанні.

Аналіз експресії генів за допомогою мікрочіпів заснований на гібридизації мРНК, отриманої з досліджуваної тканини або клітинної лінії, з певними зондами на мікрочіпі. Зонди є короткими фрагментами ДНК або РНК, які представляють конкретні гени або групи генів.

Процедура аналізу експресії генів за допомогою мікрочіпів може бути такою:

- отримання РНК із досліджуваної тканини або клітинної лінії;
- мічення мРНК флуоресцентним нуклеотидом;
- інкубація РНК з мікрочіпом, що містить зонди, які представляють гени.
- вимивання незв'язаних РНК та сканування мікрочіпа для визначення рівня гібридизації РНК з кожним зондом.
- аналіз даних, отриманих у результаті сканування, визначення рівня експресії кожного гена у зразку.
- порівняння рівнів експресії генів в експериментальній та контрольній групах дає змогу виявити відмінності в експресії генів, пов'язані з певними біологічними процесами чи патологіями [14].

Розробка мікрочіпа для аналізу експресії генів включає в себе планування та дизайн мікрочіпа. На цьому етапі визначають цілі дослідження та вибирають підхід до вибору зондів для мікрочіпа. Також визначають кількість генів, які будуть аналізуватись, та кількість повторень кожного зонда на мікрочіпі. Дизайн мікрочіпа включає розміщення зондів на поверхні чіпа та розробку протоколів гібридизації та сканування.

Зонди можуть бути розроблені на основі послідовностей генів, отриманих з геномних баз даних, або на основі анотованого геному та бути синтезовані методами фотохімічного синтезу коротких фрагментів ДНК або

РНК, зонди можуть бути розроблені на основі послідовностей генів, отриманих з геномних баз даних, або на основі анотованого геному [15].

Зазвичай користуються стандартними чіпами фірм Affymetrix і Illumina, а зазначену вище технологію використовують для спеціальних цілей.

1.3.2 Отримання даних РНК-секвенування для визначення експресії генів

РНК-секвенування – це техніка, яка використовується для дослідження експресії генів.

Дані РНК-секвенування використовують для кількісного визначення у досліджуваному зразку індивідуальних РНК в абсолютних одиницях.

Збірка транскриптома відбувається або за допомогою референсного геному або *de novo*, без нього [16].

Процес RNA-Seq включає такі етапи:

1) Отримання РНК: першим кроком є отримання РНК з досліджуваної тканини або клітинної лінії. Це може бути здійснено за допомогою методів, таких як виділення тотальної РНК або вилучення лише мРНК. Вилучення мРНК з усіх РНК здійснюється за допомогою полі(А). Цей метод використовує специфічне зв'язування полі(А)-хвостів мРНК з олігодТ-матрицями. Після зв'язування некодуючі РНК та інші компоненти видаляються, а мРНК зберігаються для подальшого аналізу.

2) Приготування бібліотеки: з отриманої РНК створюють бібліотеку. На матриці РНК за допомогою зворотної транскриптази синтезують комплементарний ланцюг ДНК (кДНК). На матриці отриманої ДНК синтезують комплементарний ланцюг ДНК. За допомогою ДНК-лігази до ДНК приєднуються адаптери – маленькі фрагменти ДНК, із двох олігонуклеотидів, частково комплементарних один одному. Це необхідно для забезпечення стійкості фрагментів в процесі секвенування, а також для

забезпечення приєднання фрагментів на пластину секвенатора для їх ампліфікації та секвенування наступними кроками.

3) Секвенування. Найбільш поширеними методами є метод Sanger sequencing та метод на основі масиву, такий як Illumina sequencing. У випадку методу Illumina отримана бібліотека з роз'єднаними фрагментами секвенується паралельно, утворюючи мільйони коротких рідів [17].

Довжина риду варіюється від 25 до 200 нуклеотидів залежно від обраного методу секвенування.

Дані підрахунку рідів після РНК-секвенування, представляють собою кількість зчитувань послідовності, яка походить від певного гена.

Ріди картуються (або вирівнюються) на референсний геном.

Найпростіший підхід – розглядати лише однозначно вирівняні ріди на анотовані моделі геномів.

Анотовані моделі геномів є результатом процесу анотації геному, тобто ідентифікації різних функціональних елементів, таких як гени, інтрони, екзони, промотори та інші, у геномній послідовності.

Метою диференційного аналізу експресії генів є визначення тих з них, які по-різному експресуються за різних умов функціонування тканини, клітини тощо (у нашому випадку між першим і другим триместрами фізіологічної вагітності) шляхом порівняння профілів РНК між двома зразками.

Використовуючи такий аналіз, можна відповісти на такі питання, як:

- Чи є якась різниця в експресії генів у різні часові проміжки або за різних умов функціонування тканини, клітини?
- Які групи генів змінюються подібним чином з часом або в різних умовах?

Першим кроком у робочому процесі визначення диференційної експресії є нормалізація даних, які порівнюються.

Нормалізація даних – це практика організації записів даних для забезпечення їх схожості в усіх полях і записах, що полегшує пошук, групування й аналіз інформації.

Наступним кроком є видалення технічної варіативності між даними, batch effect, з урахуванням біологічної варіативності [18]:

- кількість клітин у зразку;
- загальну ефективність виділення РНК;
- ефективність виділення молекул РНК;
- ефективність гібридизації;
- точність та чутливість вимірювання сигналу.

1.4 Функціональний аналіз

Вихідним результатом аналізу диференційної експресії Microarray, RNA-Seq, а також інших генетичних/геномних аналізів є список значущих генів. Для більш глибокого аналізу отриманих результатів визначають:

- чи є збагачення відомих біологічних функцій, взаємодій або шляхів;
- кластеризацію генів за їх участі у шляхах передачі інформації або мережах. Таким чином, створюються «модулі» співвиразних генів, які часто відображають групи функціонально подібних генів.

1.4.1 Диференційно експресовані гени

Диференційно експресовані гени – це гени, які експресуються по-різному у двох або більше групах зразків. Диференційний аналіз експресії генів – це статистичний метод, який використовується для ідентифікації генів з різною експресією. Диференційний аналіз експресії генів є потужним інструментом для вивчення експресії генів у різних умовах, включаючи діагностику захворювань, відкриття ліків і фундаментальні дослідження.

Ця інформація може бути використана для розробки діагностичних тестів і цільового лікування.

У даній роботі були використані дані про диференційно експресовані гени в плаценті людини між першим і другим триместрами фізіологічної

вагітності. Головною метою було розкрити, як саме ці гени експресуються в нормальних умовах без наявності патологій. Це забезпечує базове порівняння для подальшого аналізу експресії генів у плаценті людини з прееклампсією.

Також це важливо для визначення нормальної молекулярної рівноваги в процесі вагітності.

1.4.2 Функціональний аналіз експресованих генів. Використання бази даних String-db

У той час як про окремі білки, про їх структуру і функції накопичено багато інформації, про їх взаємодії є лише фрагментарна інформація.

Деякі взаємодії дуже добре задокументовані і зрозумілі, наприклад, в контексті тривимірних реконструкцій великомасштабних клітинних механізмів, тоді як для інших взаємодій доступні лише генетичні спостереження або статистичні прогнози.

Крім того, поле потенційних білок-білкових взаємодій набагато ширше і залежне від контексту, ніж поле природних молекулярних функцій окремих молекул. Мало того, що взаємодії обмежені певними типами клітин і фізіологічними умовами, їх специфічність та інтенсивність також можуть варіювати від високоспецифічних і стабільних до відносно неспецифічних.

Тому бажано зібрати всі типи білок-білкових взаємодій і об'єднати їх в єдину конструкцію. Це підтримає конвеєри аналізу даних у різних сферах – від ідентифікації модулів захворювання до відкриття біомаркерів, і дасть змогу переглядати анотації вручну.

Дані про білок-білкові взаємодії можна збирати з багатьох онлайн-баз даних, а також за допомогою окремих високопродуктивних інструментів.

База даних STRING (Search Tool for Interacting Genes/Proteins) [19] виділяється, головним чином:

- широким охопленням;
- простотою використання.

Наразі вона містить найбільше інформації про різні організми (5090) і білки (24,6 млн), має дуже великі і різноманітні джерела порівняльних даних, а також пропонує інтуїтивно зрозумілий і швидкий перегляд в режимі онлайн.

Веб-сайт надає право користувачам входити в систему для пошуку і збереження наборів генів, а також містить програму, яка дає змогу переглядати докази, що лежать в основі будь-якої взаємодії.

Цей ресурс також надає користувачам інформацію про вхідні та пошукові дані, включаючи статистику збагачення мережі та виявлення функціонального збагачення (gene set enrichment analysis).

База даних STRING нещодавно отримала статус Європейського основного ресурсу даних від ELIXIR, загальноєвропейської ініціативи з біоінформатики, присвяченої створенню стійкої біоінформатичної інфраструктури.

Наслідком цього статусу є те, що всі дані про взаємодію та допоміжна інформація в базі даних STRING тепер є у вільному доступі без обмежень.

Основною одиницею взаємодії в STRING є "функціональна спорідненість", тобто зв'язок між двома білками, які спільно виконують певну біологічну функцію; два білки не обов'язково повинні фізично взаємодіяти, щоб бути пов'язаними таким чином. Згідно з цим визначенням, навіть взаємно антагоністичні білки, такі як інгібітори та активатори одного і того ж шляху, можуть бути функціонально пов'язані.

Всі докази зв'язку в базі даних STRING класифікуються за одним із семи незалежних "каналів". Три канали прогнозування базуються на інформації про геномний контекст, інші канали для (i) коекспресії, (ii) текстового аналізу, (iii) біохімічних/генетичних даних ("експериментальний") і (iv) інформації про попередньо відібрані шляхи і білкові комплекси.

Користувач може активувати всі канали окремо або в будь-якій комбінації. Для кожного каналу доступна окрема оцінка взаємодії. Загалом, оцінки взаємодії STRING не відображають силу або специфічність

конкретної взаємодії, а скоріше мають на меті виразити приблизну впевненість у достовірності асоціації за шкалою від нуля до одиниці, беручи до уваги всі наявні докази. З цією метою карти шляхів KEGG використовується як золотий стандарт і, таким чином, вони також неявно визначають деталізацію функціональної асоціації.

STRING використовує ієрархічно впорядковані ортологічні групові зв'язки, визначені за допомогою eggNOG (evolutionary genealogy of genes: Unsupervised Orthologous Groups), коли це можливо, для передачі взаємозв'язків між організмами.

Канали коекспресії ґрунтуються на тестах кореляції генів на великій кількості наборів даних про експресію генів (з використанням як транскриптома, так і протеома). У випадку транскриптомних даних STRING обробляє і мапує велику кількість експериментів, що зберігаються в NCBI Gene Expression Omnibus, які потім нормалізуються [20].

1.4.3 Візуалізація даних. Побудова volcano plot для диференційно експресованих генів

Зазвичай для будь-якого аналізу диференційної експресії корисно інтерпретувати отримані списки генів за допомогою безкоштовних інструментів на основі веб- та R-технологій.

Volcano plot – це графічний інструмент, який зазвичай використовують в аналізі рівнів експресії мРНК, отриманих за допомогою технології мікрочипів. Volcano plot є діаграмою розсіювання, де кожна точка представляє ген, а координата x є логарифмом-зміни кратності (Fold change), а координата y є t-статистикою або $-\log(p\text{-value})$ з t-тесту.

p-value – ймовірність отримати значення випадкової величини для даної ймовірнісної моделі розподілу.

Fold change – це математичне значення описує як кількісно відрізняються дані експресії генів між собою.

У статистичних методах (t-test), заснованих на перевірці гіпотез, р-значення є основою відбору генів

Причина популярності volcano plot в аналізі даних мікрочипів пов'язана з його одночасним відображенням двох частин інформації – зміни кратності та t-статистики. Ранжування генів за кратною зміною та за t-тестом не обов'язково приводить до того самого порядку в списку диференційно експресованих генів і може призвести до різних біологічних висновків.

Одна з переваг volcano plot в аналізі мікрочипів полягає в тому, що він забезпечує «подвійне фільтрування» за допомогою $\lgFC -\log(p\text{-value})$ [21].

Отже, volcano plot – це тип діаграми розсіювання, що об'єднує статистичний показник значущості (p-value), отриманий у результаті статистичного тесту, з величиною кратності зміни (FC, Fold change) експресії генів і використовується для швидкого виявлення найбільш виражених і надійніших змін у великих наборах повторюваних даних [22, 23, 24]. Повторювані дані – це дані, які містяться у декількох записах або документах.

Основне використання volcano plot полягає в дослідженні генів, які можуть бути відібрані за одним критерієм диференційної експресії, але не за іншим. Це – так зване «подвійне фільтрування».

Гени, що не пройшли подвійну фільтрацію, не обираються з поважних причин, оскільки гени з великими значеннями зміни кратності, але з незначним результатом t-тесту мають низьку значущість, а гени зі значним результатом t-тесту та низькою зміною кратності мають сумнівне біологічне обґрунтування значущості (Рисунок 1).

Тема преєклампсії та її лікування ще потребує додаткового вивчення. На жаль, преєклампсію часто виявляють лише після 20-го тижня вагітності, що підкреслює необхідність подальших досліджень у цій області.

Розглянувши наведені вище джерела, можна зробити висновок, що за допомогою методу volcano plot можна визначити диференційно експресовані гени з найбільшою значущістю одразу за двома критеріями. Тому метою

нашої роботи було дослідження генів, які виявляють різну експресію між першим і другим триместрами фізіологічної вагітності. Це дасть змогу зрозуміти, які саме гени мають відмінну експресію в цих двох станах для здорової плаценти. Крім того, ці результати можна буде порівняти з патологічними станами для більш детального розуміння причин прееклампсії та можливостей ранньої діагностики.

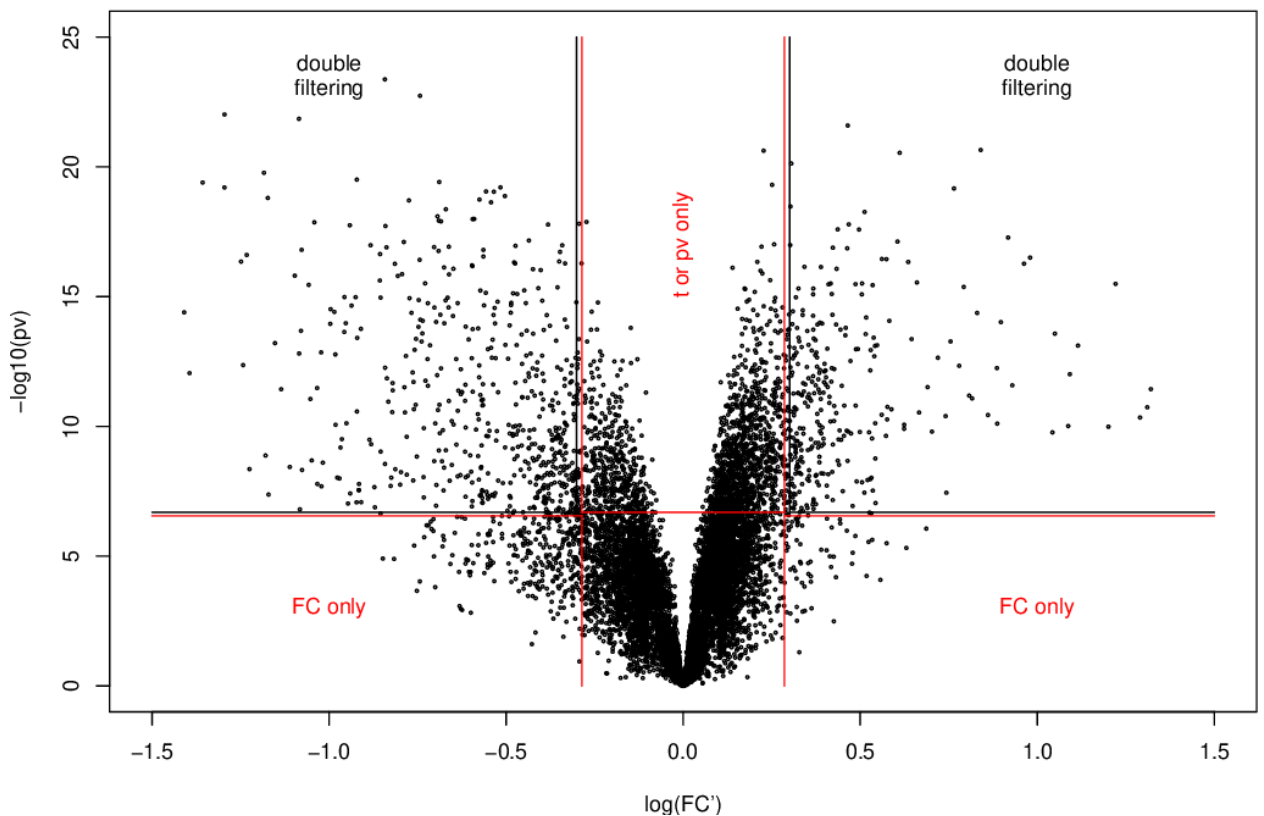


Рисунок 1 – Ілюстрація подвійної фільтрації на діаграмі volcano plot - верхній лівий і верхній правий кут; фільтрування лише за значенням FC (Fold Change) (нижній лівий і нижній правий кути) і фільтрування за t-тестом – прямокутна ділянка посередині [23].

2 МАТЕРІАЛИ І МЕТОДИ

2.1 Матеріали досліджень

Вихідні дані [21] для побудови volcano plot включають результати порівняльного дослідження експресії генів у плаценті людини між першим та другим триместром вагітності. Дані представлені таблицею переліку 3912 диференційно експресованих генів зі значеннями логарифму кратності змін експресії при основі 2 для кожного гена і скорегованого p-value з урахуванням поправки на множинну перевірку гіпотез [21].

2.2 Методи досліджень

Побудова діаграми volcano plot і визначення топ 20 найбільш значущих диференційно експресованих генів

Перш ніж побудувати volcano plot, необхідно зауважити, що вісь абсцис повинна відповідати lgFC – логарифму зміни кратності чого при основі 2, а вісь ординат – $(-\log(p\text{-value}))$. Отже, необхідно розрахувати від'ємний логарифм скорегованого p-value у програмі Microsoft Excel для усіх генів таблиці.

Після побудови графіку діаграми також доцільно відібрати 20. Для досягнення цієї цілі скористуємось таким критерієм відбору як π -value, який об'єднує lgFC і $(-\log(p\text{-value}))$.

Існує різниця між статистичним і біологічним значенням диференційної експресії, яку необхідно правильно оцінювати для оцінки та ранжування генів.

У деяких біологічних додатках гени відбирають на основі величини зміни експресії гена (використовуючи 2-кратну зміну як точку відсікання), а варіації експресії гена не враховують.

Деякі ж статистичні підходи намагаються знайти "інформативні гени", рівні експресії пов'язані зі змінами в класифікації або стані об'єктів, що досліджуються.

Ці підходи, як правило, виявляють не всі диференційовано експресовані гени, а лише деякі гени, які відрізняються за значеннями або FC (Fold change), або p-value. Для виявлення всіх диференційно експресованих генів зазвичай перевіряють гіпотези $\mu_0 = \mu_1$ (експресія гена не відрізняється між двома станами) і $\mu_0 \neq \mu_1$ (експресія гена відрізняється між двома станами).

Щоб вирішити цю проблему, зазвичай комбінують p-значення з критеріями FC (наприклад, 2-кратна зміна, $P \leq 0,05$).

Враховуючи вищезазначені труднощі, Юфей Сяо (Yufei Xiao) та ін. [28] запропонували показник генетичної значущості, який називається π -значення, що об'єднує FC та p-значення в єдиний показник.

Це забезпечує компроміс між FC і p-значеннями і пропонує новий спосіб ранжування і відбору генів. Його унікальними внесками є:

- π -значення не вводить нових статистичних тестів і є простим для обчислення;
- коригування кількості відібраних генів, що експресуються по-різному, шляхом зміни порогового значення є зручним і не потребує перерахунку. π також корисний для ранжування генів, оскільки він присвоює значення кожному гену на основі значень FC і p [28].

Код для розрахунку π -значення, відбору топ-20 генів найбільш значущих диференційно експресованих генів та побудови volcano plot

Код було розроблено у середовищі R.

R – це мова та середовище для статистичних обчислень і графіки. Це проєкт GNU (операційна система типу Unix), подібний до мови та середовища S, розробленого в Bell Laboratories Джоном Чемберсом та його колегами [30]. R можна розглядати як іншу реалізацію S. Є деякі важливі відмінності, але багато коду, написаного для S, працює без змін під R.

R надає широкий спектр статистичних (лінійне та нелінійне моделювання, класичні статистичні тести, аналіз часових рядів, класифікація, кластеризація) і графічних методів, а також дуже розширюваний. Мова S

часто є засобом вибору для досліджень статистичної методології, а R забезпечує відкритий вихідний шлях для участі в цій діяльності.

Однією з сильних сторін R є легкість, з якою можна створювати якісні графіки для публікацій, включаючи математичні символи та формули, де це необхідно. Велику увагу приділено параметрам за замовчуванням для незначних варіантів дизайну графіки, але користувачеві віддається повний контроль [30].

Було завантажено пакети з бібліотеки R, які використовували для створення графіку volcano: ggplot2, readxl, ggrepel:

```
install.packages("ggplot2", "readxl", "ggrepel") [30].
```

Потім пакети було відкрито: library(readxl), library(ggplot2), library(ggrepel); відкрили потрібний файл із даними про експресію генів, необхідними для побудови volcano plot, збережено як змінну «full»:

```
full <- read_excel("C:/Users/difexp_1_2.xlsx", sheet = "difexp 2_1 full");
```

Групи даних були позначені відповідно до значень $|\lg FC|$ та скоригованого p-value:

```
full$Expression <- "NOT SIGNIFICANT" – гени, у яких  $|\lg FC| < 1$  і  $\text{adj.p.val} > 0.05$ ;
```

```
full $Expression[full $|lgFC| > 1 & full $adj.p.val < 0.05] <- "UP", full  
$Expression[full $|lgFC| < -1 & full $adj.p.val < 0.05] <- "DOWN";
```

Обрали кольори для міток:

```
mycolors <- c("blue", "red", "grey"); names(mycolors) <- c("DOWN", "UP",  
"NOT SIGNIFICANT");
```

Створили пусту змінну, в якій зберегли підписи точок на графіку: full \$delabel <- NA.

Розрахували π -value для всіх генів: full\$ π .value <- abs(-log10(full \$adj.p.val)* full\$|lgFC|).

Розташували гени у порядку від найбільшого π -value до найменшого: full <- full [order(full \$ π .value, decreasing=TRUE),].

Перші двадцять рядків таблиці (які після попередніх операцій містили топ-20 генів) записали у окрему змінну та підписали тільки їх:

```
top20_full <- full [1:20,];
```

```
top20_full $delabel <- top20_full $SYMBOL;
```

та побудували сам графік:

```
p1 <- ggplot(data=full, aes(x=|lgFC|, y=-log10(adj.P.Val), col=Expression,
label=delabel))+geom_point() – нанесли на графік точки;
```

```
+geom_point(data=top20_full, shape=1, col="black", fill=rgb(0,0,0,
alpha=0), stroke=1, size=2.5) – точки для топ-20 генів;
```

```
+theme_minimal() – дизайн графіку мінімальний;
+geom_vline(xintercept=c(-1, 1), col="red", linetype=2) – червоні вертикальні
лінії для наочності; +geom_hline(yintercept=-log10(0.05), col="red", linetype=2)
– червона горизонтальна лінія для наочності;
```

```
+geom_text_repel(data=top20_full, col="black", max.overlaps=20, force =
4, force_pull=1) – підписи для топ-20 генів;
```

```
+scale_color_manual(values=c("blue","black","red")) – налаштування
кольорів;
```

```
ggsave("plot.jpeg",plot=p1,width=10 height=8,path="C:\\Users\\Desktop") –
зберегли зображення у форматі jpeg на робочому столі комп'ютера.
```

В кінці роботи, коли графік вже було побудовано, таблицю з топ-20 диференційно експресованими генами зберегли для подальших досліджень:

```
full <- full [order(full $π.value, decreasing=TRUE),]; library(xlsx) df=top20_full;
write.xlsx(df, "top20_full.xlsx").
```

Створення стовпчикового графіку для таблиці топ 20 найбільш значущих диференційно експресованих генів

Стовпчиковий графік (barplot) створено за допомогою бібліотек PANDAS та SEABORN у середовищі розробки для мови Python.

Імпортували необхідні бібліотеки:

- import pandas as pd
- import seaborn as sns

```

– import matplotlib.pyplot as plt
df = pd.read_excel('/content/top20.xlsx') – відкрили таблицю з топ-20 генами
і записали як змінну.
palette = ['blue','red'] – записали необхідні кольори в змінну,
p=sns.barplot(data=df,      x='SYMBOL',      y='lgFC',      hue='Expression',
palette=palette, width=1.2) – створили графік,
p.set_xticklabels(p.get_xticklabels(), rotation=90) – налаштували підписи
генів.

```

Подальший аналіз диференційно експресованих генів передбачав *визначення функцій отриманих генів*.

На веб-сайті бази даних String-db ввели повний список топ-20 генів та отримали розписані функції для кожного білку, що кодується кожним з цих генів, та загальну картину білок-білкових взаємодій.

Після аналізу взаємодій та короткого функціонального аналізу, було доцільно більш детально проаналізувати гени, які мають докази функціонального зв'язку між собою. Для цього використали базу даних Human Protein Atlas [32].

Проект Human Protein Atlas (HPA) – це широкомасштабна ініціатива, спрямована на картографування всього протеома людини за допомогою протеоміки на основі антитіл та інтеграції різних інших omics-технологій.

Загальнодоступний ресурс знань є однією з найбільш відвідуваних у світі біологічних баз даних, яка була значно оновлена протягом останніх кількох років.

Поточна версія поділена на шість основних розділів, кожен з яких зосереджений на конкретних аспектах протеома людини:

- атлас тканин, що показує розподіл білків у всіх основних тканинах і органах людського тіла;
- клітинний атлас, що показує субклітинну локалізацію білків в окремих клітинах;

- атлас патології, що показує вплив рівня білка на виживання пацієнтів з раком;
- атлас крові, що показує профілі експресії клітин крові та активно секретованих білків;
- атлас мозку, що показує розподіл білків у мозку людини, миші та свині та
- метаболічний атлас, що показує участь білків у метаболізмі людини.

НРА є важливим ресурсом для подальшого розуміння біології людини, а загальнодоступні набори даних мають багато перспектив для інтеграції з іншими новими зусиллями, спрямованими на аналіз однієї клітини, як на транскриптомному, так і на протеомному рівнях.

Основні висновки з кожного з шести різних підатласів вичерпно узагальнено на цільових сторінках із малюнками, таблицями, списками та прикладами, які можна натиснути, для швидкого доступу до результатів, пов'язаних з кожним аспектом протеома людини. Користувачі НРА також можуть комбінувати персональні розширені пошукові запити на основі, наприклад, загальних даних про гени та білки, класів білків, специфічної експресії тканини або клітини та критеріїв підтвердження, використовуючи інформацію з кожного з різних підатласів. Результати цих запитів представлені у вигляді геноцентричних списків, які можна натиснути та завантажити в різних форматах.

Інший підхід до доступу до обширних наборів даних, знайдених у НРА, полягає в 26 різних файлах, які можна завантажити, які містять дані про весь геном у різних аналізах, що дає змогу проводити широкомасштабні біоінформатичні дослідження. НРА докладає значних зусиль для охоплення, а база даних підтримується завдяки внеску спільноти в європейську інфраструктуру ELIXIR, де НРА вказано як один із основних ресурсів даних, важливих для ширшої спільноти наук про життя. Навчальний матеріал НРА також містить гістологічний словник, який є корисним інструментом як для

студентів, так і для науковців, зацікавлених у отриманні додаткових знань про різні структури тканин і клітин, які формують основу для моделей експресії білків, характерних для тканин і клітин [33].

3 РЕЗУЛЬТАТИ І ОБГОВОРЕННЯ

3.1 Побудова діаграми volcano plot на основі даних про диференційну експресію генів у плаценті людини між першим і другим триместрами вагітності

За даними про диференційну експресію генів у плаценті людини між першим і другим триместрами вагітності методом, описаним в підрозділі 2.1.2, було побудовано діаграму volcano plot (Рисунок 2).

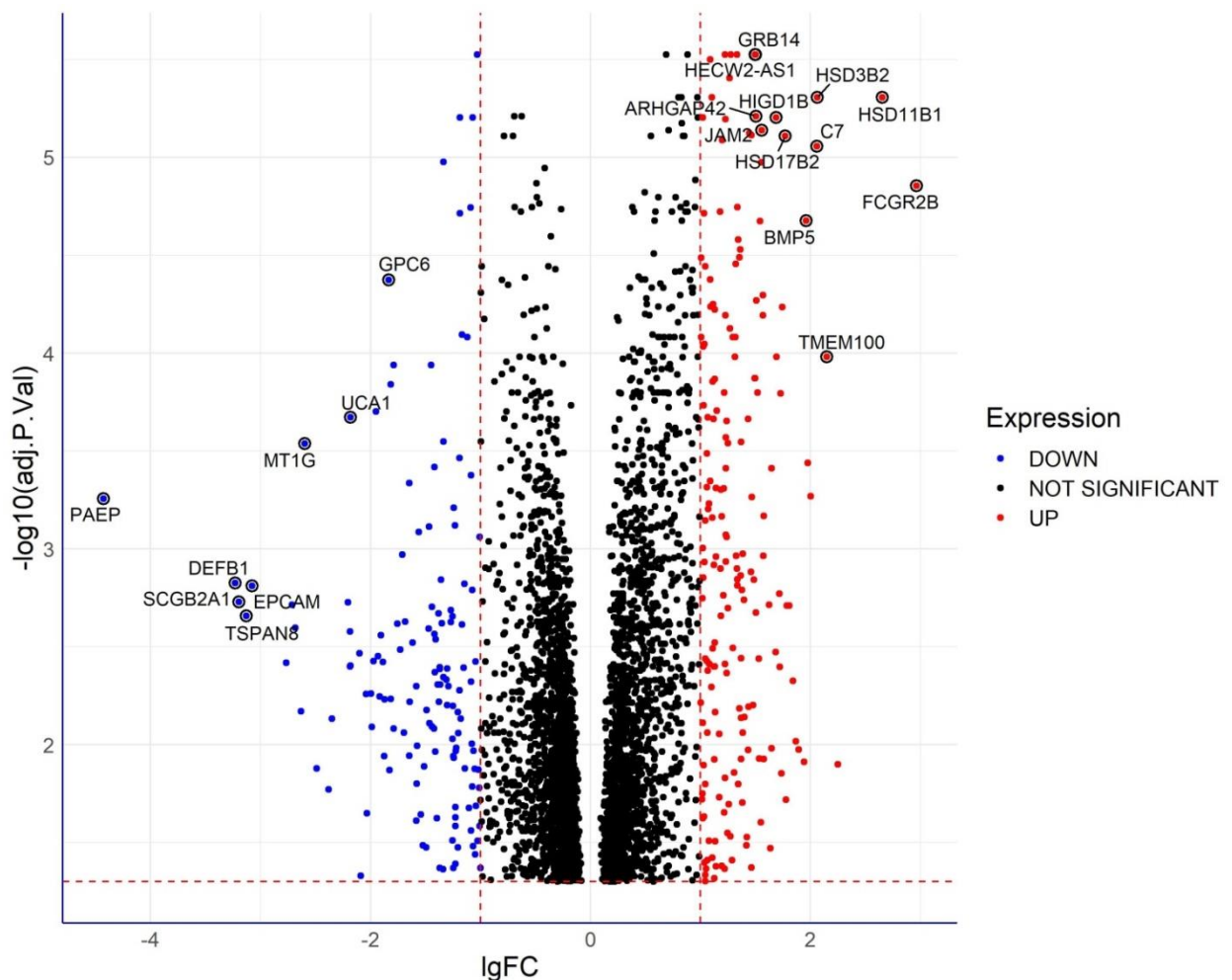


Рисунок 2 – Volcano plot на основі даних про диференційну експресію генів у плаценті людини між першим і другим триместрами вагітності:

вісь x – $\lg FC$, вісь y – $(-\log(\text{adjusted p-value}))$;

підписані точки – найбільш значущі диференційно експресовані гени.

Сині точки – це гени з $\lgFC < (-1)$, тобто зміна кратності експресії генів між другим і першим триместрами фізіологічної вагітності – від’ємна, тоді як червоні точки – це гени з $\lgFC > 1$, тобто зміна кратності експресії – додатня. Обидві групи генів задовольняють умові $-\log(\text{adjusted } p\text{-value}) < 0,05$.

Значення $p\text{-value } 0,05$ вважається стандартним. Ми відбираємо гени, для яких $p\text{-value} < 0,05$, тому що є достатньо доказів вважати, що спостережуваний ефект не може бути пояснений випадковістю.

Значення $\lgFC < (-1)$ і $\lgFC > 1$, були обрані тому що для цих генів, експресія змінилася більше ніж вдвічі між двома групами.

Так ми відсіяли незначущі зміни експресії.

Найбільш значущі гени відбирали за $\pi\text{-value}$, з найбільшими значеннями.

3.2 Розрахунок оцінки значущості для Топ-20 диференційно експресованих генів між 1 та 2 триместрами вагітності та побудова стовпчикової діаграми за показником $\pi\text{-value}$

Метод volcano plot дав змогу виявити 20 генів з найвищими абсолютними значеннями $|\lgFC|$ та $-\log(\text{скорегованого } p\text{-значення})$, які відрізнялися між другим та першим триместрами фізіологічної вагітності (Таблиця 1). Для чисельного розрахунку оцінки значущості гена було використано такі параметри:

- $\lgFC$ – логарифм кратності змін рівня експресії гена;
- $p\text{-value}$ – ймовірність отримати значення випадкової величини;
- $\text{adj.}p\text{-val}$ – скореговане значення $p\text{-value}$ з урахуванням поправки на множинну перевірку гіпотез. Поправка на множинну перевірку гіпотез використовується для контролювання ризику зробити помилкові висновки при виконанні багатьох статистичних перевірок. Ця проблема виникає, коли ми проводимо багато незалежних перевірок гіпотез.
- $-\log(\text{adj.}p\text{-val})$ – від’ємний десятковий логарифм скорегованого $p\text{-value}$;

- π -value – оцінка значущості гена, яка об'єднує FC і P-значення в одну оцінку.

Таблиця 1. Найбільш диференційно експресовані гени між 1 та 2 триместрами вагітності. $-\log(\text{adj.p-value}) > 0.05$, $|\lg FC| > 1$

SYMBOL – позначення гену, $\lg FC$ – зміна кратності рівня експресії, $\text{adj.P.Val} \cdot 10^3$ – скореговане значення p-value, помножене на 10^3 , $\log(\text{adj.P.Val})$ – десятковий логарифм скорегованого p-value.

Таблиця 1 – Найбільш диференційно експресовані гени між 1 та 2 триместрами вагітності за умови $-\log(\text{adj.p-value}) > 0.05$, $|\lg FC| > 1$

SYMBOL	$\lg FC$	$\text{adj.P.Val} \cdot 10^3$	$\log(\text{adj.P.Val})$	π-value
PAEP	-4,4275	0,5538	3,2567	14,4188
FCGR2B	2,9658	0,0140	4,8546	14,3977
HSD11B1	2,6556	0,0049	5,3066	14,0921
HSD3B2	2,0620	0,0049	5,3066	10,9422
C7	2,0594	0,0088	5,0576	10,4159
MT1G	-2,5977	0,2900	3,5375	9,1894
BMP5	1,9623	0,0211	4,6761	9,1758
DEFB1	-3,2314	1,4899	2,8269	9,1347
HSD17B2	1,7693	0,0078	5,1081	9,0377
HIGD1B	1,6871	0,0063	5,2038	8,7791
SCGB2A1	-3,1968	1,8635	2,7297	8,7263
EPCAM	-3,0775	1,5456	2,8109	8,6506
TMEM100	2,1482	0,1046	3,9806	8,5512
TSPAN8	-3,1291	2,1922	2,6591	8,3208
GRB14	1,5004	0,0030	5,5252	8,2902
HECW2-AS1	1,4961	0,0030	5,5252	8,2664
GPC6	-1,8323	0,0423	4,3740	8,0142
UCA1	-2,1809	0,2128	3,6719	8,0081

JAM2	1,5564	0,0073	5,1375	7,9961
ARHGAP42	1,5029	0,0062	5,2103	7,8303

Стовпчиковий графік дав змогу візуалізувати значення $\lg FC$ для кожного гена (Рисунок 3).

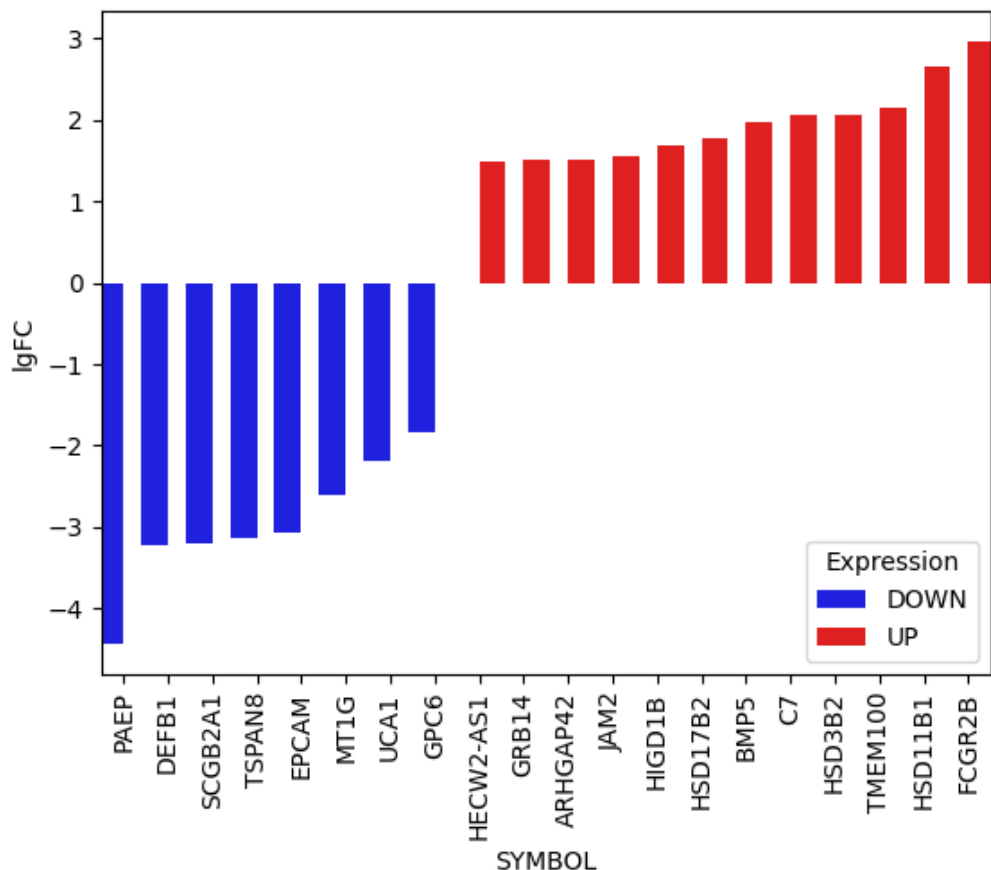


Рисунок 3 – $\lg FC$ Топ-20 диференційно експресованих генів між 1 та 2 триместрами вагітності

Він відображає значення $\lg FC$ для кожного з 20 генів. Синім кольором позначені гени, для яких $\lg FC < (-1)$, зміна експресії – від’ємна; червоним кольором позначені гени, для яких $\lg FC > 1$, зміна експресії – додатня. Отже, стовпчикова діаграма наочно показує, що між 1 і 2 триместрами вагітності експресія деяких генів у плаценті суттєво змінюється: деякі гени зменшують свою експресію, інші – збільшують.

3.3 Визначення функцій білків та їх асоціацій для Топ-20 диференційно експресованих генів між 1 та 2 триместрами вагітності

Для розуміння молекулярних механізмів дії відповідних білків за допомогою набору інструментів String-db (functional protein association networks) було проаналізовано функції генів та білок-білкові взаємодії (Рисунок 4).

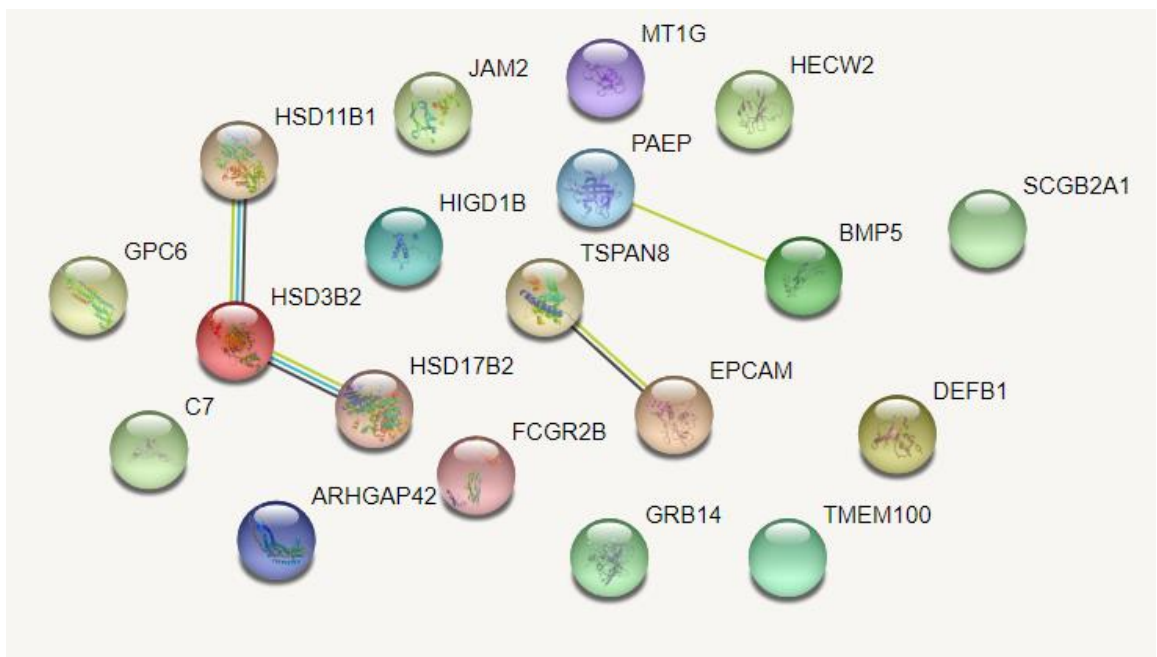


Рисунок 4 – Білки, що кодуються отриманими генами, та їх асоціація: кульки мережі представляють білки, лінії являють собою білок-білкові асоціації

Розглянемо детальніше кожен білок-білкові асоціацію

1) HSD11B1 – HSD3B2

Таблиця 2 – Докази, що свідчать про функціональний зв'язок білків HSD11B1 – HSD3B2

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,062).
Асоціація в базах даних:	так (оцінка 0,900).

Спільно згадується в рефератах Pubmed:	так (оцінка 0,444). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,066).
Загальний бал:	0,944

2) HSD17B2 – HSD3B2

Таблиця 3 – Докази, що свідчать про функціональний зв'язок білків HSD17B2 – HSD3B2

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,061).
Асоціація в базах даних:	так (оцінка 0,900).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,588). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,130).
Загальний бал:	0,961

3) BMP5 – PAEP

Таблиця 4 – Докази, що свідчать про функціональний зв'язок білків BMP5 – PAEP:

Спільно згадується в рефератах Pubmed:	так (оцінка 0,412).
Загальний бал:	0,412

4) EPCAM – TSPAN8

Таблиця 5 – Докази, що свідчать про функціональний зв'язок білків EPCAM – TSPAN8:

Спільне вираження:	так (оцінка 0,119). Крім того, передбачувані гомологи генів коекспресуються в інших
--------------------	---

	організмах (оцінка 0,061).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,450). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,104).
Загальний бал:	0,537

Отже, загалом лише для малої кількості білків є докази функціонального зв'язку. Тому ми можемо додати інші білки і побачити більш повну картину білок-білкових асоціацій. String-db дає змогу це зробити натисканням кнопки «+More» (Рисунок 5).

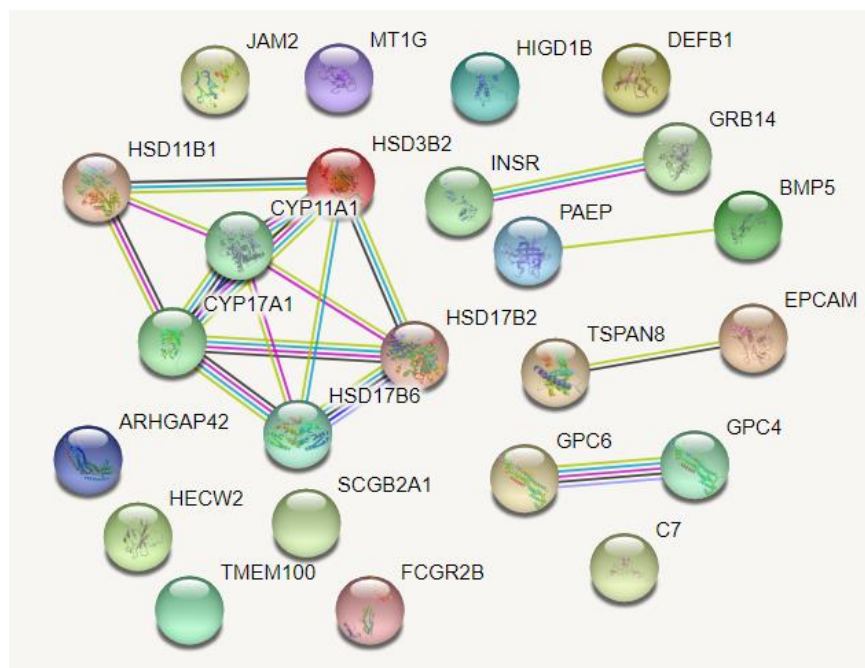


Рисунок 5 – Доповнене зображення білків, що кодуються отриманими генами, та їх асоціація: кульки мережі представляють білки, лінії являють собою білок-білкові асоціації

На рисунку можна побачити, що білків HSD11B1, HSD3B2 і HSD17B2 мають зв'язок з CYP17A1 – стероїдом, що задіяний у перетворенні прегненолону та прогестерону в їхні 17-альфа-гідроксильовані продукти, а потім у дегідроепіандростерон (DHEA) та андростендіон. Каталізує як 17-

альфа-гідроксилювання, так і 17,20-ліазну реакцію. Бере участь у статевому розвитку під час внутрішньоутробного розвитку та під час статевого дозрівання.

Проаналізуємо детальніше їх функціональний зв'язок.

1) HSD11B1 – CYP17A1

Таблиця 6 – Докази, що свідчать про функціональний зв'язок білків HSD11B1 – CYP17A1

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,062).
Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії в інших організмах (оцінка 0,092).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,557). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,050).
Загальний бал:	0,593

2) HSD3B2 – CYP17A1

Таблиця 7 – Докази, що свідчать про функціональний зв'язок білків HSD3B2 – CYP17A1

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,055).
Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії в інших організмах (оцінка 0,132).
Асоціація в базах даних:	так (оцінка 0,900).

Спільно згадується в рефератах Pubmed:	так (бал 0,838). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,351).
Загальний бал:	0,989

3) HSD17B2 – CYP17A1

Таблиця 8 – Докази, що свідчать про функціональний зв'язок білків HSD17B2 – CYP17A1

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,064).
Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії в інших організмах (оцінка 0,064).
Асоціація в базах даних:	так (оцінка 0,900).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,587). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,088).
Загальний бал:	0,960

Так само на рисунку бачимо, що білки HSD3B2 і HSD17B2 мають зв'язок з HSD17B6 – 17-бета-гідроксистероїддегідрогеназа типу 6; NAD-залежна оксидоредуктаза з широкою субстратною специфічністю, яка демонструє як окислювальну, так і відновну активність (*in vitro*). Має 17-бета-гідроксистероїддегідрогеназну активність щодо різних стероїдів (*in vitro*). Може перетворювати андростерон в епіандростерон.

То ж, подивимось на ці функціональні зв'язки.

1) HSD3B2 – HSD17B6

Таблиця 9 – Докази, що свідчать про функціональний зв'язок білків HSD3B2 – HSD17B6

Асоціація в базах даних:	так (оцінка 0,900).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,836). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,050).
Загальний бал:	0,983

2) HSD17B2 – HSD17B6

Таблиця 10 – Докази, що свідчать про функціональний зв'язок білків HSD17B2 – HSD17B6:

Спільна поява в геномах:	так (оцінка 0,238).
Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,081).
Асоціація в базах даних:	так (оцінка 0,800).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,679). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,068).
Загальний бал:	0,839

Та всі ці гени: HSD11B1, HSD3B2, HSD17B2 і CYP17A1, HSD17B6 мають зв'язок з CYP11A1 – фермент, що розщеплює бічний ланцюг холестерину, знаходиться у мітохондріях стероїдогенних тканин; каталізує реакцію розщеплення бічного ланцюга холестерину до прегненолону; сімейство цитохрому P450 11.

1) HSD11B1 – CYP11A1

Таблиця 11 – Докази, що свідчать про функціональний зв'язок білків HSD11B1 – CYP11A1

Експериментальні/біохімічні дані:	немає, але передбачувані
-----------------------------------	--------------------------

	гомологи були знайдені у взаємодії в інших організмах (оцінка 0,059).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,557). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,050).
Загальний бал:	0,569

2) HSD3B2 – CYP11A1

Таблиця 12 – Докази, що свідчать про функціональний зв'язок білків HSD3B2 – CYP11A1

Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,062).
Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії в інших організмах (оцінка 0,132).
Асоціація в базах даних:	так (оцінка 0,900).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,836). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,209).
Загальний бал:	0,987

3) HSD17B2 – CYP11A1

Таблиця 13 – Докази, що свідчать про функціональний зв'язок білків HSD17B2 – CYP11A1

Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії
-----------------------------------	---

	в інших організмах (оцінка 0,064).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,570). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,072).
Загальний бал:	0,593

4) CYP17A1 – CYP11A1

Таблиця 14 – Докази, що свідчать про функціональний зв'язок білків CYP17A1 – CYP11A1

Спільна поява в геномах:	так (оцінка 0,413).
Спільне вираження:	немає, але передбачувані гомологи генів коекспресуються в інших організмах (оцінка 0,062).
Асоціація в базах даних:	так (оцінка 0,900).
Спільно згадується в рефератах Pubmed:	так (оцінка 0,950). Крім того, передбачувані гомологи згадуються разом в інших організмах (оцінка 0,427).
Загальний бал:	0,953

5) HSD17B6 – CYP11A1

Таблиця 15 – Докази, що свідчать про функціональний зв'язок білків HSD17B6 – CYP11A1

Експериментальні/біохімічні дані:	немає, але передбачувані гомологи були знайдені у взаємодії в інших організмах (оцінка 0,064).
-----------------------------------	--

Спільно згадується в рефератах Pubmed:	так (оцінка 0,891).
Загальний бал:	0,893

За даними веб-сайту бази даних String-db було складено таблицю функцій білків 20и найбільш диференційно експресованих генів у плаценті людини між 1 та 2 триместрами фізіологічної вагітності (Таблиця 16).

Таблиця 16 – Функції білків 20и найбільш диференційно експресованих генів у плаценті людини між 1 та 2 триместрами фізіологічної вагітності

Ген	Функції
1	2
(+)ARHGAP42	Rho (ро-залежний механізм термінації) ГТФаза-активуючий білок 42; Може впливати на артеріальний тиск, функціонуючи як протеїн, що активує ГТФазу (велика родина ферментів гідролаз, які зв'язують та гідролізують гуанозинтрифосфат (GTP).
(+)BMP5	Морфогенетичний білок кісток 5; Стимулює формування хрящів і кісток; морфогенетичні білки кісток.
(+)C7	компонент комплексу – C7 (C7 має гідрофобні властивості і вставляється в ліпідний бішар клітинної мембрани, де стає високоафінним рецептором C8) ; Компонент комплексу мембранної атаки (MAC) відіграє ключову роль у вродженій та адаптивній імунній відповіді шляхом утворення пор у плазматичній мембрані клітин-мішеней.
(-)DEFB1	Бета-дефензин 1; Має бактерицидну дію. Позитивно регулює рухливість сперматозоїдів і бактерицидну

	активність залежно від CCR6 (рецептор β -хемокінів ссавців класу інтегральних мембранних білків) .
(-)EPCAM	Молекула адгезії епітеліальних клітин; Може діяти як молекула фізичної гомофільної взаємодії між епітеліальними клітинами кишечника та інтраепітеліальними лімфоцитами на епітелії слизової оболонки для забезпечення імунологічного бар'єру як першої лінії захисту від інфекції слизової оболонки.
1	2
(+)FCGR2B	Рецептор Fc-області гамма імуноглобуліну з низькою спорідненістю II-b. Бере участь у різноманітних ефекторних і регуляторних функціях, таких як фагоцитоз імунних комплексів і модуляція вироблення антитіл В-клітинами.
(-)GPC6	Гліпікан-6; Протеоглікан клітинної поверхні, який містить гепарансульфат. Передбачуваний корецептор клітинної поверхні для факторів росту, білків позаклітинного матриксу, протеаз і антипротеаз (за подібністю). Посилює міграцію та інвазію ракових клітин через передачу сигналів WNT5A (Wnt family member 5A, Задіяний у таких біологічних процесах як диференціація, сигнальний шлях Wnt, альтернативний сплайсинг. Локалізований у позаклітинному матриксі. Також секретований назовні).
(+)GRB14	Зв'язаний з рецептором фактора росту білок 14; Адаптерний білок, який модулює з'єднання рецепторних кіназ клітинної поверхні зі специфічними сигнальними шляхами. Зв'язується з активованим рецептором інсуліну

	(INSR) і пригнічує сигнали від нього. Потужний інгібітор стимульованого інсуліном фосфорилювання MAPK3 (фермент сімейства MAPK із групи кіназ, регульованих позаклітинними сигналами). Відіграє вирішальну роль у регуляції транслокації мембрани PDPK1 (Phosphoinositide-dependent kinase-1) у відповідь на стимуляцію інсуліном і служить білком-адаптером для залучення PDPK1 до активованого рецептора інсуліну.
1	2
(+)HECW2-1	Е3 убіквітин-протеїнова лігаза HECW2 (HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2), опосередковує убіквітинування TP73 (Tumor protein p73). Діє для стабілізації TP73 і посилення активації транскрипції за допомогою TP73. Бере участь у регуляції мітотичної метафази/анафазного переходу.
(+)HIGD1B	Цей ген кодує член сімейства доменів гена 1, індукованого гіпоксією (HIG1). Кодований білок локалізується на клітинній мембрані, і його пов'язують з пухлиноутворенням і прогресуванням аденоми гіпофіза. Альтернативний сплайсинг призводить до кількох варіантів транскриптів.
(+)HSD11B1	кортикостероїд 11-бета-дегідрогеназа ізофермент 1; Оборотно каталізує перетворення кортизолу в неактивний метаболіт кортизон.
(+)HSD17B2	Естрадіол 17-бета-дегідрогеназа 2; здатний каталізувати взаємоперетворення тестостерону і андростендіону, а також естрадіолу і естрону.
(+)HSD3B2	3 бета-гідроксистероїддегідрогеназа/дельта 5-->4-

	ізомераза типу 2; 3-бета-HSD є біфункціональним ферментом, який каталізує окисне перетворення дельта(5)-ен-3-бета-гідроксистероїду та окисне перетворення кетостероїдів.
(+)JAM2	Молекула з'єднання B; може відігравати роль у процесах переміщення лімфоцитів до вторинних лімфоїдних органів.
1	2
(-)MT1G	металотіонеїн-1G; металотіонеїни мають високий вміст залишків цистеїну, які зв'язують різні важкі метали; транскрипція цих білків регулюється як важкими металами, так і глюкокортикоїдами .
(-)PAEP	Глікопротеїн, який регулює важливі етапи під час запліднення, а також має імуномодулюючу дію.
(-)SCGB2A1	Маммаглобін-B; може зв'язувати андрогени та інші стероїди, також може зв'язувати естрамустин, хіміотерапевтичний засіб, що використовується для лікування раку простати. Може перебувати під регуляцією транскрипції стероїдних гормонів; відноситься до сімейства секретоглобінових.
(+)TMEM100	Трансмембранний білок 100; Відіграє певну роль під час диференціації ембріонального артеріального ендотелію та морфогенезу судин через рецептор-залежний сигнальний шлях ACVRL1 (Activin A receptor like type 1) після стимуляції кістковими морфогенетичними білками.

(-)TSPAN8	Тетраспанін-8; Тетраспанін 8; належить до сімейства тетраспанінів. Мембранна експресія в залозистих клітинах шлунково-кишкового тракту, жовчного міхура, екзокринної підшлункової залози з просвітною експресією в ниркових каналцях.
-----------	---

Примітка: (+) та (-) позначено відповідно підвищення або зниження експресії генів між другим та першим триместрами фізіологічної вагітності.

Проаналізуємо детальніше кожен ген з груп функціонально пов'язаних білків за допомогою Human Protein Atlas та порівнюємо їх.

Кластер експресії (англ. expression cluster) - це група генів або білків, які мають схожі рівні експресії (тобто виявляють подібний рівень активності) в певних умовах або в конкретних клітинах чи тканинах. Кластер експресії є результатом аналізу геномних або протеомних даних, де гени або білки групуються разом на основі їх подібності в експресії.

HSD11B1

Кластер тканинної експресії у печінці відповідає за гемостаз і ліпідний обмін (головним чином).

Кластер експресії клітинної лінії фібробластів відповідає, в основному, за метаболізм.

Білок, який кодується цим геном, є мітосомальним ферментом, який каталізує перетворення гормону стресу кортизолу в неактивний метаболіт кортизон. Крім того, закодований білок може каталізувати зворотну реакцію, перетворення кортизону в кортизол. Занадто багато кортизолу може призвести до центрального ожиріння, і особлива варіація цього гена була пов'язана з ожирінням і резистентністю до інсуліну у дітей. Мутації в цьому гені та білку H6PD (гексозо-6-фосфатдегідрогеназа (глюкозо-1-дегідрогеназа)) є причиною дефіциту кортизон редуктази. Альтернативний

сплайсинг призводить до кількох варіантів транскриптів, що кодують той самий білок [надано RefSeq, травень 2011 р.]

HSD3B2

Кластер експресії у тканинах надниркової залози відповідає за метаболізм стероїдів (переважно).

Кластер експресії клітинної лінії: Рак шкіри – процеси меланоцитів (переважно).

Білок, який кодується цим геном, є біфункціональним ферментом, який каталізує окисне перетворення дельта(5)-ен-3-бета-гідроксистероїду та окисне перетворення кетостероїдів. Він відіграє вирішальну роль у біосинтезі всіх класів гормональних стероїдів. Цей ген переважно експресується в надниркових і статевих залозах.

Мутації в цьому гені пов'язані з дефіцитом 3-бета-гідроксистероїддегідрогенази II типу. Для цього гена були знайдені альтернативно зрізані варіанти транскриптів [надано RefSeq, жовтень 2009 р.]

HSD17B2

Кластер експресії у тканинах печінці та плаценті відповідає за транспортування через ендоплазматичний ретикулум (переважно).

Кластер експресії клітинної лінії у клітинах сполучної тканини відповідає за організацію ECM (екстрацеллярна матриця).

Каталізує НАД-залежне окислення високоактивних 17бета-гідроксистероїдів, таких як естрадіол (E2), тестостерон (T) і дигідротестостерон (ДГТ), до їх менш активних форм і таким чином регулює біологічну ефективність цих стероїдів. Окислює естрадіол до естрону, тестостерон до андростендіону та дигідротестостерон до 5альфа-андростан-3,17-діону. Також має активність 20-альфа-HSD.

Ці три гени, HSD11B1, HSD3B2 і HSD17B2, відіграють важливу роль у метаболізмі стероїдів в організмі людини. Вони взаємодіють і впливають на

різні процеси, зокрема на розвиток і функціонування статевих органів, гормональний баланс та репродуктивну систему.

CYP17A1 і CYP11A1, які також функціонально пов'язані з вищеперерахованими генами, також кодують ферменти, що відіграють важливу роль у синтезі стероїдів в організмі людини.

TSPAN8

Кластер експресії у тканинах кишечника має невідому функцію (в основному). Мембранна експресія в залозистих клітинах шлунково-кишкового тракту, жовчного міхура, екзокринної підшлункової залози.

Білок, який кодується цим геном, є членом трансмембранної надродини 4, також відомої як сімейство тетраспанінів. Більшість із цих членів є білками клітинної поверхні, які характеризуються наявністю чотирьох гідрофобних доменів.

Білки опосередковують події сигнальної трансдукції, які відіграють роль у регуляції розвитку клітин, активації, росту та рухливості. Цей кодований білок є глікопротеїном клітинної поверхні, який, як відомо, утворює комплекс з інтегринами.

Цей ген експресується в різних карциномах. Для цього гена було знайдено використання альтернативних сайтів поліаденілування [надано RefSeq, липень 2008 р.].

ERCAM

Кластер експресії у тканинах кишечника має невідому функцію (в основному), але може вказувати на важливість цього білка у клітинах кишечника, можливо, в контексті регуляції клітинної адгезії та взаємодії між клітинами.

Кластер експресії клітинної лінії: поглинання ентероцитів (переважно).

Цей ген кодує асоційований з карциномою антиген і є членом сімейства, яке включає принаймні два мембранних білка типу I. Цей антиген експресується на більшості нормальних епітеліальних клітин і карциномах шлунково-кишкового тракту і функціонує як гомотипова кальцій-незалежна

молекула клітинної адгезії. Антиген використовується як мішень для імунотерапії карцином людини.

Мутації в цьому гені призводять до вродженої тафтингової ентеропатії [надано RefSeq, грудень 2008 р.].

Гени TSPAN8 і EPCAM можуть бути пов'язані через їхню роль у клітинній адгезії та взаємодії клітин. В деяких випадках TSPAN8 може взаємодіяти з EPCAM на клітинних мембранах, утворюючи комплекси або впливаючи на функції EPCAM.

Однак, конкретні деталі їхньої взаємодії та функцій в організмі можуть розглядатися в контексті конкретних досліджень і систем, в яких вони вивчаються [29].

BMP5

Кластер тканинної експресії: плацента – вагітність (в основному).

Цей ген кодує секретований ліганд надродина білків TGF-бета (трансформуючий фактор росту-бета). Ліганди цього сімейства зв'язують різні бета-рецептори TGF, що призводить до залучення та активації факторів транскрипції родини SMAD, які регулюють експресію генів. Кодований препропротеїн піддається протеолітичній обробці для вивільнення кожної субодиниці дисульфідно-пов'язаного гомодимеру, який відіграє роль у розвитку кісток і хрящів. Поліморфізм цього гена може бути пов'язаний з остеоартритом у людей. Цей ген диференційовано регулюється в багатьох ракових захворюваннях людини і кодує різні білкові ізоформи, процесінг яких відбувається протеолітичною обробкою подібним чином. [надано RefSeq, липень 2016 р.]

PAEP

Кластер тканинної експресії: фібробласти – організація ECM (в основному).

Цей ген є членом суперродина ядерних ліпокалінів, члени якої мають відносно низьку схожість послідовностей, але мають висококонсервативну структуру екзон/інтрон і тривимірне згортання білка. Більшість ліпокалінів

згруповано на довгому плечі хромосоми 9. Кодований глікопротеїн раніше називали ендометріальним альфа-2-глобуліном, асоційованим з вагітністю, плацентарним білком 14 і глікоделіном, але офіційно його назвали прогестаген-асоційованим білком ендометрію.

В амніотичній рідині, фолікулярній рідині та насінневій плазмі (seminal plasma) репродуктивної системи зустрічаються три різні форми з ідентичними білковими каркасами, але різними профілями глікозилювання. Ці глікопротеїни відіграють чітку та важливу роль у регулюванні середовища матки, придатного для вагітності, а також у визначенні часу та настанні відповідної послідовності подій у процесі запліднення. Альтернативний сплайсинг призводить до кількох варіантів транскриптів [надано RefSeq, жовтень 2015 р.].

Незважаючи на те, що BMP5 і PAEP мають різні функції і експресуються в різних тканинах, можливо, що в деяких випадках вони можуть бути пов'язані або взаємодіяти в певних біологічних контекстах. Однак, конкретні деталі їхнього взаємозв'язку та його функціональне значення вимагають додаткових досліджень і розгляду в конкретних наукових контекстах.

Згрупуємо усі гени за описаними функціями.

Гени, пов'язані з імунною відповіддю та захистом:

- (+)ARHGAP42
- (+)C7
- (+)FCGR2B

Гени, пов'язані з регуляцією росту та розвитку:

- (+)BMP5
- (+)GRB14

Гени, пов'язані з рухом та міграцією клітин:

- (-)DEFB1
- (-)EPCAM
- (-)GPC6
- (+)JAM2

Гени, пов'язані з метаболізмом гормонів та стероїдів:

- (+)HSD11B1
- (+)HSD17B2
- (+)HSD3B2

Гени, пов'язані з гіпоксією та пухлинами:

- (+)HIGD1B
- (-)MT1G

Гени, пов'язані зі структурою клітинної мембрани:

- (+)TMEM100
- (-)TSPAN8

Один з проаналізованих диференційно експресованих генів у плаценті людини між першим та другим триместрами вагітності – BMP5 – найбільше експресується саме в плаценті людини, тому було проведено літературний пошук, який показав, що він може мати відношення до гіпертонії в анамнезі матері, може бути пов'язаний з синдромом HELLP (гемоліз, підвищення рівня печінкових ферментів, пошкодження печінкової паренхіми та тромбоцитопенія) [30]. У дослідженнях також було показано, що рівень експресії цього гену значно змінюється порівняно з нормальним фізіологічним станом [31].

Експресія генів, які мають значення $\lg FC < (-1)$, а саме: PAEP, MT1G, DEFB1, SCGB2A1, EPCAM, TSPAN8, GPC6, UCA1, зменшилась у другому

триместрі, порівняно з першим, а експресія генів, які мають значення $\lg FC > 1$ навпаки збільшилась.

Загалом диференційна експресія наведених вище генів може вказувати на зміни у функціонуванні плаценти між 1 та 2 триместрами фізіологічною вагітності. Детальніші дослідження необхідні для розуміння конкретних наслідків та функцій експресії цих генів і порівняння отриманих результатів з результатами експресії цих і інших генів у аналогічний період за умов патології.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 ЗАГАЛЬНІ ПОЛОЖЕННЯ

4.1.1 До роботи у лабораторії допускаються особи, які перед початком роботи пройшли вступний і первинний інструктажі з охорони праці, спеціальне навчання, перевірку знань вимог охорони праці та стажування.

4.1.2 У процесі роботи працівники повинні проходити у встановленому порядку періодичні медичні огляди, повторні інструктажі з охорони праці, а також позапланові та цільові інструктажі.

4.1.3 Під час роботи на працівників лабораторії можуть впливати такі небезпечні й шкідливі виробничі фактори:

- підвищена загазованість повітря робочої зони;
- підвищене значення напруги в електричному ланцюзі,
- замикання якого може відбутися через тіло людини;
- недостатня освітленість робочої зони;
- підвищений рівень шуму;
- хімічні;
- нервово-психічні перевантаження (емоційні перевантаження).

4.1.4 Працівники лабораторії повинні бути забезпечені засобами індивідуального захисту.

4.1.5 Працівник лабораторії під час роботи повинен:

- використовувати лабораторне обладнання лише за призначенням;
- знати правила користування і способи перевірки справності;
- вміти користуватися первинними засобами пожежогасіння;
- стежити за справністю і цілісністю заземлення (занулення) корпусів електроприладів, електричних машин і устаткування;
- утримувати в чистоті своє робоче місце, в справному стані і чистоті обладнання, інструменти, а також спецодяг, спецвзуття та інші ЗІЗ;

- не приступати до роботи з невідомими речовинами і хімічними реактивами;
- курити у встановленому місці;
- дотримуватися правил внутрішнього трудового розпорядку та режими праці та відпочинку.

4.1.6 Лабораторному персоналу заборонено протягом 24 годин зберігати в лабораторії легкозаймисті рідини.

4.1.7 На посудинах, що працюють під тиском, може працювати лише персонал, який пройшов спеціальну підготовку та отримав відповідний сертифікат.

4.1.8 Заборонено розміщувати їжу на верстаку, розігрівати чай у лабораторному посуді, проносити їжу в лабораторію.

4.1.9 Працівники, допущені до роботи з ртуттю, повинні пройти спеціальне навчання, інструктаж і перевірку знань з питань охорони праці.

4.1.10 У разі виходу з ладу лабораторного обладнання, механізмів, інвентарю, інструментів, засобів захисту, ЗІЗ, пожежної сигналізації, засобів пожежогасіння персонал лабораторії повинен негайно доповісти завідувачу лабораторією (керівнику робіт).

4.1.11 Не дозволено залишати опалювальне обладнання без нагляду.

4.1.12 Весь персонал лабораторії повинен вміти надавати першу допомогу потерпілим.

4.1.13 Якщо персонал лабораторії отримав травму або захворів, він повинен припинити роботу, повідомити безпосереднього керівника або керівника лабораторії та звернутися за допомогою до медпункту або найближчого медичного закладу.

4.1.14 Працівники лабораторії повинні негайно повідомляти про будь-яку ситуацію, що загрожує життю чи здоров'ю працівників лабораторії, а також про нещасний випадок у лабораторії своєму безпосередньому керівнику або завідувачу лабораторією.

4.1.15 Працівники лабораторії повинні стежити за придатністю для використання спецодягу, своєчасно проводити хімчистку, а також передбачати шафи для утримання одягу в чистоті та порядку.

4.1.16 Відповідальність за недотримання вимог цієї інструкції несе персонал лабораторії.

4.2 ВИМОГИ БЕЗПЕКИ ПЕРЕД ПОЧАТКОМ РОБОТИ

4.2.1 Працівники лабораторії повинні привести в порядок спецодяг і засоби захисту, необхідні для роботи.

4.2.2 Перевірити роботу вентиляції витяжних шаф (стулки шаф повинні бути щільно закриті).

4.2.3 Перед пуском в роботу того чи іншого лабораторного приладу ретельно перевірити його технічний стан і лише після усунення всіх дефектів, що будуть виявлені його вмикати.

4.2.4 Усі електронагрівальні прилади обов'язково установлювати на листовому азбесті товщиною 8-10 мм.

4.2.5 У приміщенні, де проводяться роботи із шкідливими, токсичними та пожежонебезпечними речовинами, необхідно за 15-20 хвилин до початку роботи включити припливно-витяжну вентиляцію і вимкнути її через 20-30 хвилин після закінчення роботи з ними.

4.2.6 Всі роботи з ЛЗР, горючими рідинами (ГР) і речовинами, здатними виділяти пожежонебезпечні пари і гази, необхідно проводити тільки в справних, витяжних шафах при працюючій місцевої витяжної вентиляції.

4.2.7 Під час роботи з ртуттю у відкритих ємностях вентиляцію витяжної шафи необхідно вмикати за 20-30 хв до початку роботи і вимикати через 30 хв після закінчення роботи.

4.3 ВИМОГИ БЕЗПЕКИ ПІД ЧАС РОБОТИ

4.3.1 Працюючи з реактивами у витяжній шафі, не засовувати голову в шафи та не працювати в лабораторії наодинці. Не перевантажувати проходи, виходи та точки доступу до засобів пожежогасіння.

4.3.2 Перед збовтуванням розчинів у колбах або пробірках їх необхідно закрити притертими пробками.

4.3.3 Заборонено нахилитися над ємностями, в яких нагріваються рідини, їдкі та отруйні хімічні речовини.

4.3.4 Оцінюючи хімічну речовину за запахом, ємність з речовиною слід тримати на відстані від людини, а руку слід злегка поворухнути, щоб безпосередньо піддати впливу пари або газу, що виділяються хімічною речовиною.

4.3.5 Перекачування корозійно-агресивних рідин із пляшок у ємності проводити у витяжній шафі з увімкненою вентиляцією за допомогою скляного сифону або насоса з іско- та корозійностійких матеріалів.

4.3.6 Під час перевезення пляшок, що містять кислоти, луги та інші корозійні або токсичні речовини, особливу увагу слід приділяти умовам зберігання тари та наявності або відсутності упаковки (стружка, азбестові прокладки тощо).

4.3.7 Для приготування розчину сірчаної кислоти тонкою цівкою невеликими порціями за постійного перемішування у воду вливають розчин сульфатної кислоти. Забороняється наливати воду в сірчану кислоту, щоб уникнути можливості закипання і виплескування рідини з ємності.

4.3.8 Заповнювати ємність розчином не більше ніж на 3/4 об'єму для кип'ятіння.

4.3.9 Отруйні речовини зберігати у вентильованих та закритих шафах (сейфах) відповідно до спеціальних інструкцій. Ключ від шафи повинен зберігатися у адміністратора лабораторії.

4.3.10 Отруйні речовини нагрівати в круглodonних колбах на азбестовій решітці.

4.3.11 Отруйні речовини розтирати в закритій ступці у витяжній шафі.

4.3.12 Якщо на підлогу (стіл) пролито хімічні речовини або розчинники, їх треба негайно знешкодити їх і видалити. Цю роботу проводити із застосуванням засобів індивідуального захисту (протигаз, респіратор, ПВХ або гумові рукавички). Якщо на підлогу (стіл) розлито небезпечні речовини, їх необхідно засипати тирсою або піском, ретельно зібрати та знешкодити. Після цього підлогу і стільниці необхідно обробити хлорним вапном і ретельно промити водою.

Для нейтралізації розливої азотної кислоти слід використовувати фільтруючі респіратори.

4.3.13 Фільтри та папір, які використовуються у роботі з токсичними речовинами, повинні бути поміщені в закритий контейнер для подальшої утилізації.

4.3.14 Заборонено зливати в каналізацію хімічні відходи та органічні розчинники. Необхідно зібрати відходи в спеціальний контейнер, нейтралізувати кислотою або лугом і утилізувати.

4.3.15 Легкозаймисті речовини з низькими температурами кипіння (до 80 °C) потрібно переганяти і нагрівати у круглодонних колбах або жаростійкому склі на водяній чи піщаній бані або на закритому електричному посуді.

4.3.16 На випадок відключення електроенергії в лабораторії необхідне аварійне освітлення або вибухозахищені лампи з лужними батареями.

4.3.17 Під час роботи з легкозаймистими матеріалами або у разі ввімкнених електроплит заборонено стояти спиною до плити або біля витяжної шафи.

4.3.18 Необхідно використовувати термостійкий або керамічний посуд, щоб поглинати тепло під час змішування або розчинення хімікатів.

4.3.19 Усі роботи з вибуховими речовинами слід проводити у витяжній шафі в металевій тарі або скляній пляшці, оточеній металевою запобіжною сіткою.

4.3.20 Деякі сольові розчини (наприклад, аміак срібла) можуть утворювати вибухонебезпечні суміші під час зберігання, тому після використання їх необхідно нейтралізувати та утилізувати у зазначеному порядку.

4.4 ВИМОГИ ОХОРОНИ ПРАЦІ ПРИ КОРИСТУВАННІ ХІМІЧНИМ ПОСУДОМ

4.4.1 Під час складання устаткування або його вузлів зі скляних хімічних ємностей слід застосовувати гнучкі з'єднання та прокладки.

4.4.2 Заборонено розігрівати відкритим вогнем хімічний посуд зі звичайного скла без використання азбестової сітки.

4.4.3 Нагріту скляну пляшку необхідно охолодити до температури навколишнього середовища, а потім закрити пробкою.

4.4.4 Під час перемішування розчину хімікату скляною паличкою на кінчик палички надягають гумову трубку (для водних розчинів) або армовану тефлоном трубку.

4.4.5 Заборонено зберігати розчини лугів і концентрованих кислот у тонкостінній скляній тарі.

4.4.6 Насадження гумової трубки на скляну трубку або стрижень і вставлення скляної трубки в отвір у гумовій трубці слід здійснювати, попередньо змочивши кінець трубки водою, гліцерином або вазеліном і наблизивши трубку ближче до отвору. Поставляється з рушником.

4.4.7 Розбивати скляні трубки можна тільки після розрізання напилком або спеціальним ножом для різання скла.

4.4.8 Перед знезараженням та чищенням хімічного посуду слід одягнути захисні окуляри, гумові рукавички та гумовий фартух.

4.4.9 Дезінфікувати посуд хімічними речовинами слід у витяжній шафі.

4.4.10 Посуд, що містить хімікати, слід ставити в раковину, не залишаючи на ньому залишків.

4.4.11 Заборонено мити хімічний посуд бензином, бензолом, ефіром, ізопропіловим спиртом.

4.4.12 Під час миття хімічного посуду з тонким горлом необхідно обережно знімати чашу, щоб уникнути розбризкування вмісту ємності.

4.4.13 Заборонено зберігання в лабораторіях легких у обробці рідин понад добової норми. Якщо добова потреба перевищує 0,5 л, його слід зберігати в залізних шафах.

4.5 ВИМОГИ БЕЗПЕКИ ПІСЛЯ ЗАВЕРШЕННЯ РОБОТИ

4.5.1 Після закінчення робіт електричні та інші установки, обладнання та крани повинні бути закриті.

4.5.2 Прибрати з приміщення горючі речовини та відходи від горючих матеріалів, використані рідини (чорнослив), сміття та промаслені ганчірки.

4.5.3 Перекрити загальний газовий і водопровідний крани, відключити електроживлення.

4.6 ВИМОГИ БЕЗПЕКИ В АВАРІЙНИХ СИТУАЦІЯХ

4.6.1 Під час роботи можуть виникнути такі аварійні ситуації:

- виникнення пожежі, вибуху ураження електричним струмом;
- потрапляння кислоти або лугу на шкіру або в очі;
- виділення або утворення їдких, отруйних, вогне- або вибухонебезпечних речовин.

4.6.2 У разі виникнення аварійної ситуації працівник лабораторії зобов'язаний припинити роботу, відключити електричні прилади, електрообладнання (апаратуру, стенди), якщо необхідно, провести огорожу небезпечного місця, і негайно повідомити про подію завідувачеві лабораторії чи його заступнику.

4.6.3 Працівники лабораторії, що знаходяться поблизу, за сигналом тривоги зобов'язані негайно з'явитися до місця події і взяти участь у наданні першої медичної допомоги потерпілому або усунення виниклої аварійної ситуації.

4.6.4 У разі виникнення пожежі або загоряння займистої речовини персонал лабораторії повинен вказати точне місце пожежі або загоряння горючої речовини, негайно повідомити про подію завідувача лабораторії або

його заступника та перекрити вентиляцію. Це приміщення знеструмити, перекрити загальну газову арматуру та розпочати першочергові заходи пожежогасіння, а ємності з горючими та вибухонебезпечними речовинами евакуувати з приміщення.

4.6.5 Для гасіння пожеж і займання займистих речовин в лабораторії слід користуватися:

- у разі займання займистих рідин, що змішуються з водою, - будь-якими вогнегасниками, струменем води, піском, азбестовим або брезентовим покривалом;
- у разі займання займистих рідин, що не змішуються з водою, - вуглекислотними або порошковими вогнегасниками, піском, азбестовими або брезентовими покривалами, починаючи з периферії; при займанні фосфору - мокрим піском, вогнегасниками, водою в рясному кількості (для жовтого фосфору - 2 - 3% розчином мідного купоросу);
- у разі займання електроустановок, що знаходяться під напругою, - тільки вуглекислотними або порошковими вогнегасниками. При користуванні вуглекислотним вогнегасником не братися рукою за розтруб вогнегасника;
- у разі займання металевого натрію або калію - тільки сухими піском, магnezією або хлористим кальцієм. Застосовувати для цих цілей воду, пінний вогнегасник і вуглекислоту заборонено.

4.6.6 У разі користування пінними (вуглекислотними, порошковими) вогнегасниками струмінь піни (порошку, вуглекислоти) не направляти на людей. У разі попадання піни на незахищені ділянки тіла слід стерти її хусткою або який-небудь тканиною і змити водним розчином соди.

4.6.7 У приміщеннях з внутрішніми пожежними кранами до гасіння пожежі слід залучати дві людини: один котить шланг від крана до димоходу, а другий відкриває кран, розгортаючи шланг за командою.

4.6.8 Під час гасіння пожежі лопатою для піску не дозволено піднімати лопату на рівень очей, щоб пісок не потрапив у полум'я.

4.6.9 Якщо на досліднику загорівся одяг, його необхідно загасити якомога швидше, але не голими руками. Будь-який запалений одяг слід негайно викинути, розірвати або загасити водою. На палаючий одяг дослідника можна накинути щільну тканину, ковдру або брезент, який необхідно зняти після гасіння полум'я.

4.6.10 Ця директива застосовується до всіх осіб, які мають право працювати в лабораторіях у рамках навчання, стажування або іншої дослідницької діяльності.

ВИСНОВКИ

1) Виконано науковий літературний пошук за темою роботи, наведено методи аналізу експресії генів, обґрунтовано застосування методу volcano plot та бази даних STRING для визначення найбільш значущих диференційно експресованих генів між другим та першим триместрами фізіологічної вагітності.

2) Побудовано volcano plot для диференційно експресованих генів між першим та другим триместрами фізіологічної вагітності.

3) Визначено 20 найбільш значущих диференційно експресованих генів між 1 та 2 триместрами вагітності: PAEP, FCGR2B, HSD11B1, HSD3B2, C7, MT1G, BMP5, DEFB1, HSD17B2, HIGD1B, SCGB2A1, EPCAM, TMEM100, TSPAN8, GRB14, HECW2-AS1, GPC6, UCA1, JAM2, ARHGAP42

4) Проаналізовано функції знайдених 20 найбільш значущих диференційно експресованих генів, та їх взаємозв'язок: визначено, що це гени, пов'язані з імунною відповіддю та захистом, регуляцією росту та розвитку, рухом та міграцією клітин, гіпоксією та пухлинами, структурою клітинної мембрани. Виявлено, що один з проаналізованих диференційно експресованих генів у плаценті людини між першим та другим триместрами вагітності – BMP5 – найбільше експресується саме в плаценті людини, та, що рівень експресії цього гену під час прееклампсії значно змінюється порівняно з нормальним фізіологічним станом. Було б доцільним звернути на нього увагу при подальших дослідженнях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Brew, Obed, Mark HF Sullivan, and Anthony Woodman. "Comparison of normal and pre-eclamptic placental gene expression: a systematic review with meta-analysis." *PloS one* 11.8 (2016): e0161504.
2. Yong, Hannah EJ, and Shiao-Yng Chan. "Current approaches and developments in transcript profiling of the human placenta." *Human Reproduction Update* 26.6 (2020): 799-840.
3. Jin M, Xu Q, Li J, Xu S, Tang C. Micro-RNAs in Human Placenta: Tiny Molecules, Immense Power. *Molecules*. 2022 Sep 13;27(18):5943. doi: 10.3390/molecules27185943. PMID: 36144676; PMCID: PMC9501247.
4. Cindrova-Davies T, Sferruzzi-Perri AN. Human placental development and function. *Semin Cell Dev Biol*. 2022 Nov;131:66-77. doi: 10.1016/j.semcdb.2022.03.039. Epub 2022 Apr 4. PMID: 35393235.
5. <https://into-sana.ua/enc/puhirnij-zanesok/>
6. Kee-Hak Lim, MD Maternal-Fetal Medicine Specialist, Department of Obstetrics and Gynecology, White Plains Hospital
7. Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy. *Am J Obstet Gynecol*. 2000;183(1):S1–22.
8. Duley L, Gulmezoglu AM, Henderson-Smart DJ, Chou D. Magnesium sulphate and other anticonvulsants for women with pre-eclampsia. *Cochrane Database Syst Rev*. 2010;11, CD000025. doi:10.1002/14651858.CD000025.pub2.
9. Duley L, Henderson-Smart DJ, Walker GJ, Chou D. Magnesium sulphate versus diazepam for eclampsia. *Cochrane Database Syst Rev*. 2010;12, CD000127. doi:10.1002/14651858.CD000127. pub2. 23.
10. Duley L, Henderson-Smart DJ, Meher S, King JF. Antiplatelet agents for preventing pre-eclampsia and its complications. *Cochrane Database Syst Rev*. 2007;2, CD004659. doi:10.1002/14651858.CD004659.pub2.

11. Sarosh Rana, Elizabeth Lemoine, Joey P. Granger and S. Ananth Karumanchi Preeclampsia Pathophysiology, Challenges, and Perspectives <https://doi.org/10.1161/CIRCRESAHA.118.313276>
12. Gene Expression [Электронный ресурс]. – 2014. – Режим доступа до ресурсу: <https://www.nature.com/scitable/topicpage/gene-expression-14121669/>.
13. Intro to gene expression (central dogma) [Электронный ресурс] // Khan Academy – Режим доступа до ресурсу: <https://www.khanacademy.org/science/ap-biology/gene-expression-and-regulation/translation/a/intro-to-gene-expression-central-dogma>.
14. J Pharm Bioallied Sci. 2012 Aug; 4(Suppl 2): S310–S312. doi: 10.4103/0975-7406.100283 PMID: 23066278
15. Hacia JG. Resequencing and mutational analysis using oligonucleotide microarrays. Nat Genet. 1999 Jan;21(1 Suppl):42-7. doi: 10.1038/4469. PMID: 9915500.
16. GeneChip Expression Analysis Technical Manual. Affymetrix, Inc. 2005
17. Shomaker TS, Ward K. Microarray technology in biomedical research. Hawaii Med J. 2006 Sep;65(9):253-6. PMID: 17153668.
18. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016 Jan 26;17:13. doi: 10.1186/s13059-016-0881-8. Erratum in: Genome Biol. 2016;17(1):181. PMID: 26813401; PMCID: PMC4728800.
19. <https://string-db.org/>
20. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009 Jan;10(1):57-63. doi: 10.1038/nrg2484. PMID: 19015660; PMCID: PMC2949280.
21. Lykhenko O. K., Frolova A. O., Obolenskaya M. Yu. Changes in the human placental transcriptome during the physiological course of pregnancy. Biopolym. Cell. 2021 Feb;37(1):73-82. doi:10.7124/bc.000A4D

22. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019 Jan 8;47(D1):D607-D613. doi: 10.1093/nar/gky1131. PMID: 30476243; PMCID: PMC6323986
23. Li W, Freudenberg J, Suh YJ, Yang Y. Using volcano plots and regularized-chi statistics in genetic association studies. *Comput Biol Chem.* 2014 Feb;48:77-83. doi: 10.1016/j.compbiolchem.2013.02.003. Epub 2013 Mar 22. PMID: 23602812.
24. Li W. Volcano plots in analyzing differential expressions with mRNA microarrays. *J Bioinform Comput Biol.* 2012 Dec;10(6):1231003. doi: 10.1142/S0219720012310038
25. Li, Wentian. “Application of Volcano Plots in Analyses of mRNA Differential Expressions with Microarrays.” (2011).
26. Li W, Freudenberg J, Suh YJ, Yang Y. Using volcano plots and regularized-chi statistics in genetic association studies. *Comput Biol Chem.* 2014 Feb; 48:77-83. doi: 10.1016/j.compbiolchem.2013.02.003
27. Rahmatallah Y, Emmert-Streib F, Glazko G. Gene set analysis approaches for RNA-seq data: performance evaluation and application guideline. *Brief Bioinform.* 2016 May;17(3):393-407. doi: 10.1093/bib/bbv069
28. Xiao Y, Hsiao TH, Suresh U, et al. A novel significance score for gene selection and ranking. *Bioinformatics.* 2014;30(6):801-807. doi:10.1093/bioinformatics/btr671
29. Jimenez-Luna C, Torres C, Ortiz R, Dieguez C, Martinez-Galan J, Melguizo C, Prados JC, Caba O. Proteomic biomarkers in body fluids associated with pancreatic cancer. *Oncotarget.* 2018 Mar 27;9(23):16573-16587. doi: 10.18632/oncotarget.24654. PMID: 29662668; PMCID: PMC5893263.
30. The R Project for Statistical Computing [Электронный ресурс] – Режим доступа до ресурсу: <https://www.r-project.org/>.

31. GitHub - hbctraining/Training-modules: Materials for short, half-day workshops. GitHub. URL: <https://github.com/hbctraining/Training-modules>
32. <https://www.proteinatlas.org/>
33. Digre, A, Lindskog, C. The Human Protein Atlas—Spatial localization of the human proteome in health and disease. *Protein Science*. 2021; 30: 218– 233. <https://doi.org/10.1002/pro.3987>
34. YiLin Meng, Chuang Li & Cai-Xia Liu (2021) Immune cell infiltration landscape and immune marker molecular typing in preeclampsia, *Bioengineered*, 12:1, 540-554, DOI: 10.1080/21655979.2021.1875707
35. Uuskula L, Mannik J, Rull K, et al. Mid-gestational gene expression profile in placenta and link to pregnancy complications. *PLoS One*. 2012;7(11):e49248.