

ВИВЧЕННЯ ДИФЕРЕНЦІЙНОЇ ЕКСПРЕСІЇ ГЕНІВ У ПЛАЦЕНТІ ЗА ФІЗІОЛОГІЧНОЇ ВАГІТНОСТІ

Левченко С.О.¹, Оболенська М.Ю.²

¹КПІ ім. Ігоря Сікорського, levchenko.stanislaw@kpi.ua

²Інститут молекулярної біології і генетичної інженерії НАН України

Abstract

This study focuses on the identification and functional interpretation of genes with significantly altered expression in the placenta between the first and second trimesters and third and second trimesters of pregnancy. Differential expression analysis was performed using microarray data and visualized via MA and Volcano plots. Top 20 DE-genes were selected based on log₂FC and adjusted P-value and functionally annotated. The findings contribute to understanding molecular changes in normal placenta and may have future diagnostic value.

Keywords: placenta, microarray, gene expression, log₂FC, p-value.

Вступ. Плацента є центральним регулятором росту плода і адаптації матері до вагітності. Порушення у функціонуванні плаценти пов'язують із виникненням ускладнень вагітності, прееклампсії або затримки розвитку плоду. Природа виникнення цих ускладнень, їх патогенез, діагностика, попередження і лікування досі є нез'ясованими. Вибір плаценти для такого аналізу обґрунтований її центральною роллю у підтримці гомеостазу матері та плоду, а також тим, що молекулярні зміни в плацентарній тканині зумовлюють розвиток ускладнення. На цей час дослідження транскриптому є найбільш розвиненою технологією широкомасштабних досліджень і забезпечене відповідним біоінформатичним інструментарієм [1].

Окрім формального виділення диференційно експресованих генів за статистичними порогами, актуальним є пошуковий (exploratory) аналіз транскриптомних даних із використанням MA-plot, Volcano-plot та smooth scatter plot. Такий підхід дає змогу не лише підтвердити ключові диференційно експресовані (DE) гени, а й виявити додаткові патерни зміни експресії, які можуть стати відправною точкою для подальших функціональних досліджень

Мета дослідження. Мета роботи – використати найпростіший, найшвидший і наочний метод для аналізу диференційної експресії генів у плаценті під час переходу від I до II та від II до III триместрів фізіологічної вагітності.

Матеріали та методи. Для аналізу експресії генів було використано диференційно експресовані гени між триместрами вагітності, надані Олександром Лихенко (Група системної біології, Інститут молекулярної біології і генетики НАН України). Ці дані, збережені у форматі csv, були завантажені для подальшого аналізу та візуалізації в середовищі R. Для обробки даних та побудови графіків використовували пакет tidyverse, зокрема модуль ggplot2 для створення інформативних візуалізацій.

Були визначені порогові значення для статистичної та біологічної значущості: скориговане р-значення (adj.P.Val) – менше 0.05 та абсолютне значення логарифму двократної зміни експресії (|log₂FC|) – більше 1.

Пошуковий аналіз даних провели через побудову MA plot (залежність $\log_2FC$ від AveExpr) – середніх значень двійкового логарифму нормалізованої експресії генів, та порівняли з Volcano plot (залежність $-\log_{10}(adj.P.Val)$ від $\log_2FC$) за допомогою функцій пакету ggplot2.

Для ідентифікації 20 найбільш значущих генів було розраховано ранжуючий показник π -value, за такою формулою [2]:

$$\pi\text{-value}_i = \log_2FC_i \cdot (-\log_{10}(Adj.P.Value_i)),$$

де $\log_2FC$ – це логарифм зміни експресії для конкретного гена між триместрами;

$Adj.P.Value_i$ – це скоригована статистична значущість для множинних порівнянь для конкретного гена.

Саме перелік цих генів став об'єктом для пошукового/розвідувального аналізу даних (exploratory data analysis), який включає описову статистику, перевірку розподілів і виявлення прихованих закономірностей та аномалій у наборі даних, що дає змогу сформулювати робочі гіпотези та оптимізувати подальші етапи статистичного моделювання.

Результати та обговорення. Порівняння графіків MA plot (рис. 1), що візуалізують залежність логарифма зміни експресії від середнього рівня експресії, показало, що середня експресія генів (вісь AveExpr) охоплює схожий діапазон, переважно від 1 до 12.5. Більшість генів у кожному порівнянні не є диференційно експресованими, і їх $\log_2FC$ лежить переважно в межах від -1 до 1, скупчуючись навколо горизонталі $\log_2FC=0$.

Однак, масштаби та амплітуда значущих змін експресії суттєво різняться: на MA плоті для порівняння I та II триместрів спостерігається значна кількість генів, які демонструють виражені зміни у бік пониження, наприклад, ген *PAEP* ($\log_2FC \approx -4.0$).

На противагу цьому, на MA плоті для порівняння II та III триместрів спостерігаються окремі гени з високим рівнем підвищення експресії, наприклад, *ALPP* та *CRH*, значення $\log_2FC$ яких сягає понад +5.0.

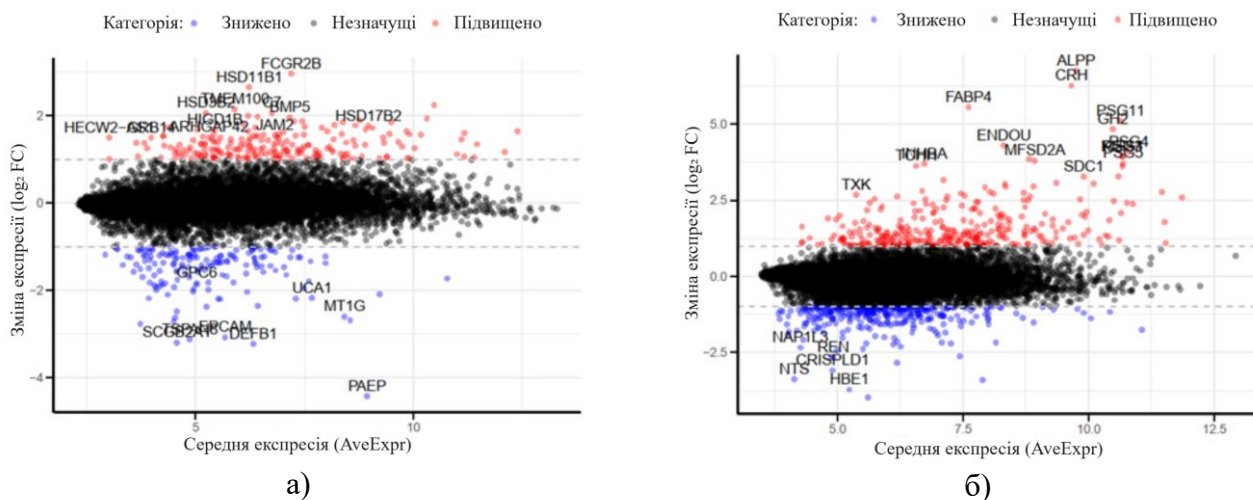


Рис. 1. MA plot диференційної експресії генів між а) першим і другим триместрами вагітності; б) другим і третім триместрами вагітності.

Volcano плоти (рис. 2) є ефективним інструментом для візуалізації результатів аналізу диференційної експресії, одночасно відображаючи зміну експресії та статистичну значущість змін для кожного гена. На Volcano плоті, що відповідає порівнянню I та II триместрів, розподіл значущих генів є майже симетричним по вісі $\log_2FC$. Проте виділяються гени з високою статистичною значущістю та сильною зниженою експресією, такі як *PAEP*.

На противагу цьому, Volcano плот для порівняння II та III триместрів виглядає інакше: помітна перевага точок у правому верхньому квадранті свідчить про схильність до збагачення генами з підвищеною експресією. На цьому плоті виділяються гени з високою статистичною значущістю та значно підвищеною експресією, зокрема *ALPP* та *CRH*.

Таким чином, порівняння Volcano плотів візуалізує відмінності в динаміці змін експресії: перехід від I до II триместру характеризується значними змінами у бік як підвищення, так і зниження, тоді як зміни між II та III триместрами включають гени з високим рівнем підвищеної експресії та статистичної значущості.

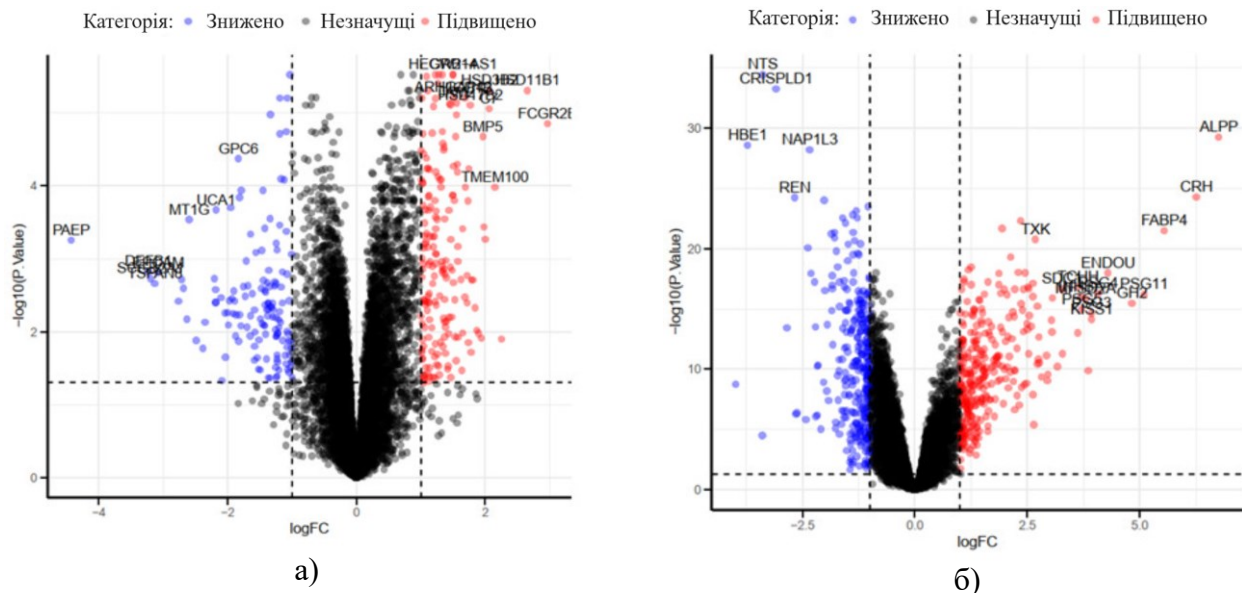


Рис. 2. Volcano plot диференційної експресії генів між а) першим і другим триместрами вагітності; б) другим і третім триместрами вагітності.

На основі розрахунку ранжуючого показника π -value було ідентифіковано топ 20 найбільш диференційно експресованих генів для кожного порівняння за результатами Volcano плоту.

Для переходу від I до II триместру до цього списку увійшли: *PAEP*, *FCGR2B*, *HSD11B1*, *HSD3B2*, *C7*, *MT1G*, *BMP5*, *DEFB1*, *HSD17B2*, *GRB14*, *HIGD1B*, *SCGB2A1*, *EPCAM*, *TMEM100*, *TSPAN8*, *ARHGAP42*, *HECW2-AS1*, *GPC6*, *JAM2*, *UCA1*.

У порівнянні II та III триместрів до топ 20 генів включено: *ALPP*, *CRH*, *FABP4*, *NTS*, *HBE1*, *CRISPLD1*, *PSG11*, *ENDOU*, *GH2*, *PSG4*, *NAP1L3*, *REN*, *TINHBA*, *MFSD2A*, *INHBA*, *PSG3*, *TXK*, *KISS1*, *PSG5*, *SDC1*. Ці гени

демонструють найбільш значущі та виражені зміни експресії на відповідних етапах вагітності і є ключовими кандидатами для подальшого дослідження [3].

Ідентифікація топ-20 найбільш диференційно експресованих генів за ранжуючим показником π -value розкрили функціональну специфіку кожного етапу. У порівнянні I і II триместру серед лідерів виявлено гени, залучені до регуляції гормонального фону (*CGA*, *GHI*), модуляції імунної відповіді (*FCGR2B*) та процесів ангиогенезу (*VEGFA*, *ANGPT2*), що відображає період інтенсивної перебудови плаценти, включаючи інвазію трофобласта та формування судинної мережі.

На противагу цьому, у порівнянні II $\rightarrow$ III триместр, до топ-генів увійшли *SLC2A1* (транспортування глюкози), *HSD11B2* (метаболізм стероїдних гормонів), *SOD2* (антиоксидантний захист), *MFAP5* (організація позаклітинного матриксу), а також гени, що демонструють надзвичайно високе підвищення експресії, такі як *ALPP* та *CRH*. Цей профіль вказує на посилення функцій забезпечення потреб плода, структурної підтримки та захисних механізмів на пізніх термінах вагітності.

Ідентифікація специфічних наборів диференційно експресованих генів та відповідних біологічних процесів для кожного переходу між триместрами розширює розуміння молекулярних механізмів нормального розвитку плаценти [4].

Висновки. Дане дослідження було присвячене аналізу диференційної експресії генів у плаценті протягом фізіологічної вагітності шляхом порівняння транскриптомних профілів між I і II, а також II і III триместрами. Візуалізація результатів за допомогою MA плотів та Volcano плотів чітко показала, що перехід від першого до другого триместру супроводжується значно більшою кількістю генів зі значущими змінами експресії та вищою амплітудою цих змін, ніж перехід від другого до третього триместру.

Ранжування за π -value й аналіз топ-20 генів підтверджують центральну роль *CGA*, *GHI*, *FCGR2B*, *VEGFA* та *ANGPT2* на ранній стадії та *SLC2A1*, *HSD11B2*, *SOD2* і *MFAP5* наприкінці вагітності. Функціональний GO-аналіз висвітлив специфічні біологічні процеси (від імуномодуляції й ангиогенезу до гормональної регуляції й метаболічної адаптації), що забезпечують нормальний розвиток плода. Отримані результати створюють надійну основу для подальшого вивчення потенційних біомаркерів ускладнень вагітності та розробки превентивних стратегій.

Список використаної літератури:

1. Lykhenko O.K., Frolova A.O., Obolenskaya M.Yu. Changes in the human placental transcriptome during the physiological course of pregnancy. *Biopolymers and Cell*. 2021; 37(1):73–82. doi:10.7124/bc.000a4d
2. Than N.G., Balogh A., Romero R. et al. Impaired placental immunoregulation and the pathogenesis of preeclampsia. *Nature Reviews Nephrology*. 2022; 18:466–484
3. Burton G.J., Fowden A.L., Thornburg K.L. Placental origins of chronic disease. *Physiol. Rev.* 2016; 96(4):1509–1565.
4. Xiao, Y. et al. A novel significance score for gene selection and ranking. *Bioinformatics*, 2014; 30(6), 801–807. doi:10.1093/bioinformatics/btr671
5. Ritchie M.E., Phipson B., Wu D. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. 2015; 43(7):e47.