

Анализ коэкспрессии генов плацентарного транскриптома как основа для поиска ключевых сигнальных путей и биомаркеров преэклампсии

Бабовская А.А., Трифонова Е.А., Зарубин А.А., Марков А.В., Степанов В.А.

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук
634050, г. Томск, ул. Набережная Ушайки, 10

Проблема профилактики и ранней диагностики преэклампсии (ПЭ) продолжает оставаться одной из ведущих в акушерстве, поскольку данное осложнение беременности несет большой риск материнской и младенческой смертности. Считается, что основная причина ПЭ – это нарушение этапов формирования плаценты, а регуляции экспрессии генов является значимым механизмом развития плацентарной патологии. Классический подход в транскриптомных исследованиях экспрессии основан на поиске дифференциально-экспрессирующихся генов при заболевании, однако такой подход рассматривает гены изолированно, не учитывая их возможные взаимодействия. Более перспективным подходом является анализ коэкспрессии, который описывает гены, вовлеченные в единые биологические пути патологического процесса, а также позволяет выделять в каждом из кластеров наиболее функционально значимый ген в сети – центральный (hub gene).

Ключевые слова: коэкспрессия, транскриптом, плацента, преэклампсия

Для цитирования: Бабовская А.А., Трифонова Е.А., Зарубин А.А., Марков А.В., Степанов В.А. Анализ коэкспрессии генов плацентарного транскриптома как основа для поиска ключевых сигнальных путей и биомаркеров преэклампсии. *Медицинская генетика* 2020; 19(11): 71-72.

DOI: 10.25557/2073-7998.2020.11.71-72

Автор для корреспонденции: Бабовская Анастасия Александровна; **e-mail:** anastasia.babovskaya@medgenetics.ru

Финансирование. Работа выполнена при финансовой поддержке РФФИ (проекты №18-29-13045, №18-44-700007).

Конфликт интересов. Авторы декларируют отсутствие конфликта интересов.

Поступила: 20.05.2020.

Co-expression network analysis of placental transcriptome as the basis for searching signal ways and preeclampsia biomarkers

Babovskaya A.A., Trifonova E.A., Zarubin A.A., Markov A.V., Stepanov V.A.

Research Institute of Medical Genetics, Tomsk National Research Medical Center, Russian Academy of Sciences
Naberezhnaya Ushaiki 10, Tomsk, 634050, Russia

The problem of prevention and early diagnosis of preeclampsia continues to be one of the leading in obstetrics. It's a major problem that contributes substantially to maternal and perinatal morbidity and mortality worldwide. Gene expression contributes significantly to the pathogenesis of placental diseases. Traditional methods of studying gene expression are based on the search of differentially expressed genes in a disease, but this approach considers genes in isolation. Coexpression analysis describes the genes involved in the unified biological pathways of the pathological process and also allows you to select in each of the clusters the most functionally significant gene in the network – the hub gene.

Keywords: coexpression, transcriptome, placenta, preeclampsia

For citation: Babovskaya A.A., Trifonova E.A., Zarubin A.A., Markov A.V., Stepanov V.A. Co-expression network analysis of placental transcriptome as the basis for searching signal ways and preeclampsia biomarkers. *Medical genetics*. 2020; 19(11): 71-72. (In Rus).

DOI: 10.25557/2073-7998.2020.11.71-72

Corresponding author: Babovskaya Anastasia; **e-mail:** anastasia.babovskaya@medgenetics.ru

Funding: The reported study was funded by RFBR according to the research projects №18-29-13045, №18-44-700007

Conflict of interest. The authors declare no conflict of interests.

Accepted: 20.05.2020.

Проблема профилактики и ранней диагностики преэклампсии (ПЭ) продолжает оставаться одной из ведущих в акушерстве, поскольку данное осложнение беременности несет большой риск материнской и младенческой смертности. Считается, что основная причина ПЭ – это нарушение этапов формирования

плаценты, а регуляции экспрессии генов является значимым механизмом развития плацентарной патологии. Классический подход в транскриптомных исследованиях экспрессии основан на поиске дифференциально-экспрессирующихся генов при заболевании, однако такой подход рассматривает гены изолированно, не учитывая

их возможные взаимодействия. Более перспективным подходом является анализ коэкспрессии, который описывает гены, вовлеченные в единые биологические пути патологического процесса, а также позволяет выделять в каждом из кластеров наиболее функционально значимый ген в сети – центральный (hub gene).

Материалы и методы

Был проведен полногеномный анализ экспрессии генов плацентарной ткани представителей двух этнических групп (русские ($n=24$) и якуты ($n=23$)) на микрочипах Illumina HumanHT-12 v3 Expression BeadChip. Полученные в нашем эксперименте данные по экспрессии были объединены с данными экспериментов, доступных в Gene Expression Omnibus (GEO) [1] (GSE25906, GSE30186, GSE35574, GSE44711, GSE60438, GSE6573, GSE73374, GSE94643). Всего в эксперимент вошли 53 образца от пациенток с ПЭ и 76 контрольных образцов. Взвешенный сетевой анализ коэкспрессии WGCNA осуществлялся в среде R, функциональную интерпретацию генов проводили в базе DAVID [2], построение сети геновых взаимодействий и оценку функциональной значимости генов внутри кластера осуществляли в STRING [3], центральными генами (hub genes) кластера считали гены, имеющие score $\geq 0,7$.

Результаты

С помощью WGCNA проанализировано 3986 генов и получен 31 кластер коэкспрессирующихся генов у больных ПЭ, и 34 кластера выделено в контрольной группе. Мы отобрали 86 генов, образующих 6 кластеров, ассоциированных с заболеванием, не показавших перекрытия с контрольной группой. При построении общей сети всех 86 генов только четыре гена не вошли в неё, что говорит о высокой степени связи между рассматриваемыми локусами. Внутрикластерный анализ белок-белковых взаимодействий выделил только три кластера из шести (C1, C4, C5), которые характеризуются высоким коэффициентом корреляции между генами.

В первый кластер C1 вошли 12 генов (*RAD21*, *YY1*, *ZFP30*, *GHR*, *LACCI*, *SEPT11*, *SHFM1*, *TMEM161A*, *TMEM258*, *UBQLN3*, *UBE2H*, *ZNF12*). Центральным геном кластера выступает *RAD21* (score = 0,7), участвующий в процессах репарации двойных разрывов, недостаточная экспрессия данного гена тормозит переход клетки из фазы S в G2[4]. Также центральным геном в данном кластере является транскрипционный фактор *YY1* (score = 0,7), который регулирует активность матриксных металлопротеиназ и, тем самым, может играть ключевую роль в миграции и инвазии трофобlasta.

Кластер (C4) включает 17 генов, вовлеченные в GO категорию, связанные с иммунным ответом как клеточ-

ного (*GSTP1*, *CXCL10*, *B2M*, *RGS1*, *CXCL9* ($p=0,005$)), так и гуморального типа (*IFI44L*, *IFI44*, *IFIH1*, *CXCL10*, *CXCL9* ($p=0,01$)). Отмечается участие белков *IFIH1*, *CXCL10* в RIG IR-подобном сигнальном пути; рецепторов *CXCL9*, *CXCL10* в каскаде активации Toll-подобных рецепторов; молекул ферментов *GSTP1*, *DPYD* в путях метаболизма ксенобиотиков, а также ферментов *LAP3*, *GSTP1* в метаболизме глутатиона ($FDR=0,033$). При построении сети белок-белковых взаимодействий, центральными в модуле выступают гены семейства хемокиновых рецепторов *CXCL9* (score=0,988), *CXCL10* (score=0,869), которые участвуют в хемоаттракции и пролиферации лимфоидных клеток, а также в процессе ангиогенеза. Известно, что нарушения в ангиогенезе являются одним из патогенетических компонентов ПЭ. Кроме того гены семейства интерферона *IFI44* (score=0,921), *IFI44L* (score=0,906) являются центральными в данном кластере. Дисбаланс в медиаторах иммунной системы приводит к срыву иммунотollerантности и также может выступать одним из факторов развития ПЭ.

Гены C5 кластера в основном относились по GO к группе белок-связывающих молекул ($p=0,003$). По данным Reactome продукты генов *CSGALNACT2*, *GPC4* участвуют в метаболизме хондроитинсульфата ($FDR=0,039$), который, как известно, стимулирует инвазию клеток трофобlasta. Главными функционально-активными генами кластера являются *GYP4*, *GYPB* (score=0,904), которые относятся к компонентам мембранных эритроцитов и ответственны за контроль инвазии малярийного плазмодия ($FDR=0,021$).

Взвешенный сетевой анализ коэкспрессии генов показывает в целом сходное распределение транскриптов по модулям, обнаруженным как в контрольной группе, так и при ПЭ. Однако были найдены модули коэкспрессии, характерные только для ПЭ, включающие 86 генов. Согласно функциональной аннотации данные локусы вовлечены в процессы регуляции жизненного цикла и метаболизма клетки, репарации ДНК и иммунной системы. С помощью анализа в сети белок-белковых взаимодействий мы выявили новые потенциальные биомаркеры ПЭ, такие как *RAD21*, *YY1*, *GYP4*, *GYPB*, *CXCL9*, *CXCL10*, *IFI44*, *IFI44L*, *AHSP*, связанные с процессами опухолеобразования, ангиогенеза, миграции и инвазии трофобlasta, функциональная значимость которых должна быть подтверждена экспериментально в последующих исследованиях.

Литература / References

1. Gene Expression Omnibus. URL: <https://www.ncbi.nlm.nih.gov/geo/>
2. DAVID Bioinformatics Resources 6.8. URL: <https://david.ncifcrf.gov/>
3. STRING. URL: <https://string-db.org/>
4. Wu Y. et al. Identification of novel biomarkers for preeclampsia on the basis of differential expression network analysis. Experimental and therapeutic medicine 2016; 12(1): 201–207.