Changes in the human placental transcriptome during the physiological course of pregnancy

O. K. Lykhenko, A. O. Frolova, M. Yu. Obolenskaya

Institute of Molecular Biology and Genetics, NAS of Ukraine
150, Akademika Zabolotnoho Str., Kyiv, Ukraine, 03143
Lykhenko.olexandr@gmail.com

Aim. To determine the changes of gene expression in human placenta during the physiological course of pregnancy. Methods the integrative analysis of publicly available data. Results. We have revealed that between the first and second trimesters of gestation the main changes relate to immune processes, morphogenesis and intercellular communication through the surface receptors of the cellular membrane. In the interval between the second and third trimesters, the main changes concern the regulation of the response to external stimuli, metabolic processes, morphogenesis of individual tissues, regulation of signaling pathways via transmembrane serine / threonine protein kinase receptors. Conclusion. The changes in gene expression in human placenta in the course of physiological gestation will serve as a reliable control with the changes during the pathological course of gestation.

Keywords: human placenta, transcriptome, integrative analysis

Introduction

The placenta is a temporary organ that develops ab ovo to a multifunctional organ and provides a close relationship between mother and fetus throughout pregnancy. The disruption in placenta’s functioning is the cause of several serious complications of pregnancy such as primarily preeclampsia, a common complication in the world (5–10 % of pregnancies) with the highest maternal and fetal mortality [1, 2]. The beginning of preeclampsia is connected with the adverse placentation and its clinical manifestations do not appear until the 20th week [3]. For a detailed study of the preeclampsia pathogenesis at the molecular level, it is necessary to know what changes in the molecular processes of the placenta occur during the physiological course of pregnancy. The work is dedicated to clarifying this issue.

Rapid development of systems biology and high-performance technologies has led to a large-scale study of the processes occurring in the placenta and its structures during the
physiological and pathological course of pregnancy [4, 5]. The vast majority of large-scale gene expression data have been so far obtained using the microarray technology for the samples of the bulk placental tissue and its components [6, 7]. These data are stored in the Gene Expression Omnibus (GEO) and Array Express databases and are available to scientists as a source of information and as a material for comparative studies. The available studies tend to use the samples of small size due to a high cost of the analysis, which affects the statistical reliability of the findings. There are two ways to increase the samples size: the meta-analysis, by which the results of each analysis are combined, and the integrative analysis, which consists in combining the original data into a new larger group with normalization of the previous data and subsequent simultaneous analysis of all the data in a newly created group. The first approach provides a list of differentially expressed genes [8] whereas the second provides quantitative characteristics of the differentially expressed genes, but it requires the use of a more complex mathematical apparatus [9]. The aim of the study was to determine how the gene expression changes in the bulk tissue of the human placenta in the intervals from the first to the second and from the second to the third trimesters of pregnancy by the integrative analysis of the publicly available data on gene expression.

**Material and Methods**

1. **General description of the data**

From the open databases GEO [10] and ArrayExpress [11], we downloaded the data from 8 studies of gene expression in the human placenta from three trimesters of pregnancy using the R package ArrayExpress (Table 1). From the specialized database for microarray experiments’ metadata developed by us, we downloaded the relevant standardized metadata [12].

| GSE accession number | Samples count | Gene count |
|----------------------|---------------|------------|
|                      | Total Trimesters |            |
| GSE122214            | 4 4 4           | 20261      |
| GSE22490             | 6 5 1           | 20261      |
| GSE35574             | 26 26           | 24547      |
| GSE37901             | 4 4             | 20342      |
| GSE6573              | 1 1             | 20326      |
| GSE73374             | 12              | 29509      |
| GSE73685             | 12              | 20624      |
| GSE9984              | 12 4 4 4        | 20342      |
| Total                | 77 13 9 55      |            |

2. **Pipeline: data processing and analysis**

We used the affy [13] R package to access and preprocess the expression data from Affymetrix raw CEL files that store the results of the intensity calculations on the pixel values. A single representative intensity value is stored per cell (feature) of the image. The Affy package uses a robust multi-array average (RMA) algorithm [14] based on the quantile normalization algorithm [15]. RMA maps probes to genes, then the raw intensity values are background corrected, log2 transformed and quantile normalized. For Illumina arrays we used a similar lumi R package [16]. We cross-normalized the data between datasets eliminating non-biological variability (batch-effect) using the empirical Bayes method implemented as a ComBat function from the sva R package [17].
We identified differentially expressed genes using generalized linear models implemented in the limma R package [18]. A linear model was fit to each gene with moderated t-statistics computed. P-value was adjusted to account for multiple gene comparisons with Benjamini & Hochberg method (FDR) [19]. We took genes with adjusted p-value < 0.05 and |logfc|>1 for the analysis. Differentially expressed genes were clustered using a type of greedy algorithm (fastgreedy) [20] based on String data [21]. In these clusters the functionally enriched groups of differentially expressed genes were found using Gene Ontology [22]. To increase the Gene Ontology results readability, categories of biological processes were grouped based on semantic proximity using REVIGO web tool (revigo.irb.hr) [23].

**Results and Discussion**

The clustering of differentially expressed genes revealed the presence of 4 clusters of functionally enriched genes from a total of eight clusters when comparing the data from the first and second trimesters, and 7 clusters of functionally enriched genes from a total of fourteen clusters when comparing the data from the second and third trimesters (Tables 2 and 3). Both tables show the clusters of genes with enrichment coverage value in the range of 0–1.0. This value shows the proportion of genes in a cluster, which are included in at least one enriched biological process according to Gene Ontology.

We identified 253 differentially expressed genes between the first and second trimester placenta (comparison 1_2); 152 up-regulated and 101 down-regulated in the second trimester. Table 2 contains a detailed breakdown by clusters. Comparison between the second and third trimester pregnancies (comparison 2_3) yielded 489 differentially expressed genes between; 221 genes are up-regulated and 268 are down-regulated in the third trimester. There are almost twice as many differentially expressed genes in 2_3 vs 1_2 most likely due to a larger gestational age gap in 2_3. Namely, the gestational age values in the datasets we used are in the range of 7–11 weeks for the first trimester, 13–18 weeks for the second and 38–42 for the third trimester (term). Majority of genes are up-regulated in 1_2 but down-regulated in 2_3. Furthermore, from 54 genes that are differentially expressed in both 1_2 and 2_3, 21 genes

| Clusters                       | Genes count | Log Fc Min | Log Fc Max | Enrichment coverage |
|--------------------------------|-------------|------------|------------|---------------------|
| 1. Immune system processes     | 114         | -3,23      | 2,97       | 0,87                |
| 2. Organ and blood vessel development | 22         | -3,2       | 2,15       | 0,88                |
| 3. Cellular response to zinc ion | 10          | -4,43      | 1,55       | 1                   |
| 4. Cell surface receptors signalling pathways | 7          | -2,68      | 2,25       | 1                   |
| Total                          | 253         | 152        | 101        |                     |

Note: Up — the number of genes, the expression of which is upregulated between the first and second trimesters; Down — the number of genes, the expression of which is downregulated between the first and second trimesters; Min — minimal value; Max - maximal value; Log FC — log fold change.
are up-regulated in 1_2 and down-regulated in 2_3 (UpDown). Respectively, 18 genes are UpUp, 12 genes are DownDown and only 3 are DownUp. These numbers suggest the existence of biological processes that peak in mid-gestation which requires additional research.

Shown below are the GO biological processes, grouped based on semantic proximity by Revigo. The cluster names in GO terms are given by the names of the biological processes with the lowest p-values.

1. *Comparison of gene expression in human placenta from the first and second trimesters of gestation*

When comparing gene expression between the first and second trimesters of pregnancy, the largest number of differentially expressed genes belongs to the first two clusters.

**Cluster 1. Immune system processes**

The placenta functions as an immunologically privileged barrier between mother and fetus. Their mutual immune tolerance is ensured by three main mechanisms: the placenta separates the tissues of the mother and the fetus; syncytiotrophoblast does not express leukocyte antigens of I and II classes, and extravillous trophoblast expresses typical HLA-C and atypical HLA-E and HLA-G leukocyte antigens and therefore avoids the action of maternal cytotoxic T-, NK-cells [24].

The proteins coded by differentially expressed genes between the first and the second trimesters are involved in humoral immune response; response to other organism; regulated exocytosis; inflammatory response; response to stress; chemokine-mediated signaling pathway; positive regulation of ERK1 and ERK2 cascade[s]; killing of cells of other organism; modification of morphology or physiology of other organism; multi-organism process; locomotion

Therefore, the immune system processes have more versatile functions for the placental development than supporting immune tolerance.

**Cluster 2. Organ and blood vessel development**

At the end of the 10–12 weeks of pregnancy, the maternal blood enters the intervillous space, and the histiotrophic nutrition of the placenta changes to hemotrophic [25]. The following processes are ongoing at this time: vasculogenesis and angiogenesis as well as the organ development, which includes morphogenesis or the formation of a three-dimensional spatial structure, and tissue growth occurring mainly through the cell division. The categories of this cluster represent the mentioned functions: regulation of angiogenesis; blood circulation and chemotaxis; regulation of the cAMP-dependent protein kinase activity and regulation of growth; regulation of the anatomical structure size; G-protein coupled receptor signaling pathway; multicellular organismal process; cell communication; response to external stimulus; movement of cell or subcellular component.

**Cluster 3. Cellular response to zinc ion**

Zinc is the most abundant microelement in the body after iron. In addition to its role in the catalytic activity of enzymes, zinc is involved in the regulation of innate and adaptive immunity and guards the immune response [26]. The peculiarity of zinc ions is the lack of redox potential and the ability to form coordination bonds in the structure of proteins. Zink is involved in the enzymes, in the structural
**Fig. 1.** Interaction graph for differentially expressed genes in between first and second trimesters of pregnancy. Clusters are named based on the most representative enriched biological processes according to Gene Ontology. Clusters 5–8 are unnamed because no enriched processes are found for them.

**Fig. 2.** Interaction graph for differentially expressed genes in between second trimester and term. Clusters are named based on the most representative enriched biological processes according to Gene Ontology.
proteins for which redox processes would lead to their destruction [27]. Among the differentially expressed genes of this cluster, the following processes are distinguished: the cellular response to zinc ion; the response to chemicals; negative regulation of growth; plasma lipoprotein particle remodeling; cholesterol homeostasis; reactive oxygen species metabolic process; positive regulation of cytokine secretion.

**Cluster 4.** Cell surface receptors signalling pathways

Signaling pathways from the receptors on cellular membranes are essential for the processes of morphogenesis. Every cell type has different number of adhesion molecules of different types on the cell membrane. The interaction between the adhesion molecules of different cells determines the mutual cellular localization in the tissue [28]. The genes of this cluster are involved in several signalling pathways: interleukin-35-mediated signaling pathway; response to peptide hormone; JAK-STAT cascade; positive regulation of cytokine production; peptidyl-tyrosine phosphorylation; positive regulation of phosphatidylinositol 3-kinase signaling; cytokine production and positive regulation of developmental growth.

Full list of differentially expressed genes and enriched biological processes between the 1 and 2 trimesters can be found in Supplement 1 (.xlsx file, pages 1, 3).

2. Comparison of gene expression in the human placenta between the second and third trimesters of pregnancy

The time interval from the second to the third trimester is characterized by intensive placental and fetal growth. The duration of gestation, the fetal length and weight at birth linearly depend on the thickness and area of placenta. These parameters are recommended to use to clarify the duration of pregnancy [29]. As can be seen from the list of categories below, the genes are involved in the general processes of development and in morphogenesis of specific structures. Unlike the differentially expressed genes between the first and second trimesters, the differentially expressed genes

| Cluster names                                      | Number of genes | LogFc Min | LogFc Max | Enrichment coverage |
|---------------------------------------------------|-----------------|-----------|-----------|---------------------|
| 1. Regulation of response to external stimulus    | Total 145       | 56        | 89        | -3.94               | 6.14                | 0.95                |
| 2. Metabolic processes                            | Total 119       | 43        | 76        | -1.91               | 2.97                | 0.93                |
| 3. O-glycan processing                            | Total 36        | 24        | 12        | -2.13               | 6.65                | 0.14                |
| 4. Tissue morphogenesis                           | Total 168       | 84        | 84        | -2.68               | 3.58                | 0.91                |
| 5. Sodium ion transmembrane transport             | Total 7         | 3         | 4         | -1.83               | 1.48                | 0.57                |
| 6. Chloride transmembrane transport               | Total 8         | 5         | 3         | -2.01               | 1.44                | 1                   |
| 7. Pregnancy specific beta-1-glycoproteins        | Total 6         | 6         | 0         | 3.16                | 4.97                | 1                   |
| Total                                             | 489             | 221       | 268       |                     |                     |                     |

Note: Up — the number of genes, the expression of which is upregulated between the second and third trimesters; Down — the number of genes, the expression of which is downregulated between the second and third trimesters. For other notations — see Table 2.
between the second and third trimesters contain more genes, the expression of which is downregulated in the term placenta (38–42 weeks of pregnancy).

**Cluster 1. Regulation of response to external stimulus**

This cluster contains the following categories: humoral immune response; platelet degranulation; inflammatory response; response to cytokine; plasma lipoprotein particle remodeling; regulation of angiogenesis; cellular response to chemokine; modification of morphology or physiology of other organism. At this stage, the cytokines, which include the chemokines, interferon, lymphokines and tumor necrosis factor, play the special role. They are involved in the cellular and humoral immunity. The signaling via chemokines is important for the coordinated cellular migration in space and time [30].

**Cluster 2. Metabolic processes**

The fully formed placenta is an active metabolizing organ. The following categories refer to this cluster: steroid metabolic process; organic substance metabolic process; cellular hormone metabolic process; cellular response to xenobiotic stimulus; nitric oxide biosynthetic process; regulation of neurotransmitter level.

**Cluster 3. O-glycan processing.** This small cluster is enriched with only one category – O-glycan processing. Glycans are the sequences of carbohydrates that are added to proteins and lipids to modulate their structure and function. In placenta, glycans modify the proteins required for trophoblast function, and alterations have been associated with pathological conditions [31].

This cluster also contains the genes involved in interferon gamma signalling pathway (HLA-G, TRIM62, IFI27) and the genes involved in organophosphate biosynthetic processes (AGPAT4, IP6K3, ELOVL7), the relations between which require further investigation.

**Cluster 4. Tissue morphogenesis**

According to Revigo and GO, this cluster contains the categories, reflecting the participation of its genes in: anatomical structure morphogenesis; heart development; tissue morphogenesis; negative regulation of cell differentiation; reproductive structure development; regulation of epithelial cell proliferation; odontogenesis; limb development. Epidemiological, clinical and experimental data confirmed the existing connection between placenta and the heart development [32]. The mechanistic basis for this connection is not clear – whether this connection is the result of the general role of placenta providing the nutrient substances to the fetus and removing the waste products, the adequate blood supply, the active placental metabolic activity, or there is a specific interaction between placenta and the fetal morphogenesis. The regulation of transmembrane receptor protein serine/threonine kinase signaling pathway is also enriched in this cluster. Transmembrane serine/threonine protein kinase receptors making up about a quarter of all protein kinases possess the catalytic and receptor functions. The binding of the extracellular ligand to these receptors activates the enzyme at the intracellular side of the cell membrane, which leads to phosphorylation of the corresponding proteins and signal transmission to the nucleus. The proteins of this group are involved in the MAPK signaling pathway, in the
interaction of cytokine receptors with cytokines, the TGFβ signaling pathway, and the formation of adhesive intercellular contacts in epithelium and endothelium [33]. The categories of this cluster expand the list in the previous cluster by including the categories connected with the development of various organs and systems, cellular response to growth factors and hormones, regulation of non-canonical and canonical Wnt signaling pathway, regulation of pathway-restricted SMAD protein phosphorylation, mesenchymal cell differentiation and extracellular matrix organization.

Clusters 5 and 6. Sodium transport and transmembrane transport of chloride

The genes of the 5th cluster refer to the nucleoside salvage synthesis and the transmembrane transport of sodium ions. The salvage nucleoside synthesis in contrast to de novo synthesis, is characteristic to the tissues with the low proliferative activity typical for the term placenta [34]. Transmembrane chloride transport and the general category transport are included in the 6th cluster.

Cluster 7. Pregnancy specific beta-1-glycoproteins

The pregnancy-specific glycoproteins (PSG) are complexes consisting of carbohydrate and protein, which are prevalent in the maternal bloodstream at the later stages of pregnancy. There are 10 protein-coding PSG genes (PSG1 - PSG9, PSG11) and 9 of them (no PSG5) are represented in differentially expressed genes. These proteins have immunoregulatory, proangiogenic and antithrombotic functions. They interact [with] and activate anti-inflammatory cytokines, growth factors TGFβ1 and TGF-β2, making them one of the few known biological activators of these important cytokines. TGF-β regulates many biological processes necessary for successful pregnancy [35].

Conclusion

Between the first and second trimesters of gestation the main changes relate to the immune processes, morphogenesis and intercellular communication through the surface receptors of the cellular membrane. In the interval between the second and third trimesters, the main changes concern the regulation of the response to external stimuli, metabolic processes, morphogenesis of individual tissues and systems and regulation of signaling pathways via the transmembrane serine/threonine protein kinase receptors. The described changes in gene expression will serve as a reliable control for comparison with any pathological changes that occur in the human placenta.

Full list of differentially expressed genes between the 2 and 3 trimesters can be found in Supplement 1 (.xlsx file, pages 2, 4).

Funding

The study was conducted within the budget project #2.2.4.18 and supported by the National Academy of Sciences of Ukraine. The authors declare that they have no conflicts of interest.

Additional files

Supplement1.xlsx - Differentially expressed genes and enriched biological processes in 1 vs 2 and 2 vs 3 trimesters, human placenta.

REFERENCES

1. Vousden N, Lawley E, Seed PT et al. Incidence of eclampsia and related complications across 10 low- and middle-resource geographical regions: Secon-
Changes in the human placental transcriptome during the physiological course of pregnancy

dary analysis of a cluster randomised controlled trial. PLoS Med. 2019;16(3):e1002775.
2. Ivanov II, Cheripco MV, Kosolapova NV, Prochas EN. Preeclampsia: speciality of pathogenesis and therapeutic tactics. Tavri Med Biol Vestn. 2012; 15(2), 2(58): 273–86.
3. Rana S, Lemoine E, Granger JP, Karumanchi SA. Preeclampsia: Pathophysiology, Challenges, and Perspectives. Circ Res. 2019;124(7):1094–1112.
4. Jóźwik M, Lipka A. Recent progress in human placental transcriptomics. Dev Period Med. 2019;23(2): 104–108.
5. Yong HEJ, Chan SY. Current approaches and developments in transcript profiling of the human placenta. Hum Reprod Update. 2020;26(6):799–840.
6. Knöfler M, Haider S, Saleh L, Pollheimer J, Gamage TKJB, James J. Human placenta and trophoblast development: key molecular mechanisms and model systems. Cell Mol Life Sci. 2019;76(18):3479–3496.
7. Szilagyi A, Gelencser Z, Romero R et al. Placenta-specific genes, their regulation during villous trophoblast differentiation and dysregulation in preterm preeclampsia. Int J Mol Sci. 2020; 21(2):628.
8. Brew O, Sullivan MH, Woodman A. Comparison of Normal and Pre-Eclamptic Placental Gene Expression: A Systematic Review with Meta-Analysis. PLoS One. 2016;11(8):e0161504.
9. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra JJ, Wishart D. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. Metabolites. 2019;9(4):76.
10. Clough E, Barrett T. The Gene Expression Omnibus Database. Methods Mol Biol. 2016;1418:93–110.
11. Brazma A, Parkinson H, Sarkans U et al. ArrayExpress—a public repository for microarray gene expression data at the EBI. Nucleic Acids Res. 2003;31(1):68–71.
12. Lykhenko OK, Frolova AO, Obolenskaya MY. Creation of gene expression database on preeclampsia-affected human placenta. Biopolym Cell. 2017; 33(6): 442–52.
13. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):307–15.
14. Irizarry RA, Bolstad BM, Collin F et al. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 2003;31(4):e15.
15. Zhao Y, Wong L, Goh WWB. How to do quantile normalization correctly for gene expression data analyses. Sci Rep. 2020;10(1):15534.
16. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008; 24(13):1547–8.
17. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012; 28(6):882–3.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
19. Smyth GK. limma: linear models for microarray data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (eds) Bioinformatics and computational biology solutions using R and bioconductor. Statistics for Biology and Health. Springer, New York, NY. 2005
20. Clauset A, Newman ME, Moore C. Finding community structure in very large networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2004;70(6 Pt 2):066111.
21. 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;47(1):D607–D613.
22. Mi H, Muruganujan A, Eberth D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1): D419–D426.
23. Šupek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.
24. Tersigni C, Meli F, Neri C, Iacoangeli A, Franco R, Lanzone A, Scambia G, Di Simone N. Role of Human Leukocyte Antigens at the Feto-Maternal Interface in Normal and Pathological Pregnancy: An Update. Int J Mol Sci. 2020;21(13):4756.
Зміни транскриптому в плаценті людини впродовж фізіологічного перебігу вагітності
О. К. Лихенко, А. А. Фролова, М. Ю. Оболенська

Мета. Визначити як змінюється експресія генів у цілісній тканині плаценти людини впродовж фізіологічного перебігу вагітності. Методи. Інтегративний аналіз наявних у відкритому доступі даних. Результати. Показано, що в інтервалі між першим і другим триместром основні зміни стосуються імунних процесів, морфогенезу і міжклітинного спілкування через поверхневі рецептори клітинної мембрани. В інтервалі між другим і третьим триместром основні зміни стосуються регуляції відповіді на зовнішній стимул, метаболічних процесів, морфогенезу окремих тканин, регуляції сигнальних шляхів через трансмембранні серин/треонін протеїнкиназні рецептори. Висновки. Зміни в експресії генів плаценти впродовж фізіологічного перебігу вагітності слугуватимуть надійним контролем для порівняння зі змінами впродовж патологічного перебігу вагітності.

Ключові слова: плацента людини, транскриптограма, інтегративний аналіз

Изменения в плацентарном транскриптотоме человека в течение физиологического течения беременности
О. К. Лихенко, А. А. Фролова, М. Ю. Оболенская

Цель. Определить как меняется экспрессия генов в плаценте человека в течение физиологического беременности. Методы. Интегративный анализ имеющихся в открытом доступе данных. Результаты. Показано, что в интервале между первым и вторым триместром основные изменения касаются иммунных процессов, морфогенеза и межклеточного общения через поверхностные рецепторы клеточной мембраны. В интервале между вторым и третьим триместром основные изменения касаются регуляции ответа на внешний стимул, метаболических процессов, морфогенеза отдельных тканей, регуляции сигнальных путей через трансмембранные серин/треонин протеинкиназные рецепторы. Выводы. Изменения в экспрессии генов плаценты на протяжении физиологического течения беременности будут служить надежным контролем для сравнения с изменениями при патологическом течении беременности.

Ключевые слова: плацента человека, транскриптограма, интегративный анализ