CLASSIFICATION AND FUNCTIONAL ANALYSIS OF GENE EXPRESSION DYNAMICS IN THE HUMAN PLACENTA DURING UNCOMPLICATED PREGNANCY

Kukuruza Y.O.[1], Obolenska M.Yu.[2]
[1]Igor Sikorsky Kyiv Polytechnic Institute, kukuruza.yevhenii@lll.kpi.ua
[2]Institute of Molecular Biology and Genetics of NASU

Abstract
This article analyzes placental gene expression dynamics during pregnancy. 9 expression dynamics clusters were grouped into 4 "mirrored" groups. Functional analysis revealed that genes in these groups partake in immune, metabolic, structural, and developmental processes, offering insight into coordinated function of genes in placenta.
Keywords: *placenta, gene, expression, dynamics, clustering.*

Introduction. Placenta is a temporary organ that forms as a part of embryo-placenta-mother system during pregnancy, mainly functioning as a barrier and facilitator of nutrient, oxygen, CO_2, and waste product exchange [1]. Gene expression dynamics, which control the proper development of the human placenta and thus severely impact the health of both the baby and the mother throughout and after pregnancy, are poorly studied. Most studies focus on comparison of two time points or term placenta [2]. Proper understanding of gene expression changes in placenta in healthy individuals is crucial for studying points of failure that result in pathology.

The aim of this study was to identify variants of gene expression dynamics by clustering genes with similar continuous changes in expression, analyze the interactions between these variants and investigate what biological processes are impacted by their simultaneous action.

Materials and methods. Placental gene expression data in healthy individuals, extracted from Gene Expression Omnibus database [3] from series GSE9984 [4], GSE22490 [5], GSE122214 [6], GSE6573 [7], GSE73374 [8], GSE73685 [9], GSE35574 [10], and GSE37901 [11] was normalized and combined into time series with 4 time points, from which differentially expressed genes were selected and clustered using the DPGP method [12], which jointly models cluster number via a Dirichlet process and temporal dependencies with Gaussian processes. Based on the similarity of gene expression dynamics curves (qualitatively similar and "mirrored" along the vertical axis), the clusters were combined into Groups. List of genes that belonged to each Group were then used to perform over-representation analysis (ORA) using Gene Ontology terms of "Biological process" aspect [13].

Cytoscape environment [14] and *stringApp* [15] were used to build protein-protein interaction (PPI) networks of gene expression products for each Group. Nodes, which represent individual genes, were colored according to their gene expression dynamics. Functional clusters obtained from *treeplot* graphs, built in *R* environment using *enrichplot* package [16], were used to color inner portion of nodes based on which functional cluster they belong to.

Results and discussion. According to the described algorithm, a list of 369 differentially expressed genes was identified. Based on this list, 9 clusters were

obtained, each containing genes with similar expression dynamics across gestation. It was found that each cluster had a “mirror pair” (Fig. 1.) – a different cluster, dynamics of gene expression in which were essentially reversed compared to the former. The exception was Cluster 2, which had two “mirror” clusters (6 and 8). 4 groups of clusters were identified (Group I-IV).

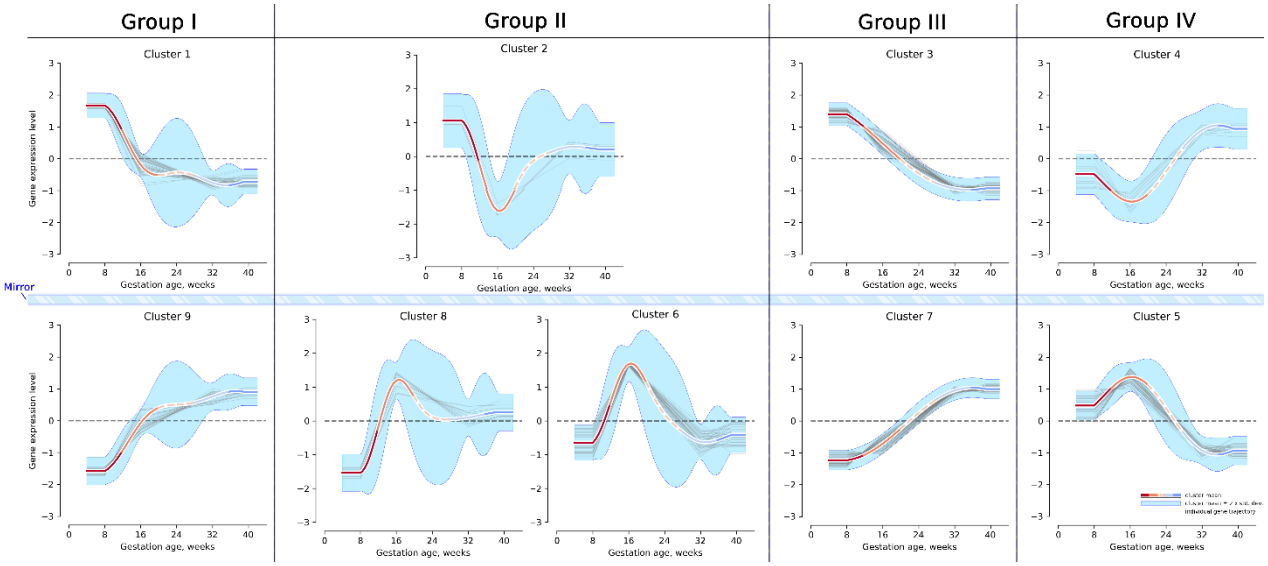

Fig. 1. Four groups of clusters were identified (Group I-IV), which had clusters with gene expression dynamics “mirrored” to each other. An imaginary mirror is shown in the middle of the figure for better understanding of the relationship between the graphs.

The genes which belong to clusters of one Group were combined to form 4 gene lists respectively and analyzed together to investigate the effects of concurrent upregulation and downregulation of genes with “mirrored” expression dynamics in human placenta during pregnancy.

Enriched pathways in Group I, according to the results of ORA, refer to such processes as immune response, cellular defense against chemical agents, such as metal ions, maintenance of physiological stability and barrier functions, tissue organization, tissue and organ development, hormone metabolism and endocrine regulation. Group I contained Clusters 1 and 9, which represent genes whose expression levels severely changes from first to second trimester and remains stable thereafter.

Enriched pathways in Group II referred to: cytokine, chemokine, and lipid responses, innate immune recognition of microbial components, immune defense mechanisms against pathogens, indicating a role in cell signaling and inflammatory stimuli recognition; chemotaxis and leukocyte migration and broader processes of coordination and regulation of adaptive immunity. Group II contained Clusters 2, 6 and 8 – genes, expression of which sharply changes from first to second trimester, drastically rebounding afterwards – rising or falling respectively by preterm.

Group III was enriched in biological pathways that referred to hormone regulation, lipoprotein, triglyceride and steroid metabolism, lipid localization and cholesterol transport, vesicle-mediated transport and phagocytosis regulation, as well as general reproductive and multicellular organismal processes. Group III contained

Clusters 3 and 7, in which the expression of genes either slowly rises or decreases throughout pregnancy.

Group IV was enriched in biological pathways that referred to development and morphogenesis of the heart and its valves, mesenchymal transformations and tissue development events, regulatory mechanisms acting at the systemic and organismal levels, as well as intracellular signaling pathways, cellular mobility and proliferation, including immune signaling and movement. Group IV contained Clusters 4 and 5, which represented genes whose expression only moderately changed from first trimester to second, but drastically increased or decreased by preterm.

PPI networks of gene expression products for each Group shows a high count of interactions between proteins coded by genes, whose expression changes in opposite ways for Groups I-III (Fig. 2. a, b, c). In Groups II and IV (Fig. 2. b, d) there is a high ratio of genes with one type of dynamics and their "mirrored" counterparts. Only in Group IV it was found that genes with one type of expression dynamics (red nodes) were not central to the network and were found mostly on the outside. Separate PPI networks (figures not shown) which combined this dynamics classification with functional clusters obtained from *treeplot* showed that genes, expression of which is moving in opposite directions, were partaking in the enrichment of the same and related GO terms.

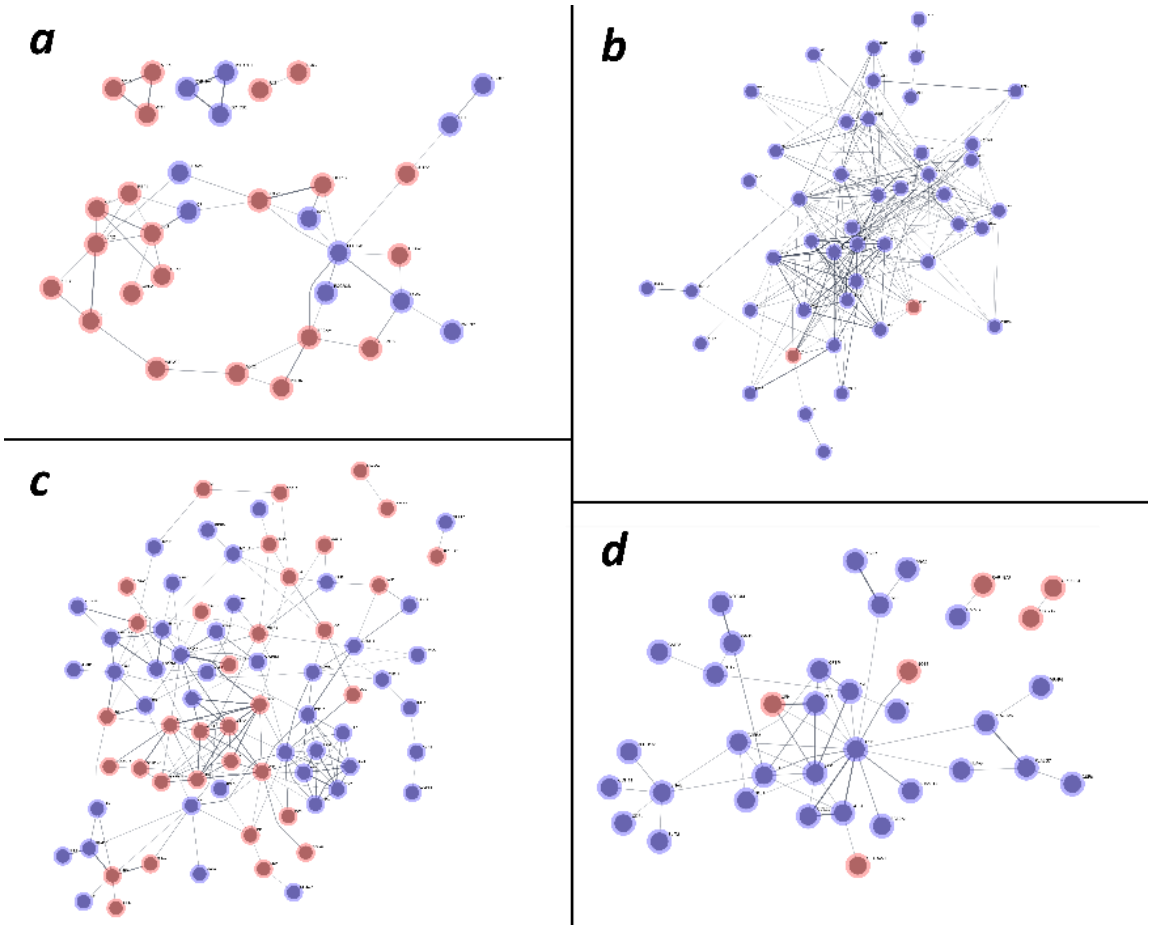

Fig. 2. PPI networks for products of gene expression for Groups I (*a*), II (*b*), III (*c*) and IV (*d*). Red nodes represent genes from Clusters 1, 2, 3, and 4, while blue nodes represent genes from their "mirror" counterparts – Clusters 9, 8 and 6, 7, and 5.

Conclusions. The defined variants of gene expression dynamics, which are represented by 9 clusters, function simultaneously during physiological development of the human placenta. Surprisingly, these obtained clusters can be grouped into 4 distinct groups of "mirrored" types of gene expression dynamics, meaning that any positive change in the gene expression in one cluster coincides with negative change in gene expression in the mirrored cluster and vice versa. This, together with the results of ORA and obtained PPI networks, allows to suggest can be said that genes in the obtained groups synergistically and antagonistically partake in processes such as immune system regulation, maintenance of physiological stability, cellular defense against toxins and pathogens, dynamic immune signaling, regulation of metabolism and reproductive processes, developmental and structural mechanisms. These generalizing interpretations of time series gene expression data, along with obtained individual gene expression dynamics, may provide a framework to study systematic changes in the human placenta during pregnancy from both a high-level (e.g. impacted biological functions) and a low-level (e.g. specific gene or protein) angle at once.

References:

1. Turco M. Y., Moffett A. Development of the human placenta. Development. 2019. Vol. 146, no. 22. DOI: 10.1242/dev.163428.

2. Dynamic genome-wide gene expression and immune cell composition in the developing human placenta / H. Suryawanshi et al. Journal of Reproductive Immunology. 2022. P. 103624. DOI: 10.1016/j.jri.2022.103624.

3. NCBI GEO: archive for functional genomics data sets–update / T. Barrett et al. Nucleic Acids Research. 2012. Vol. 41, no. D1. P. D991–D995. DOI: 10.1093/nar/gks1193.

4. Profiling Gene Expression in Human Placentae of Different Gestational Ages: An OPRU* Network and UW SCOR Study / A. M. Mikheev et al. Reproductive Sciences. 2008. Vol. 15, no. 9. P. 866–877. DOI: 10.1177/1933719108322425.

5. Increased placental expression and maternal serum levels of apoptosis-inducing TRAIL in recurrent miscarriage / K. Rull et al. Placenta. 2013. Vol. 34, no. 2. P. 141–148. DOI: 10.1016/j.placenta.2012.11.032.

6. The placental transcriptome of the first-trimester placenta is affected by in vitro fertilization and embryo transfer / L. Zhao et al. Reproductive Biology and Endocrinology. 2019. Vol. 17, no. 1. DOI: 10.1186/s12958-019-0494-7.

7. Dysregulation of the Circulating and Tissue-Based Renin-Angiotensin System in Preeclampsia / F. Herse et al. Hypertension. 2007. Vol. 49, no. 3. P. 604–611. DOI: 10.1161/01.HYP.0000257797.49289.71.

8. Epigenetics and Preeclampsia: Defining Functional Epimutations in the Preeclamptic Placenta Related to the TGF-β Pathway / E. Martin та ін. PLOS ONE. 2015. Т. 10, № 10. С. e0141294. DOI: 10.1371/journal.pone.0141294.

9. Onset of human preterm and term birth is related to unique inflammatory transcriptome profiles at the maternal fetal interface / R. Bukowski et al. PeerJ. 2017. Vol. 5. P. e3685. DOI: 10.7717/peerj.3685.

10. Differentially expressed microRNAs and affected biological pathways revealed by modulated modularity clustering (MMC) analysis of human preeclamptic and IUGR placentas / L. Guo et al. Placenta. 2013. Vol. 34, no. 7. P. 599–605. DOI: 10.1016/j.placenta.2013.04.007.

11. Mid-Gestational Gene Expression Profile in Placenta and Link to Pregnancy Complications / L. Uusküla et al. PLoS ONE. 2012. Vol. 7, no. 11. P. e49248. DOI: 10.1371/journal.pone.0049248.

12. Clustering gene expression time series data using an infinite Gaussian process mixture model / I. C. McDowell et al. PLOS Computational Biology. 2018. Vol. 14, no. 1. P. e1005896. DOI: 10.1371/journal.pcbi.1005896.

13. Gene Ontology overview. Gene Ontology Resource. URL: https://geneontology.org/docs/ontology-documentation/.

14. Shannon P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research. 2003. Vol. 13, no. 11. P. 2498–2504. DOI: 10.1101/gr.1239303.

15. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data / N. T. Doncheva et al. Journal of Proteome Research. 2018. Vol. 18, no. 2. P. 623–632. DOI: 10.1021/acs.jproteome.8b00702.

16. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data / T. Wu et al. The Innovation. 2021. Vol. 2, no. 3. P. 100141. DOI: 10.1016/j.xinn.2021.100141.