Gene review

Mira Welner

April 2024

Project Goal

The goal of this literature search is to identify a broad set of publications that use a wide variety of methodologies to identify genes that are differentially expressed between pre-eclamptia and non-pre-eclamptia samples. While there are many such papers, the process of identifying these genes is at times non-trivial. Our counts database uses the ensembl code nomenclature and many of the papers do not. The ensembl codes do not always have a one-to-one correlation with the genes identified in this paper. This search describes the process in which our genes were identified.

Jung Paper

Some publications listed genes in the same format accepted by ensembl and thus the translation was very straightforward. For example, one study on 82 pregnant women who underwent amniocentesis for fetal chromosomal analysis after finding a soft marker for Down syndrome [12] yeilded 10 genes that could be directly copied from table 4. The 10 genes consisted of 7 upregulated genes and 3 downregulated genes based on fold-change value in PE. Most of the upregulated genes were associated with the ribosome pathway.

Wang Paper

Another paper [10] utilized the GSE75010 profile database which contains 80 preeclampsia and 77 non-preeclampsia pregnancies, and used machine

learning to identify five genes for which the $\log(|FC|) > 0.4$. One of these genes, CXCL8 (ensembl code ENSG00000169429), caused a non-convergence error when run through PIRM. There is not obvious reason as to why and the counts were not overwhelmingly large numbers or counts of 0, but due to causing an error it was removed.

Textoris Paper

The Textoris paper[8] utilized a microarray to identify very many genes however there were far too many genes discovered (over 100) for there to be a lot of precision. However in the section "Transcriptional signature of severe PE" it listed genes that were associated with severe preeclamptia specifically. As opposed to what was identified using the microarray, these 13 genes had a very high correlation with severe preeclamptia and were very selected. This is where the ribosomal genes come from.

The gene EEF1B2 had two identical entries in ensembl: ENSG00000114942 and ENSG00000283391. Of these two, only ENSG00000114942 was found in our dataset so it was used.

Schnettler Paper

This publication[5] focused on the ratio between two genes: plasma soluble fms-like tyrosine kinase-1 (sFlt1) and placenta growth factor (PlGF). However they are very rarely refered to by sFlt1 and PlGF, the more common names for those two genes are FLT1 and PGF. They are synonyms according to GeneCard. This paper's database was Women (n=176) presenting to the hospital at ¡34 weeks' gestation for evaluation of possible preeclampsia during 2009–2010. Cases without complete cost or outcome data (n=9) and re-enrollments (n=18) were excluded.

Gray Paper

This publication[2] was less of an investigative study into various causes of preeclampsia and more of an analysis on how the gene PLEKHG1 was related to preeclamptia. The assumption was it was already relevant.

Yu Paper

Similar to the Gray publication, this study [11] was focusing on how the previously identified H19 gene was relevant to preeclampsia rather than identifying previously unknown genes. The H19 gene also appears in "The H19 Gene Imprinting in Normal Pregnancy and Pre-eclampsia". It is about H19. It is relevant that the H19 gene was also identified as relevant in the Tarca paper[7]. The Tarca paper was published subsequently to the Yu paper and used very different methods.

Zusterzeel Paper

The methods section of this paper[13] lists the three genes glutathione S-transferase P1, glutathione S-transferase, M1 and glutathione S-transferase T1. The more common names GSTP1, GSTM1, and GSTT1 were used in our database. However a fourth gene, epoxide hydrolase, is listed as EPHX. EPHX in not in ensemble however EPHX1, EPHX2, and EPHX3, which are three specific polymorphisms, are listed so these are included. The gene CYP1A1 translates directly.

Goddard Paper

This publication[1] utilized 775 SNPs in 190 candidate genes selected for a potential role in obstetrical complications. SNP discovery was performed by DNA sequencing, and genotyping was carried out in a high-throughput facility using the MassARRAY(TM) System. Women with PE (n = 394) and their offspring (n = 324) were compared with control women (n = 602) and their offspring (n = 631) from the same hospital-based population. Haplotypes were estimated for each gene using the EM algorithm, and empirical p values were obtained for a logistic regression-based score test, adjusted for significant covariates. An interaction model between maternal and offspring genotypes was also evaluated.

Peng Paper

This publication[4] used qRT-PCR testing to identify the PLEK and LEP genes. The LEP gene was seperately identified here[3].

Tarca Paper

This publication[7] profiled gene expression at exon-level resolution in whole blood samples collected longitudinally from 49 women with normal pregnancy (controls) and 13 with early preeclampsia (delivery ¡34 weeks of gestation). After preprocessing and removal of gestational age-related trends in gene expression, data were converted into Z-scores based on the mean and standard deviation among controls for six gestational-age intervals. The average Z-scores of mRNAs in each previously established signature considered herein were compared between cases and controls at 9-11, 11-17, 17-22, 22-28, 28-32, and 32-34 weeks of gestation. It also identifed the LEP gene.

Timofeeva Paper

This publication[9] was unusual in that rather than looking for standard genes it specifically looked for RNA. The MRNA samples mentioned in the paper are: miR-532-5p, miR-423-5p, miR-127-3p, miR-539-5p, miR-519a-3p, miR-629-5p miR-let-7c-5p miR-423-5p, miR-519a-3p, and miR-629-5p and miR-let-7c-5p. These are not in ensemble. However the genecards website translates them to MIR532, MIR423, MIR127, MIR539, MIR519a, MIR629, MIRLET7C, MIR423, MIR519a, MIR629, all of which were in ensembl except for MIR519a. This is because it has two alleles, MIR519a1 and MIR519a2 which were in ensembl and those two codes were used. the -5p and -3p indicated the 5' and 3' 'arms', and ensembl dataset classifies both RNA 'arms' together in the same entry, so these were ignored.

Seamon

Previous studies identified ERAP1 and ERAP2 as signficant and this study[6] confirmed it.

Kaartokallio

Thi sstudy used RNA sequencing to identify multiple genes. The nomenclature was straightforward except they refered to the gene PVRL4 which is more commonly known as NECTIN4 the gene LOC100129345 which is more commonly known as LINC02291. They were translated before being passed into ensemble.

Williams

This publication used clustering to identify genes related to preeclamsia as well as other hypertensive disorders. There was a bit of complexity in recording the genes however. The genes which related to preeclamptia were in table two. It refereed to Glycoprotein IIIa as GPIIIA, and the more common name is ITGB3 (they were marked as identical in genecards)

Other translations made through genecards were eNOS3 to NOS3, and VEGFR1 to FLT1.

One gene mentioned, vascular endothelial growth factor (VEGF) has four more specific version in ensembl: VEGFA, VEGFB, VEGFC, and VEGFD. Similarly, EPHX was split into EPHX1, EPHX2, and EPHX3

References

[1] Katrina A.B. Goddard, Gerard Tromp, Roberto Romero, Jennifer M. Olson, Qing Lu, Zhen Xu, Neelam Parimi, Jyh Kae Nien, Ricardo Gomez, Eric Behnke, Mario Solari, Jimmy Espinoza, Jaime Santolaya, Tinnakorn Chaiworapongsa, Guy M. Lenk, Karen Volkenant, Mary K. Anant, Benjamin A. Salisbury, Jeffrey Carr, M. Sean Lee, Gary F. Vovis, and Helena Kuivaniemi. Candidate-gene association study of mothers with pre-eclampsia, and their infants, analyzing 775 snps in 190 genes. *Human Heredity*, 63(1):1–16, 2007.

[2] Kathryn J. Gray, Vesela P. Kovacheva, Hooman Mirzakhani, Andrew C. Bjonnes, Berta Almoguera, Andrew T. DeWan, Elizabeth W. Triche, Audrey F. Saftlas, Joseph Hoh, Dale L. Bodian, Emily Klein, Katharine C. Huddleston, Sue Ann Ingles, Charles J. Lockwood, Hakon

Hakonarson, Thomas F. McElrath, Jeffrey C. Murray, Melissa L. Wilson, Errol R. Norwitz, S. Ananth Karumanchi, Brian T. Bateman, Brendan J. Keating, and Richa Saxena. Gene-centric analysis of preeclampsia identifies maternal association at plekhg1. *Hypertension*, 72(2):408–416, Aug 2018.

[3] Tea Kaartokallio, Alejandra Cervera, Anu Kyllönen, et al. Gene expression profiling of pre-eclamptic placentae by rna sequencing. *Scientific Reports*, 5:14107, 2015.

[4] Yu Peng, Hong Hong, Na Gao, et al. Bioinformatics methods in biomarkers of preeclampsia and associated potential drug applications. *BMC Genomics*, 23(711), 2022.

[5] William T. Schnettler, Dmitry Dukhovny, Julia Wenger, Saira Salahuddin, Steven J. Ralston, and Sarosh Rana. Cost and resource implications with serum angiogenic factor estimation in the triage of preeclampsia. *BJOG: An International Journal of Obstetrics & Gynaecology*, 120(10):1224–1232, Sep 2013.

[6] Kyle Seamon, Lesley O. Kurlak, Meghan Warthan, Efstratios Stratikos, Jerome F. III Strauss, Hiten D. Mistry, and Eun D. Lee. The differential expression of erap1/erap2 and immune cell activation in pre-eclampsia. *Frontiers in Immunology*, 11:396, Mar 10 2020.

[7] Adi L. Tarca, Roberto Romero, Offer Erez, Dereje W. Gudicha, Nandor Gabor Than, Neta Benshalom-Tirosh, Percy Pacora, Chaur-Dong Hsu, Tinnakorn Chaiworapongsa, Sonia S. Hassan, and Nardhy Gomez-Lopez. Maternal whole blood mrna signatures identify women at risk of early preeclampsia: a longitudinal study. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians*, 34(21):3463–3474, Nov 2021.

[8] Julien Textoris, Delphine Ivorra, Amina Ben Amara, Florence Sabatier, Jean-Pierre Ménard, Hélène Heckenroth, Florence Bretelle, and Jean-Louis Mege. Evaluation of current and new biomarkers in severe preeclampsia: A microarray approach reveals the vsig4 gene as a potential blood biomarker. *PLoS One*, 8(12):e82638, Dec 9 2013.

[9] Angelika V. Timofeeva, Vladyslava A. Gusar, Nataliya E. Kan, Kseniya N. Prozorovskaya, Anna O. Karapetyan, Oleg R. Bayev, Vitaliy V. Chagovets, Sergei F. Kliver, Daria Yu. Iakovishina, Vladimir E. Frankevich, and Gennadiy T. Sukhikh. Identification of potential early biomarkers of preeclampsia. *Placenta*, 61:61–71, 2018.

[10] Y. Wang, B. Li, and Y. Zhao. Inflammation in preeclampsia: Genetic biomarkers, mechanisms, and therapeutic strategies. *Frontiers in Immunology*, 13:883404, July 8 2022.

[11] L. Yu, M. Chen, D. Zhao, P. Yi, L. Lu, J. Han, X. Zheng, Y. Zhou, and L. Li. The h19 gene imprinting in normal pregnancy and pre-eclampsia. *Placenta*, 30(5):443–447, 2009.

[12] Jung YW, Shim JI, Shim SH, Shin YJ, Shim SH, Chang SW, and Cha DH. Global gene expression analysis of cell-free rna in amniotic fluid from women destined to develop preeclampsia. *Medicine (Baltimore)*, 98(3):e13971, 2019.

[13] Petra L. M. Zusterzeel, Wilbert H. M. Peters, Graham J. Burton, Willy Visser, Hennie M. J. Roelofs, and Eric A. P. Steegers. Susceptibility to pre-eclampsia is associated with multiple genetic polymorphisms in maternal biotransformation enzymes. *Gynecologic and Obstetric Investigation*, 63(4):209–213, 2007.