IECBM 2022

MDPI

Proceedings

In Silico Determination of Changes in Transcription factor Binding Sites for the Preeclampsia Risk Haplotype in the Regulatory Region of the FLT1 Gene †

Nataliia Karpova *, Olga Dmitrenco and Ekaterina Arshinova

Institute of General Pathology and Pathophysiology, Moscow, Russia (NIIOPP); nataliiakarpova.sp@gmail.com

* Correspondence: nataliiakarpova.sp@gmail.com;

† Presented at the 2nd International Electronic Conference on Biomolecules: Biomacromolecules and the Modern World Challenges, 1–15 Nov 2022; Available online: https://iecbm2022.sciforum.net/

Abstract: Preeclampsia (PE) is one of the most common complications of pregnancy that occurs in 3-8% of pregnant women, being one of the top five causes of maternal morbidity and mortality. We found that PE-associated polymorphisms near the FLT1 gene are located in the same regulatory region. In combination, these polymorphisms can be a genetic pattern which affects the development of pathology by forming a PE risk haplotype. When analyzing the changes in TPFS, which are characteristic only for the risk haplotype with a prevalence in the European population of 0.0825, we found that 5 TFBs change. The number of TFBSs for ELF1, SPIB increases, while the amount of TFBS POLR2A, KLF15 decreases (not expressed in the placenta).The newly emerged transcription factor binding site KAT5 acquires a promoter signature only after 118 days of pregnancy in the placenta, while before 118th day only DNase signature is observed. Theoretically, the appearance of a new TFBS can increase the expression of FLT1, causing an imbalance of angiogenic - antiangiogenic factors, characteristic of PE.

Keywords: Preeclampsia SNP; polymorphism; FLT1; risk haplotype; TFBs; TF

Citation: Karpova, N.; Dmitrenco O.; Archinova, K.; In silico determination of changes in transcription factor binding sites for the preeclampsia risk haplotype in the regulatory region of the FLT1 gene. *Biol. Life Sci. Forum* **2022**, *2*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Published: date

1. Introduction

One of the most serious pregnancy disorders, preeclampsia (PE), affects both the mother and the fetus [1]. According to Wang et al(2021) meta-analysis, preeclampsia affects 2–10% of pregnancies globally, with an average of 4.6 % for 2021 [2-3]. No matter the level of blood pressure in the patient's history, PE is characterized by a rise in blood pressure that typically occurs after the 20th week of pregnancy, with systolic blood pressure > 140 mm Hg Art and diastolic blood pressure > 90 mm Hg Art and proteinuria or at least one other parameter indicating the addition of multiple organ failure [1, 4-5].

Researchers at the Margaret Haig Maternity Hospital in New Jersey discovered the genetic predisposition to PE in the early 1960s. They found that preeclampsia is more prevalent in preeclamptic moms' sisters and daughters [6]. Later studies by Swedish researchers revealed that the heritability of PE is estimated to be around 55% [7]. But changes from the side of which gene or other genomic region remained unknown at that time.

The fact that trisomy of the 13 chromosomes has historically been associated with PE allows you to look for the reason for preeclampsia to develop on this chromosome [8]. The causes of preeclampsia were searched for by analyzing changes in the concentration of various proteins in the blood. At 15-20 weeks gestation, Flt-1 (which is located on 13

chromosomes), endoglin, and placental growth factor mRNA transcripts were examined by Sekizawa et al. (2010) in maternal whole blood [9]. Galaziou et al. (2021) and Lim et al. (2021) confirmed these data that PE had higher levels of sFlt-1 mRNA expression in their blood [10, 11]. Data obtained by Phupong et al. revealed that PE is the only pregnancy problem related with sFlt1/PlGF [12].

Genome-wide association studies (GWASs), were carried out by McGinnis et. al. (2019) originally discovered a connection between the T allele rs4769613 with preeclampsia (p-value of 5.4 10-11), which was later supported by many case-control association studies [13-14]. Allele T rs7318880 (p-value 8 x 10-8) in the analysis of the mother's genotype and allele C rs4769612 (p-value 4 x 10-14) in the analysis of the child's genotype were later found to be related with preeclampsia, according to Steinthorsdottir et al. (2020). Close linkage disequilibrium exists between the variants rs4769612 and rs7318880 (r2=0.98) [15]. All these Single Nucleotide Polymorphisms (SNPs) are located in the FLT1 gene regulatory area.

According to this data, the aim of our study was to determine regulatory potential of these SNPs, analysation of possible risk haplotype of PE, changing transcription factor binding sites (TFBSs) and evaluate getting information in FLT1 excess theory of PE etiology and pathogenesis.

2. Materials and Methods

2.1. Determination of SNPs associated with preeclampsia

We used the GWAS catalog [16] and PubMed to identify SNPs in FLT1 associated with preeclampsia in the maternal and child genomes.

2.2. Determination regulatory areas overlapped with SNPs associated with preeclampsia

Detected SNPs were mapped to the human genome in the UCSC genome browser to identify overlapping polymorphisms with regulatory regions [17].

2.3. Study of the enhancer signature of regulatory areas in the placenta

For determination of the enhancer signature of regulatory areas overlapped with SNPs was used cCRE details at ENCODE SCREEN [18].

2.4. Selection of SNPs

For SNPs in the regulatory region, we selected SNPs with a MAF of 1%, according UCSC genome browser Short Genetic Variants from dbSNP release 155 track [19].

2.5. Preeclampsia Risk Haplotype Determination

For selected SNPs were identified possible haplotypes in European populations using the LDhap Tool [20].

2.6. Obtaining the reference DNA sequence for the region of interest

For further analysis, the region chr13:28,563,912-28,564,785 (873 base pairs) was chosen, which overlaps with the regulatory regions of the FLT1 gene. The DNA sequence was obtained with the addition of 100 base pairs from the 5' and 3' ends using the Ensembl Genome Browser to assemble the human genome (GRCh38) [21]. To obtain sequences for all possible haplotypes, the corresponding nucleotides were changed manually using a standard text editor.

2.7. Determination of Change in Transcription Factor Binding Sites (TFBS)

Identification of changes in TTFS was performed using Human TFDB for the reference genotype and for the risk genotype, with further identification of the change from the side of TFBS only for the on - DNA strand, where the FLT1 gene was located.

Additionally, a comparison was made for the genotype with reference alleles, with the exception of rs4769612 and rs4769613 (for which the reference allele is a risk allele) [22].

2.8. Determination of Transcription Factor expression (TF)

Determination of the each TF expression were performed using The Human Protein Atlas [23].

3. Results and discussion

In the GWAS-catalog, we found 2 polymorphisms in FLT1 associated with preeclampsia: rs4769612-C (p-value 4 x 10-14) and rs7318880-T (p-value 8 x 10-8), with rs4769612 associated with preeclampsia in analysis of the child's genotype, and rs7318880 in the analysis of the mother's genotype. PubMed was able to detect one more SNP in the FLT1 gene regulatory area (rs4769613), associated with PE as a consequence of a genome-wide search for a relationship with PE. Additionally, the pathophysiology is more affected by fetal polymorphism than by maternal genetic variables [24].

The mapping of these SNPs was performed according to the latest assembly of the human genome (GRCh38/hg38) in UCSC genome browser. All 3 SNPs (rs7318880, rs4769612, rs4769613) are located in the regulatory region of FLT1 gene, according to the cCREs ENCODE project and oRegAnno. According to the UCSC genome browser (oRegAnno) data, there are 4 regulatory elements overlapped with SNPs: OREG1191996, OREG1658246, OREG1688336, OREG1537828. According to cCRE details at ENCODE SCREEN, this region contains the putative regulatory element EH38E1663332, the largest distal enhancer signature of which sharply increases at 16 weeks of gestation in the placenta and embryonic tissues, which can lead to changes in FLT1 expression [18]. In addition to rs4769612 and rs7318880, the EH38E1663332 region contains 7 more polymorphisms with MAF > 1%: rs7320190, rs12867370, rs4769613, rs74623647, rs7321138, rs76592233, rs9579193 (Table 1). Of these, only rs7320190 and rs4769613 are mentioned in scientific articles and rs4769612 and rs7318880 in the GWAS database, as well as rs12867370 associated with the risk of developing schizophrenia in offspring born to mothers with PE [25].

Table 1. Description for selected SNPs in FLT1 gene regulation areas (OREG1191996, OREG1658246, OREG1688336, OREG1537828 and EH38E1663332).

Gene	Variant (SNP)	Chr location (GRCh38.p13)	MAF*	Consequence	Allele Substitution
FLT1	rs7320190	chr13:28564119	0.20121	None	T > C
	rs7318880	chr13:28564148	0.50888	None	C > T
	rs12867370	chr13:28564261	0.06396	None	G > A
	rs4769612	chr13:28564361	0.45778	None	C > T
	rs4769613	chr13:28564472	0.475826	None	C > A, C > T
	rs74623647	chr13:28564495	0.00021	None	G > A, G > T
	rs7321138	chr13:28564568	0.187966	None	T > C, T > G
	rs76592233	chr13:28564624	0.00021**	None	C > A, C > G, C > T
	rs9579193	chr13:28564631	0.19905	None	G > A, G > T

* Minor Allele Frequency (MAF) for total population (Release Version: 20201027095038), according dbSNP. ** MAF for rs76592233 prevalence varies by source. According 1000Genomes for global population MAF of rs76592233 is 0.0104.

Searching possible haplotypes in European populations applying LDhap Tool allowed us to identify 4 possible haplotypes for given SNPs (Figure 1).

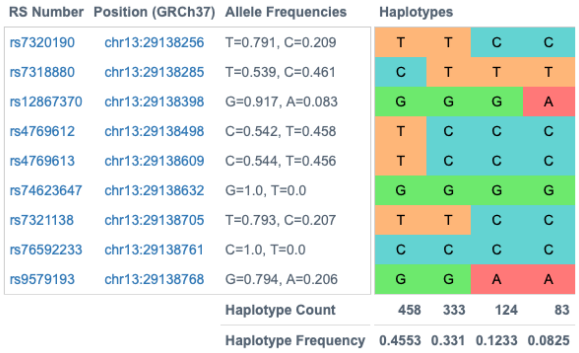

Figure 1. For polymorphisms rs7320190, rs7318880, rs12867370, rs4769612, rs4769613, rs74623647, rs7321138, rs76592233, and rs9579193, the prevalence of potential haplotypes was determined for EUR populations (SEU, TSI, FIN, GBR, IBS). In addition, the risk haplotype (C T A C C G C C A) occurs at 8.25%.

A potential preeclampsia risk haplotype with a prevalence of 8.25% (rs7320190-C, rs7318880-T, rs12867370-A, rs4769612-C, rs4769613-C, rs74623647-G, rs7321138-C, rs76592233-C). was identified after examining the prevalence of four different haplotypes (considering the risk alleles of maternal and fetal PE). Additionally, the frequency of this risk haplotype in the homozygous condition is 0.68%, which is comparable to the prevalence of early-onset preeclampsia, which is 0.38% [26].

Theoretically, the appearance of a new TFBS can increase the expression of FLT1, causing an imbalance of angiogenic - antiangiogenic factors, characteristic of PE. Therefore, we identified changes from 83 transcription factor binding sites (TFBs) to minus DNA strands in the analysis of all 4 possible haplotypes using HumanTFDB. According to Proteomic DB only 40 transcription factors (TFs) are expressed in the placenta. When analyzing the changes in TFBS, which are characteristic only for the risk haplotype with a prevalence in the European population of 0.0825, we found that 5 TFBs change (Table 2).

Table 2. TFBSs change for each haplotype and its expression according human protein atlas.

TFBS	**TFBSs change for each haplotype (hap.) with frequency**				**Type of TF**	**RNA expression in placenta (The human protein atlas, nTPM)**	**Protein expression in placenta (The human protein atlas, types of cells)**
	Hap. 1 (0.4553)*	**Hap. 2 (0.331)**	**Hap. 3 (0.1233)**	**Hap. 4 (0.0825)**			
KAT5	**0**	**0**	0	1	Activator, Acyltransferase, Chromatin regulator, Transferase	29.9	Decidual cells: Medium Trophoblastic cells: High
ELF1	2	2	2	3	Activator, DNA-binding	48.6	Trophoblastic cells: Medium
POLR2A	12	12	12	11		0.2	NA
KLF15	3	3	3	2	Activator, DNA-binding	1.7	not detected
SPIB	2	2	2	3	Activator, DNA-binding	0.2	Decidual cells: Medium

	Trophoblastic cells: Medium

The number of TFBSs for ELF1, SPIB increases, while the amount of TFBS POLR2A, KLF15 decreases (not expressed in the placenta).The newly emerged transcription factor binding site KAT5 acquires a promoter signature only after 118 days of pregnancy in the placenta, while before 118th day only DNase signature is observed. Theoretically, the appearance of a new TFBS can increase the expression of FLT1, causing an imbalance of angiogenic - antiangiogenic factors, characteristic of PE.

4. Conclusions

As a results we were able to identify a potential preeclampsia risk haplotype (C T A C C G C C A), which has a prevalence of 0.68% for homozygotes and a rate of 0.38% for the start of preeclampsia in its early stages.

Additionally, we discovered that the most critical event is the formation of a novel TFBS KAT5, for whose promoter only a DNase signature is seen in the placenta up to day 118 of pregnancy, after which it gains a promoter signature. According to theory, the emergence of a new TFBS can boost FLT1 expression, leading to an imbalance of angiogenic and antiangiogenic factors that is typical of PE.

Author Contributions: The following statements should be used "Conceptualization, K.N. and Y.Y.; methodology, K.N.; software, K.N.; validation, K.N.; formal analysis, K.N.; investigation, K.N.; resources, K.N.; data curation, K.N.; writing—original draft preparation, K.N.; writing—review and editing D.O., K.N. and A.E.; visualization, K.N.; supervision, A.E.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Please refer to suggested Data Availability Statements in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Acknowledgments: The authors express their gratitude to the supervisors of the phd study: Nurbekova Malik Kubanychbekov and Budykina Tatiana Sergeevna for their help with the research and comprehensive support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Preeclampsia. Eclampsia. Edema, proteinuria and hypertensive disorders during pregnancy, childbirth and the postpartum period, Clinical guidelines, Moscow, 2021 (in Russian)
2. Osungbade KO, Ige OK. Public health perspectives of preeclampsia in developing countries: implication for health system strengthening. J Pregnancy. 2011;2011:481095. doi: 10.1155/2011/481095. Epub 2011 Apr 4.
3. Wang, W., Xie, X., Yuan, T. et al. Epidemiological trends of maternal hypertensive disorders of pregnancy at the global, regional, and national levels: a population-based study. BMC Pregnancy Childbirth 21, 364 (2021).
4. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. Obstet Gynecol. 2013;122(5):1122-1131. doi:10.1097/01.AOG.0000437382.03963.88
5. ACOG Practice Bulletin No. 202 Summary: Gestational Hypertension and Preeclampsia. Obstet Gynecol. 2019;133(1):1. doi:10.1097/AOG.0000000000003019
6. CHESLEY LC, COSGROVE RA, ANNITTO JE. Pregnancies in the sisters and daughters of eclamptic women. Obstet Gynecol. 1962;20:39-46. doi:10.1097/00006250-196207000-00004

7. Lie RT, Rasmussen S, Brunborg H, Gjessing HK, Lie-Nielsen E, Irgens LM. Fetal and maternal contributions to risk of pre-eclampsia: population-based study. BMJ. 1998;316(7141):1343-1347. doi:10.1136/bmj.316.7141.1343
8. Maynard, S., Venkatesha, S., Thadhani, R. et al. Soluble Fms-like Tyrosine Kinase 1 and Endothelial Dysfunction in the Pathogenesis of Preeclampsia. Pediatr Res 57, 1–7 (2005). https://doi.org/10.1203/01.PDR.0000159567.85157.B7
9. Sekizawa A, Purwosunu Y, Farina A, et al. Prediction of pre-eclampsia by an analysis of placenta-derived cellular mRNA in the blood of pregnant women at 15-20 weeks of gestation. BJOG. 2010;117(5):557-564. doi:10.1111/j.1471-0528.2010.02491.xGalaziou A, Filidou E, Spathakis M, et al. Imbalance of growth factors mRNA expression associated with oxidative stress in the early pregnancy loss [published online ahead of print, 2021 Apr 5]. J Matern Fetal Neonatal Med. 2021;1-7. doi:10.1080/14767058.2021.1907337
10. Lim S, Li W, Kemper J, Nguyen A, Mol BW, Reddy M. Biomarkers and the Prediction of Adverse Outcomes in Preeclampsia: A Systematic Review and Meta-analysis. Obstet Gynecol. 2021;137(1):72-81. doi:10.1097/AOG.0000000000004149
11. Phupong V, Areeruk W, Tantbirojn P, Lertkhachonsuk R. Soluble fms-like tyrosine kinase 1 and placental growth factor ratio for predicting preeclampsia in elderly gravida. Hypertens Pregnancy. 2020 May;39(2):139-144. doi: 10.1080/10641955.2020.1745227. Epub 2020 Mar 22. PMID: 32202171
12. Ratnik K, Rull K, Aasmets O, Kikas T, Hanson E, Kisand K, Fischer K, Laan M. Novel Early Pregnancy Multimarker Screening Test for Preeclampsia Risk Prediction. Front Cardiovasc Med. 2022 Jul 27;9:932480. doi: 10.3389/fcvm.2022.932480. PMID: 35966513; PMCID: PMC9363612.
13. Kikas T, Inno R, Ratnik K, Rull K, Laan M. C-allele of rs4769613 Near FLT1 Represents a High-Confidence Placental Risk Factor for Preeclampsia. Hypertension. 2020 Sep;76(3):884-891. doi: 10.1161/HYPERTENSIONAHA.120.15346. Epub 2020 Aug 3. PMID: 32755415.
14. Steinthorsdottir V, McGinnis R, Williams NO, Stefansdottir L, Thorleifsson G, Shooter S, Fadista J, Sigurdsson JK, Auro KM, Berezina G, Borges MC, Bumpstead S, Bybjerg-Grauholm J, Colgiu I, Dolby VA, Dudbridge F, Engel SM, Franklin CS, Frigge ML, Frisbaek Y, Geirsson RT, Geller F, Gretarsdottir S, Gudbjartsson DF, Harmon Q, Hougaard DM, Hegay T, Helgadottir A, Hjartardottir S, Jääskelaïnen T, Johannsdottir H, Jonsdottir I, Juliusdottir T, Kalsheker N, Kasimov A, Kemp JP, Kivinen K, Klungsøyr K, Lee WK, Melbye M, Miedzybrodska Z, Moffett A, Najmutdinova D, Nishanova F, Olafsdottir T, Perola M, Pipkin FB, Poston L, Prescott G, Saevarsdottir S, Salimbayeva D, Scaife PJ, Skotte L, Staines-Urias E, Stefansson OA, Sørensen KM, Thomsen LCV, Tragante V, Trogstad L, Simpson NAB, Laivuori H, Heinonen S, Kajantie E, Kere J, Kivinen K, Pouta A, Morgan L, Pipkin FB, Kalsheker N, Walker JJ, Macphail S, Kilby M, Habiba M, Williamson C, O'Shaughnessy K, O'Brien S, Cameron A, Redman CWG, Farrall M, Caulfield M, Dominiczak AF, Aripova T, Casas JP, Dominiczak AF, Thorsteinsdottir U, Iversen AC, Feenstra B, Lawlor DA, Boyd HA, Magnus P, Zakhidova N, Laivuori H, Svyatova G, Stefansson K, Morgan L. Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. Nat Commun. 2020;11. doi:10.1038/s41467-020-19733-6
15. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, Suveges D, Vrousgou O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H, Flicek P, Burdett T, Hindorff LA, Cunningham F and Parkinson H. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research, 2019, Vol. 47 (Database issue): D1005-D1012.
16. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006
17. ENCODE Project Consortium, Jill E. Moore, Michael J. Purcaro, Henry E. Pratt, Charles B. Epstein, Noam Shoresh, Jessika Adrian, et al. 2020. "Expanded Encyclopaedias of DNA Elements in the Human and Mouse Genomes." Nature 583 (7818): 699–710.

18. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001 Jan 1;29(1):308-11. PMID: 11125122; PMC: PMC29783
19. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015 Jul 2.
20. Rigden, Daniel J., and Xosé M. Fernández. "The 2022 Nucleic Acids Research database issue and the online molecular biology database collection." *Nucleic acids research* 50.D1 (2022): D1-D10.
21. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. Hui Hu, Ya-Ru Miao, Long-Hao Jia, Qing-Yang Yu, Qiong Zhang and An-Yuan Guo*. Nucl. Acids Res. Volume 47, Issue D1, 08 January 2019, Pages D33-D38, https://doi.org/10.1093/nar/gky822
22. Karlsson M, Zhang C, Méar L, Zhong W, Digre A, Katona B, Sjöstedt E, Butler L, Odeberg J, Dusart P, Edfors F, Oksvold P, von Feilitzen K, Zwahlen M, Arif M, Altay O, Li X, Ozcan M, Mardinoglu A, Fagerberg L, Mulder J, Luo Y, Ponten F, Uhlén M, Lindskog C. A single-cell type transcriptomics map of human tissues. Sci Adv. 2021 7(31):. PubMed: 34321199 DOI: 10.1126/sciadv.abh2169
23. McGinnis R, Steinthorsdottir V, Williams NO, et al. Variants in the fetal genome near FLT1 are associated with risk of preeclampsia. *Nat Genet.* 2017;49(8):1255-1260. doi:10.1038/ng.3895
24. Dolores Malaspina, Thorsten Kranz, Karine Kleinhaus, Sulaima Daboul, Karen Rothman, Caitlin Gilman, Mara Getz, Susan Harlap, Yechiel Friedlander. Short duration of marriage at conception as an independent risk factor for schizophrenia. Schizophrenia Research, 2019; DOI: 10.1016/j.schres.2019.03.001
25. Lisonkova S., Joseph K.S. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *Am J Obstet Gynecol*. 2013;209(6):544.e1-544.e12.