Published in final edited form as:

Hum Genet. 2011 January ; 129(1): 25–34. doi:10.1007/s00439-010-0893-5.

Partial correlation network analyses to detect altered gene interactions in human disease: using preeclampsia as a model

Åsa Johansson,

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Kvinne-barn senteret, 1.etg. Øst, 7006 Trondheim, Norway; Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Mari Løset,

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Kvinne-barn senteret, 1.etg. Øst, 7006 Trondheim, Norway

Siv B. Mundal,

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Kvinne-barn senteret, 1.etg. Øst, 7006 Trondheim, Norway

Matthew P. Johnson,

Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Katy A. Freed,

Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Mona H. Fenstad,

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Kvinne-barn senteret, 1.etg. Øst, 7006 Trondheim, Norway

Eric K. Moses,

Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Rigmor Austgulen, and

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Kvinne-barn senteret, 1.etg. Øst, 7006 Trondheim, Norway

John Blangero

Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Abstract

Differences in gene expression between cases and controls have been identified for a number of human diseases. However, the underlying mechanisms of transcriptional regulation remain largely unknown. Beyond comparisons of absolute or relative expression levels, disease states may be associated with alterations in the observed correlational patterns among sets of genes. Here we use partial correlation networks aiming to compare the transcriptional co-regulation for 222 genes that are differentially expressed in decidual tissues between preeclampsia (PE) cases and non-PE controls. Partial correlation coefficients (PCCs) have been calculated in cases ($N = 37$) and controls ($N = 58$) separately. For all PCCs, we tested if they were significant non-zero in the cases and controls separately. In addition, to examine if a given PCC is different between the cases and

controls, we tested if the difference between two PCCs were significant non-zero. In the group with PE cases, only five PCCs were significant (FDR p value ≤ 0.05), of which none were significantly different from the PCCs in the controls. However, in the controls we identified a total of 56 statistically significant PCCs (FDR p value ≤ 0.05), of which 31 were also significantly different (FDR p value ≤ 0.05) from the PCCs in the PE cases. The identified partial correlation networks included genes that are potentially relevant for developing PE, including both known susceptibility genes (*EGFL7*, *HES1*) and novel candidate genes (*CFH*, *NADSYN1*, *DBP*, *FIGLA*). Our results might suggest that disturbed interactions, or higher order relationships between these genes play an important role in developing the disease.

Introduction

Preeclampsia (PE), with more than four million incidences and 50,000 deaths per year, is a leading cause of maternal and perinatal mortality in the world (Lain and Roberts 2002; Roberts et al. 2003). PE is a pregnancy-specific disorder, diagnosed by new onset of hypertension and proteinuria in the latter half of pregnancy (Roberts and Gammill 2005; Roberts et al. 2003; Witlin and Sibai 1997). There is neither a reliable predictive test, nor effective treatment for PE available, except delivery of the child and placenta. Consequently, PE accounts for approximately 20% of all preterm births (Goldenberg and Rouse 1998).

PE is a common complex disease that involves the contribution of multiple genetic and environmental components and their interactions and the heritability has been estimated to 0.54 (Salonen Ros et al. 2000). One potential source of such genetic effects may involve the regulation of gene expression that may play a major role in the development of this disorder. Transcriptional profiling has recently been shown to be a powerful method to identify genetic variants influencing human traits (Goring et al. 2007). Many genes have also been shown to be differentially expressed in the placental or decidual tissue between women with PE and women with normal pregnancies (Enquobahrie et al. 2008; Winn et al. 2009).

Despite the many genes that have been suggested to be differentially expressed between groups of individuals, little is known about the underlying transcriptional regulation of these genes. Studies of differential expression focus specifically on differences in mean levels (either absolute or relative). However, the relationship between transcripts (as measured by covariances or correlations) may also be altered in disease states. These higher order effects on correlations may reflect alterations in network coherence that is important for understanding the biological nature of pathological mechanisms.

Correlation between levels in genome-wide transcription data can be substantiated and illustrated by correlation networks. However, correlation networks only describe the correlation between the expression levels, without any information regarding causality. It is difficult to determine the primary dependencies among many transcripts, and the true interactions between genes are still poorly understood. One reason why these interactions are hard to identify is the overfitting problem generated by measuring a large number of transcripts in a small number of samples. However, for smaller sets of genes exhibiting altered transcription, it is possible to examine the complex inter-relationships amongst genes more closely. One way to extract information on these interactions is to use partial correlation analysis. Correlation networks represent correlation coefficients that are calculated for pairs of transcripts, regardless of all other variables (transcripts). In contrast, partial correlation networks are represented by partial correlation coefficients that are calculated for pairs of transcripts when all other variables (transcripts) are taken into account. Consequently, partial correlations represent direct associations, whereas correlation analyses do not distinguish between indirect and direct associations. In addition, partial

correlation analyses may allow identifying the direction of a partial correlation, which enables to distinguish between the response variables and the covariates. Partial correlation analysis is an appropriate method to use to detect relationship between genes in datasets when the number of variables (e.g. transcripts) is much larger than the sample size. The partial correlation approach has successfully been developed and applied to transcription data from yeast (de la Fuente et al. 2004; Han and Zhu 2008), *Arabidopsis thaliana* (Ma et al. 2007; Opgen-Rhein and Strimmer 2007), HeLa cells (Fujita et al. 2007), and breast cancer tumors (Schafer and Strimmer 2005a), but to our knowledge, it has not been used to study the difference in gene interactions between case and control groups of any human disease.

Recently, analyses of transcription data from placental or decidual tissue in PE case/control cohorts have identified a large number of genes to be differentially expressed (Enquobahrie et al. 2008; Løset et al. 2010; Winn et al. 2009). In this study, we have examined the partial correlation structure of differentially expressed genes by comparing partial correlation networks between PE cases and controls, in order to identify PE-associated alterations in gene–gene interaction or co-regulation of genes.

Materials and methods

Study groups

The women were recruited at St. Olavs' University Hospital (Trondheim, Norway) and Haukeland University Hospital (Bergen, Norway) from 2002 to 2006. PE was defined as persistent hypertension (blood pressure of $\geq 140/90$ mmHg) plus proteinuria (≥ 0.3 g/day or $\geq 1+$ according to a dipstick test), developing after 20 weeks of pregnancy (Gifford et al. 2000). Descriptive characteristics can be found in Table 1 and information about the study population has been published previously (Fenstad et al. 2010; Løset et al. 2010). Women with preeclamptic pregnancies ($N = 37$) had cesarean section performed for medical indications, and none of them were in labor prior to cesarean section. Exclusively healthy women with no history of PE were accepted in the control group ($N = 58$). Cesarean section was also carried out on everyone in the control group for reasons such as breech presentation, previous cephalopelvic disproportion in previous delivery or maternal choice. Controls with previous PE and/or fetal growth restriction were excluded. Multiple pregnancies, pregnancies with chromosomal aberrations, fetal and placental structural abnormalities, or suspected perinatal infections were excluded. Informed consent was obtained from all participants prior to collection of decidual samples and the study was approved by the Norwegian Regional Committee for Medical Research Ethics.

Sample preparation

Decidua basalis samples were obtained by vacuum suction of the placental bed and immediately incubated in RNA-later and stored at -80°C . RNA was isolated by Trizol extraction protocol and purified with an RNeasy Mini Kit using spin technology (Qiagen, Valencia, CA, USA). Ethical approval for total RNA processing and decidua expression analysis was obtained from the Institutional Review Board at The University of Texas Health Science Center in San Antonio. Illumina TotalPrep RNA Amplification Kit was used for synthesizing, amplifying, and purifying anti-sense RNA according to the manufacturer's instructions (Ambion, Austin, TX, USA). Specific information pertaining to the decidual sample collection, sample preparation and transcriptional profiling can be found at ArrayExpress (<http://www.ebi.ac.uk/microarray-as/ae/>; accession code E-TABM-682).

Transcriptional profiling

Transcriptional profiling were performed on 37 cases and 58 controls using Illumina's Human-6 v2 Expression Bead-Chip® (Illumina, San Diego, CA, USA) according to Illumina's standard protocols. For each sample, 48,095 transcripts were interrogated. Illumina's BeadStudio Gene Expression software module (version 3.2.7) was used to subtract background noise signals and generate an output file. In total, 26,504 transcripts were detected and used for statistical analysis. To make all transcription values comparable across individuals as well as across genes, these were normalized as described previously (Goring et al. 2007).

Statistical analyses

Out of the 26,504 transcripts detected, differentially expressed genes were identified using the stats library in R. A linear regression model (transcription level ~ PE-status + RIN, where RIN is the RNA integrity number) was fitted for each transcript using the `lm()` function. Summary statistics were computed for the fitted linear model using the `summary.lm()` function, where the p values were extracted based on t statistics. FDR (false discovery rate) were calculated using the `fdrtool()` function implemented in the `fdrtool` R library (Strimmer 2008). For downstream statistical analyses only the differentially expressed genes (FDR p value ≤ 0.05), between PE cases and controls, were included.

Partial correlation coefficients (PCCs) between all pairs of transcripts and the significance of the derived edges and directions were calculated for cases and controls separately, using a graphical Gaussian model for small sample estimation of partial correlation implemented in the GeneNet R package (version 1.2.0) (Opge-Rhein and Strimmer 2007; Schafer and Strimmer 2005b). For all pairs of transcripts, the difference between the PCCs in the cases and controls ($PCC_{cases} - PCC_{controls}$) were calculated. To test if this difference were significantly non-zero, we used the same approach as for the PCCs (Schafer and Strimmer 2005a), which is implemented in GeneNet. This approach, which can be used on small samples sizes with many covariates (transcripts), includes an estimation of the degree of freedom of the null distribution from the data and a subsequent estimation of p values and FDR p values (Hochberg and Benjamini 1990; Strimmer 2008). In order to control for the many tests performed, an FDR of 0.05 was employed as the cut-off for statistical significance in all comparisons.

Results

In total, 222 transcripts were identified as being differentially expressed (FDR p value ≤ 0.05) between PE cases and controls (Supplementary material Table S1). For these 222 transcripts, 24,531 ($222 \times 221/2$) pairwise partial correlations were estimated between the transcription levels.

Partial correlations were estimated for the PE cases and controls separately and the number of non-zero partial correlations (unadjusted p value ≤ 0.05) was 1,832 in the control and 1,549 in the case group. However, after adjusting for multiple testing, 56 of the partial correlations were significant non-zero (FDR p value ≤ 0.05) in the control group, compared to five in the case group (Table 2). The 56 significant partial correlations in the controls represent 75 transcripts and can be divided into 19 non-overlapping networks (Fig. 1). In contrast, the five partial correlations observed in the case group represent ten transcripts, and can be connected in a pairwise manner into five non-overlapping partial correlation networks (Fig. 2).

None of the directions for the edges in the partial correlation networks were significant (FDR p value ≤ 0.05) in cases or controls, but three of the directions of the edges were

nominally significant ($p \leq 0.05$) in the controls (Fig. 1). The overlapping set of significant partial correlations between PE cases and controls includes only two pairwise relationships: a positive partial correlation between *TMEM140* and *APOL3* and another positive partial correlation between two probes in *SPCS2*.

We further statistically compared the difference between the PCCs between PE cases and controls. Among all PCCs, 115 were significantly different (FDR p value ≤ 0.05) between cases and controls. However, only 31 of these PCCs were also significant (FDR p value ≤ 0.05) in either cases or controls (Fig. 3). Interestingly, all 31 of these transcript pairs exhibited significant partial correlations in the controls (Fig. 3), whereas none of the partial correlations were significant in the cases. Thus, the disease state appears to be associated with a diminished correlation amongst this set of genes. The 31 partial correlations that were present in the controls were distributed over 16 unconnected gene-clusters (networks) including a total of 46 genes (Fig. 3).

Discussion

While differentially expressed genes between case and control groups for different human diseases has been routinely identified, this is the first study to formally assess differences in the observed partial correlation patterns of mRNA levels or gene co-regulation using transcription data from PE disease cases and controls. Such interactions between genes are known to play an important role in the development of human diseases, and recently a number of gene–gene interactions have been identified as determinants of several different diseases, including prostate cancer (Cooper et al. 2008), age-related macular degeneration (Seitsonen et al. 2008), ischemic stroke (Szolnoki et al. 2009), and Alzheimer's disease (Infante et al. 2010).

In this study, we have identified partial correlations in gene expression that differ substantially between PE cases and controls. We identified 5 significant partial correlations in the case versus 56 in the control group. However, some partial correlations that are significant in one of the groups can be just below the threshold of significance in the other group. To examine if a given PCC is different between the groups, we tested if the difference between two PCCs were significantly non-zero. In total, we identified 31 partial correlations that were significant only in the control group and that were also statistically different from the case group. These observed partial correlations may reflect gene–gene interactions as well as co-regulation of the genes. However, it is interesting that we only identified partial correlations that were present in the controls but absent in the cases, not vice versa. This might reflect a general loss of regulation or genetic coherence resulting from the disease. Alternatively, and more interestingly, such a dysregulation may represent a causal factor that directly contributes to the development and risk of PE. However, since we are assessing the difference in gene expression after disease onset we cannot separate these two alternatives, and prospective studies are needed.

While some of the partial correlations that we have identified occurred between genes that have relatively unknown functions, others that have functions known to be relevant for the development of PE will be discussed briefly below. One interesting partial correlation is between *CFH* (complement factor H) and *NADSYN1* (nicotinamide adenine dinucleotide synthetase1). *CFH* is known to be a regulator of the alternative pathway of the complement system, which plays an important role in the innate immunity against microbial infection. *CFH* has also been suggested to be crucial to the inflammatory response and atherosclerosis (Oksjoki et al. 2003). Generalized inflammatory reaction, including the complement system is central in PE pathogenesis (Redman and Sargent 2003). However, a polymorphism in *CFH* (Y402H) that has been shown to be a major risk factor for age-related macular

degeneration in several populations (Edwards et al. 2005; Haines et al. 2005; Klein et al. 2005; Seitsonen et al. 2006), was not associated with PE in a previous study (Kaare et al. 2008). SNPs in the *NADSYN1* gene have been associated with the circulating level of vitamin D (Ahn et al. 2010) and vitamin D deficiency has been shown to increase the risk of PE (Bodnar et al. 2007). Circulating D vitamin levels has even been suggested to explain the link between PE and cardiovascular diseases (Grant 2010).

A positive partial correlation was also found between *FIGLA* (folliculogenesis specific basic helix–loop–helix) and *DBP* [D site of albumin promoter (albumin D-box) binding protein]. *FIGLA* is required for normal folliculogenesis and mutations in the gene have been associated with premature ovarian failure (Zhao et al. 2008). *FIGLA* is known to be a transcription factor that regulates other oocyte-specific genes (Bayne et al. 2004; Huntriss et al. 2002; Liang et al. 1997). *DBP* is also a transcription factor whose expression has been shown to be induced by both estrogen and testosterone (Eyster et al. 2007). *DBP* is known to regulate the expression levels of factor IX (Boccia et al. 1996), which is an essential vitamin K-dependent serine protease that participates in the intrinsic pathway of coagulation. Even though none of these genes have any direct correlation to the characteristics of PE, the function of *FIGLA* in folliculogenesis and the fact that *DBP* is hormone induced still makes them interesting empirical candidate genes in relation to a pregnancy-specific disease such as PE.

The partial correlations described above are between genes with known functions that can to some extent be related to the development of PE. Most of the cluster of partial correlations, however, includes none or only one gene with a function that can be linked to PE. *EGFL7* (Epidermal growth factor-like domain 7) has been suggested to protect endothelial cells from hyperoxia-induced cell death, which is a possible link between *EGFL7* and PE. However, for the other genes in the same cluster as *EGFL7* (Fig. 3), we cannot find any function related to PE. Similarly, *HES1* is part of the Notch signaling pathway, which has been suggested to be an important factor in the development to PE (Cobellis et al. 2007).

One of the major goals using partial correlation analyses is to elucidate if the transcription level between genes are directly correlated or whether they are mediated by an additional gene or other factor. Unfortunately, we were not able to assess the directions of the partial correlations in our study and consequently we cannot determine if one gene in a network directly influences the transcription level of another. Each significant PCC only indicate that there is a direct correlation between the transcription levels of two genes. This correlation might be due to the fact one gene regulates the expression of the other, or due to co-regulation of these genes. However, we are not able to determine whether this co-regulator is an environmental factor such as disease status, or if it is another gene that has not been targeted in our study. It is also known that gene–gene interactions might influence disease status and the PCCs identified might as well represent gene–gene interactions that are important to promote the development of PE.

One limitation in our study is the fact that the PE cases and controls are not matched regarding gestational age (Table 1). It has been reported that the expression level changes dramatically for some genes between mid gestation and term (Winn et al. 2007). While cesarean section were performed earlier in our PE group (mean = 31.9 weeks, standard deviation = 3.9) compared to our control group (mean = 38.7 weeks, standard deviation = 0.8), this needs to be considered as a possible bias in our analyses. However, among the genes discussed in this manuscript (Table 2) only *SULT2B1*, *IL6STm*, and *SRPRB* have been shown to differ between term and mid gestation previously (Winn et al. 2007). Other factors that might influence gene expression are the presence of labor prior to cesarean section, parity, and infant sex. However, in our study cesarean section were carried out in the

absence of labor and we do not find any significant difference (Table 1) between the case and control group, neither regarding sex of the infants nor the fraction of primiparous.

A factor that might influence the number of significant PCCs is higher in the controls compared to the cases might be somewhat related to the larger sample size (58 controls compared to 37 cases). By repeating a resampling of a random set of controls ($N = 37$) 100 times, calculating PCCs and assessing the significance of the edges, we still found significantly higher number [Wilcoxon (Mann–Whitney) signed rank test p value $\ll 0.05$] of significant PCCs in controls (mean = 41.5, standard deviation = 15.2). The lower number of genes that are upregulated compared to downregulated in the PE cases (Supplemental Table S1) might also influence the difference in number of significant PCCs between the two groups.

The problem with non-independent or circular analyses has gained a lot of interest lately (Kriegeskorte et al. 2009, 2010). We should point out that the transcripts used for the partial correlation analyses were not independently selected. Due to the overfitting problem and restrictions in computer capacity, it is not possible to perform partial correlation analyses on the complete dataset including 26,504 transcripts. Consequently, we have chosen to analyze 222 transcripts, representing the ones that were differentially expressed (FDR p value ≤ 0.05) between the case and control group. The selection process of which subset of transcripts to analyze will influence the results dramatically. If a low threshold is used for this selection process, also the number of false positives might be elevated. However, using a stringent threshold for which transcripts included in the analyses with many true partial correlations that play an important role in the development of a disease might be ignored.

In summary, using partial correlation networks we have identified partial correlations that are present in non-PE decidua tissue, but absent in PE cases. To our knowledge, this is a novel approach to identify possible disease-related mechanisms in human disease. Our findings suggest that beyond the transcription levels of different genes, the pattern of observed gene–gene partial correlations representing the interaction or co-regulation of genes may also be important in the development of disease. However, to determine if the partial correlations seen are due to causal alterations in gene–gene interactions or co-regulation will require further prospective and functional studies of these genes and gene products.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by grants from FUGE—Functional genomics in Norway within the Research Council of Norway (Åsa Johansson) and motility grants from NTNU—The Norwegian University of Science and Technology (Åsa Johansson) and Sven och Dagmar Saléns stiftelse (Åsa Johansson). Statistical analysis was supported in part by NIH grant MH59490 (Blangero).

References

- Ahn J, Yu K, Stolzenberg-Solomon R, Simon KC, McCullough ML, Gallicchio L, Jacobs EJ, Ascherio A, Helzlsouer K, Jacobs KB, Li Q, Weinstein SJ, Purdue M, Virtamo J, Horst R, Wheeler W, Chanock S, Hunter DJ, Hayes RB, Kraft P, Albanes D. Genome-wide association study of circulating vitamin D levels. *Hum Mol Genet.* 2010; 19:2739–2745. doi: 10.1093/hmg/ddq155. [PubMed: 20418485]

- Bayne RA, Martins da Silva SJ, Anderson RA. Increased expression of the FIGLA transcription factor is associated with primordial follicle formation in the human fetal ovary. *Mol Hum Reprod.* 2004; 10:373–381. [PubMed: 15044608]
- Boccia LM, Lillicrap D, Newcombe K, Mueller CR. Binding of the Ets factor GA-binding protein to an upstream site in the factor IX promoter is a critical event in transactivation. *Mol Cell Biol.* 1996; 16:1929–1935. [PubMed: 8628259]
- Bodnar LM, Catov JM, Simhan HN, Holick MF, Powers RW, Roberts JM. Maternal vitamin D deficiency increases the risk of preeclampsia. *J Clin Endocrinol Metab.* 2007; 92:3517–3522. doi: 10.1210/jc.2007-0718. [PubMed: 17535985]
- Cobellis L, Mastrogiacomio A, Federico E, Schettino MT, De Falco M, Manente L, Coppola G, Torella M, Colacurci N, De Luca A. Distribution of Notch protein members in normal and preeclampsia-complicated placentas. *Cell Tissue Res.* 2007; 330:527–534. doi: 10.1007/s00441-007-0511-6. [PubMed: 17955263]
- Cooper ML, Adami HO, Gronberg H, Wiklund F, Green FR, Rayman MP. Interaction between single nucleotide polymorphisms in selenoprotein P and mitochondrial superoxide dismutase determines prostate cancer risk. *Cancer Res.* 2008; 68:10171–10177. [PubMed: 19074884]
- de la Fuente A, Bing N, Hoeschele I, Mendes P. Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics.* 2004; 20:3565–3574. [PubMed: 15284096]
- Edwards AO, Ritter R 3rd, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science.* 2005; 308:421–424. doi: 10.1126/science.1110189. [PubMed: 15761121]
- Enquobahrie DA, Meller M, Rice K, Psaty BM, Siscovick DS, Williams MA. Differential placental gene expression in preeclampsia. *Am J Obstet Gynecol.* 2008; 199(566):e1–e11. [PubMed: 18533121]
- Eyster KM, Mark CJ, Gayle R, Martin DS. The effects of estrogen and testosterone on gene expression in the rat mesenteric arteries. *Vascul Pharmacol.* 2007; 47:238–247. [PubMed: 17652031]
- Fenstad MH, Johnson MP, Loset M, Mundal SB, Roten LT, Eide IP, Bjorge L, Sande RK, Johansson AK, Dyer TD, Forsmo S, Blangero J, Moses EK, Austgulen R. STOX2 but not STOX1 is differentially expressed in decidua from preeclamptic women. *Mol Hum Reprod.* 2010 doi: 10.1093/molehr/gaq064.
- Fujita A, Sato JR, Garay-Malpartida HM, Yamaguchi R, Miyano S, Sogayar MC, Ferreira CE. Modeling gene expression regulatory networks with the sparse vector autoregressive model. *BMC Syst Biol.* 2007; 1:39. [PubMed: 17761000]
- Gifford RW, August PA, Cunningham G, Green LA, Lindheimer MD, McNellis D, Roberts JM, Sibai BM, Taler SJ. Report of the National High Blood Pressure Education Program Working Group on high blood pressure in pregnancy. *Am J Obstet Gynecol.* 2000; 183:S1–S22.
- Goldenberg RL, Rouse DJ. Prevention of premature birth. *N Engl J Med.* 1998; 339:313–320. [PubMed: 9682045]
- Goring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JB, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, MacCluer JW, Kissebah AH, Collier GR, Moses EK, Blangero J. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet.* 2007; 39:1208–1216. [PubMed: 17873875]
- Grant WB. Low vitamin D may explain the link between preeclampsia and cardiovascular disease. *Am Heart J.* 2010; 159:e19. doi: 10.1016/j.ahj.2009.12.007. [PubMed: 20211290]
- Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Nouredine M, Gilbert JR, Schnetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA. Complement factor H variant increases the risk of age-related macular degeneration. *Science.* 2005; 308:419–421. doi: 10.1126/science.1110359. [PubMed: 15761120]
- Han L, Zhu J. Using matrix of thresholding partial correlation coefficients to infer regulatory network. *Biosystems.* 2008; 91:158–165. [PubMed: 17919808]
- Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med.* 1990; 9:811–818. [PubMed: 2218183]

- Huntriss J, Gosden R, Hinkins M, Oliver B, Miller D, Rutherford AJ, Picton HM. Isolation, characterization and expression of the human factor in the germline alpha (FIGLA) gene in ovarian follicles and oocytes. *Mol Hum Reprod*. 2002; 8:1087–1095. [PubMed: 12468641]
- Infante J, Rodriguez-Rodriguez E, Mateo I, Llorca J, Vazquez-Higuera JL, Berciano J, Combarros O. Gene-gene interaction between heme oxygenase-1 and liver X receptor-beta and Alzheimer's disease risk. *Neurobiol Aging*. 2010; 31:710–714. doi: 10.1016/j.neurobiolaging.2008.05.025. [PubMed: 18597895]
- Kaare M, Seitsonen S, Jarvela I, Meri S, Laivuori H. Complement factor H variant Y402H is not a risk factor for preeclampsia in the Finnish population. *Hypertens Pregnancy*. 2008; 27:328–336. doi: 10.1080/10641950801955691. [PubMed: 19003634]
- Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005; 308:385–389. doi: 10.1126/science.1109557. [PubMed: 15761122]
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 2009; 12:535–540. doi: 10.1038/nn.2303. [PubMed: 19396166]
- Kriegeskorte N, Lindquist MA, Nichols TE, Poldrack RA, Vul E. Everything you never wanted to know about circular analysis, but were afraid to ask. *J Cereb Blood Flow Metab*. 2010; 30:1551–1557. doi: 10.1038/jcbfm.2010.86. [PubMed: 20571517]
- Lain KY, Roberts JM. Contemporary concepts of the pathogenesis and management of preeclampsia. *JAMA*. 2002; 287:3183–3186. [PubMed: 12076198]
- Liang L, Soyal SM, Dean J. FIGalpha, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. *Development*. 1997; 124:4939–4947. [PubMed: 9362457]
- Løset M, Mundal SB, Johnson MP, Fenstad MH, Freed KA, L B, Blangero J, Moses EK, Austgulen R. A transcriptional profile of the decidua in preeclampsia. *AJOG*. 2010 doi: 10.1016/j.ajog.2010.08.043.
- Ma S, Gong Q, Bohnert HJ. An Arabidopsis gene network based on the graphical Gaussian model. *Genome Res*. 2007; 17:1614–1625. [PubMed: 17921353]
- Oksjoki R, Kovanen PT, Pentikainen MO. Role of complement activation in atherosclerosis. *Curr Opin Lipidol*. 2003; 14:477–482. doi: 10.1097/01.mol.0000092627.86399.7b. [PubMed: 14501586]
- Opgen-Rhein R, Strimmer K. From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Syst Biol*. 2007; 1:37. [PubMed: 17683609]
- Redman CW, Sargent IL. Pre-eclampsia, the placenta and the maternal systemic inflammatory response—a review. *Placenta*. 2003; 24(Suppl A):S21–S27. [PubMed: 12842410]
- Roberts JM, Gammill HS. Preeclampsia: recent insights. *Hypertension*. 2005; 46:1243–1249. [PubMed: 16230510]
- Roberts JM, Pearson G, Cutler J, Lindheimer M. Summary of the NHLBI Working Group on Research on hypertension during pregnancy. *Hypertension*. 2003; 41:437–445. [PubMed: 12623940]
- Salonen Ros H, Lichtenstein P, Lipworth L, Cnattingius S. Genetic effects on the liability of developing pre-eclampsia and gestational hypertension. *Am J Med Genet*. 2000; 91:256–260. [PubMed: 10766979]
- Schafer J, Strimmer K. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics*. 2005a; 21:754–764. [PubMed: 15479708]
- Schafer J, Strimmer K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat Appl Genet Mol Biol*. 2005b; 4 Article 32.
- Seitsonen S, Lemmela S, Holopainen J, Tommila P, Ranta P, Kotamies A, Moilanen J, Palosaari T, Kaarniranta K, Meri S, Immonen I, Jarvela I. Analysis of variants in the complement factor H, the elongation of very long chain fatty acids-like 4 and the hemicentin 1 genes of age-related macular degeneration in the Finnish population. *Mol Vis*. 2006; 12:796–801. [PubMed: 16885922]

- Seitsonen SP, Onkamo P, Peng G, Xiong M, Tommila PV, Ranta PH, Holopainen JM, Moilanen JA, Palosaari T, Kaarniranta K, Meri S, Immonen IR, Jarvela IE. Multifactor effects and evidence of potential interaction between complement factor H Y402H and LOC387715 A69S in age-related macular degeneration. *PLoS ONE*. 2008; 3:e3833. [PubMed: 19048105]
- Strimmer K. A unified approach to false discovery rate estimation. *BMC Bioinformatics*. 2008; 9:303. doi: 10.1186/1471-2105-9-303. [PubMed: 18613966]
- Szolnoki Z, Maasz A, Magyari L, Horvatovich K, Farago B, Kondacs A, Bodor A, Hadarits F, Orosz P, Ille A, Melegh B. Galectin-2 3279TT variant protects against the lymphotoxin-alpha 252GG genotype associated ischaemic stroke. *Clin Neurol Neurosurg*. 2009; 111:227–230. doi: 10.1016/j.clineuro.2008.09.027. [PubMed: 19013708]
- Winn VD, Haimov-Kochman R, Paquet AC, Yang YJ, Madhusudhan MS, Gormley M, Feng KT, Bernlohr DA, McDonagh S, Pereira L, Sali A, Fisher SJ. Gene expression profiling of the human maternal-fetal interface reveals dramatic changes between midgestation and term. *Endocrinology*. 2007; 148:1059–1079. doi:en.2006-0683[pil]10.1210/en.2006-0683. [PubMed: 17170095]
- Winn VD, Gormley M, Paquet AC, Kjaer-Sorensen K, Kramer A, Rumer KK, Haimov-Kochman R, Yeh RF, Overgaard MT, Varki A, Oxvig C, Fisher SJ. Severe preeclampsia-related changes in gene expression at the maternal-fetal interface include sialic acid-binding immunoglobulin-like lectin-6 and pappalysin-2. *Endocrinology*. 2009; 150:452–462. [PubMed: 18818296]
- Witlin AG, Sibai BM. Hypertension in pregnancy: current concepts of preeclampsia. *Annu Rev Med*. 1997; 48:115–127. [PubMed: 9046950]
- Zhao H, Chen ZJ, Qin Y, Shi Y, Wang S, Choi Y, Simpson JL, Rajkovic A. Transcription factor FIGLA is mutated in patients with premature ovarian failure. *Am J Hum Genet*. 2008; 82:1342–1348. [PubMed: 18499083]

Fig. 1. Partial correlation networks of transcription levels in the control group. Only significant correlations (FDR p value ≤ 0.05) are included. *Solid lines* represent positive correlations and *dashed lines* represent negative correlations. The *arrows* in the figures indicate the suggested direction of the edges (nominal p value ≤ 0.05)

Fig. 2. Partial correlation networks of transcription levels in the PE case group. Only significant correlations (FDR p value ≤ 0.05) are included. *Solid lines* represent positive correlations and *dashed lines* represent negative correlations

Fig. 3.

Partial correlation networks of transcription levels in the control group. Only significant correlations (FDR p value ≤ 0.05) for which also the difference in PCCs between cases and controls ($PCC_{cases} - PCC_{controls}$) are significant non-zero (FDR p value ≤ 0.05) are included. *Solid lines* represent positive correlations and *dashed lines* represent negative correlations

Table 1

Descriptive characteristics of study population

	Mean (SD)		p value (37 cases, 58 controls)
	Cases (N = 37)	Controls (N = 58)	
Birth weight (g)	1,556 (769)	3,619 (469)	1.8E-26
Gestational Age (weeks)	31.89 (3.88)	38.7 (0.76)	2.8E-21
DBP (diastolic blood pressure)	96.4 (10.7)	70.1 (8.9)	4.5E-20
SBP (systolic blood pressure)	151.6 (16.3)	116.5 (10.5)	8.7E-20
Maternal weight (kg)	77.5 (19.7)	71.2 (12.5)	0.068
Infant sex (% male infants)	45	41	0.459
Parity (% primiparous)	55	65	0.261

Table 2

Partial correlations in the PE cases and controls

Gene name	Gene name	PE cases		Controls		Difference between PE cases and controls	
		PCC	p value	FDR p value	PCC	p value	FDR p value
SPCS2	SPCS2	0.133	7.58E-09	0.000	0.130	1.74E-05	0.016
TMEM140	APOL3	0.123	9.34E-08	0.001	0.125	3.80E-05	0.022
TRAPPC10	COQ10A	0.108	2.61E-06	0.021	0.014	6.45E-01	0.960
ATP6V1A	ZNF563	0.104	5.70E-06	0.032	0.021	4.84E-01	0.948
KLF5	SLC39A8	0.103	7.58E-06	0.037	0.023	4.43E-01	0.943
FIGLA	DBP	-0.054	1.81E-02	0.595	0.154	3.14E-07	0.002
TMEM130	MAP6	-0.050	2.87E-02	0.677	0.149	7.57E-07	0.002
GPIHBP1	C7orf26	-0.043	6.24E-02	0.813	0.139	4.39E-06	0.007
ILMN_I880210	VPS33B	0.030	1.96E-01	0.932	-0.150	6.96E-07	0.002
ILMN_I834017	PDZD4	0.033	1.47E-01	0.911	-0.145	1.53E-06	0.003
LOC731196	PDE12	-0.047	4.17E-02	0.743	0.129	1.93E-05	0.017
ILMN_I834017	UBFD1	-0.046	4.60E-02	0.762	0.128	2.27E-05	0.018
DLG2	DNAJB11	0.003	9.06E-01	0.984	-0.165	4.31E-08	0.001
FERMT1	PDZD4	0.042	6.96E-02	0.829	-0.125	3.57E-05	0.022
FERMT1	NPNT	0.033	1.49E-01	0.912	-0.127	2.60E-05	0.019
LOC731196	TAS2R1	-0.023	3.22E-01	0.957	0.136	6.22E-06	0.009
NPTX2	SCG2	-0.036	1.16E-01	0.890	0.118	1.02E-04	0.045
ILMN_I834017	TMEM130	-0.035	1.30E-01	0.900	0.117	1.07E-04	0.046
GNAI3	FERMT1	0.024	3.00E-01	0.954	-0.125	3.50E-05	0.022
HES1	LOC644596	0.023	3.18E-01	0.957	-0.122	5.65E-05	0.029
ILMN_I837017	C12orf59	0.012	5.97E-01	0.976	-0.131	1.55E-05	0.015
MYOC	SHC2	-0.014	5.50E-01	0.975	0.128	2.19E-05	0.017
ZNF154	CSTF2T	-0.010	6.58E-01	0.979	0.132	1.32E-05	0.014
CFH	NADSYN1	-0.019	4.09E-01	0.966	0.123	4.87E-05	0.026
SEMA5A	RUNX1T1	-0.009	6.99E-01	0.980	-0.150	7.11E-07	0.002
HIGD1B	SULT2B1	-0.013	5.77E-01	0.976	0.126	3.20E-05	0.020
SLN	ILMN_I837017	-0.016	4.98E-01	0.972	-0.151	5.79E-07	0.002

Gene name	Gene name	PE cases		Controls		Difference between PE cases and controls	
		PCC	p value	PCC	p value	p value	FDR p value
<i>IL6ST</i>	<i>ALDH8A1</i>	0.016	4.80E-01	0.150	7.12E-07	3.05E-05	0.014
<i>EGFL7</i>	<i>TAS2R1</i>	0.013	5.70E-01	-0.119	8.17E-05	3.55E-05	0.016
<i>TMEM130</i>	<i>UBFD1</i>	-0.011	6.39E-01	-0.142	2.48E-06	3.98E-05	0.017
<i>ILMN_1880210</i>	<i>MARCH2</i>	-0.006	8.03E-01	0.119	7.97E-05	9.28E-05	0.027
<i>SULT2B1</i>	<i>SLC39A10</i>	-0.003	9.08E-01	-0.127	2.61E-05	1.01E-04	0.028
<i>APOBEC3B</i>	<i>LOC731196</i>	-0.009	6.83E-01	-0.134	9.72E-06	1.03E-04	0.029
<i>NPTX2</i>	<i>MXRA5</i>	-0.006	7.92E-01	0.118	9.45E-05	1.04E-04	0.029
<i>C10orf10</i>	<i>CSDA</i>	-0.031	1.73E-01	-0.155	2.65E-07	1.09E-04	0.030
<i>FIGLA</i>	<i>TRAPPC10</i>	-0.023	3.27E-01	-0.146	1.40E-06	1.20E-04	0.031
<i>ZNF614</i>	<i>EARS2</i>	0.069	2.76E-03	0.159	1.31E-07	4.90E-03	0.218
<i>ILMN_1880210</i>	<i>HSPC047</i>	-0.049	3.51E-02	-0.151	5.94E-07	1.44E-03	0.122
<i>ALDH8A1</i>	<i>ZNF154</i>	-0.041	7.29E-02	-0.148	9.38E-07	8.68E-04	0.092
<i>CSDA</i>	<i>MARCH2</i>	0.020	3.94E-01	0.134	9.15E-06	3.54E-04	0.060
<i>SEMA5A</i>	<i>NADSYN1</i>	-0.055	1.67E-02	-0.134	9.26E-06	1.40E-02	0.351
<i>SEMA5B</i>	<i>NOTCH3</i>	0.051	2.56E-02	0.131	1.41E-05	1.29E-02	0.339
<i>TMEM130</i>	<i>FERMT1</i>	0.022	3.31E-01	0.130	1.79E-05	8.14E-04	0.089
<i>ST7OT1</i>	<i>NAALAD2</i>	0.032	1.65E-01	0.129	1.89E-05	2.41E-03	0.155
<i>KLF15</i>	<i>LOC644596</i>	0.021	3.53E-01	0.129	2.05E-05	8.07E-04	0.088
<i>PDE12</i>	<i>SHC2</i>	0.018	4.44E-01	0.128	2.20E-05	5.51E-04	0.073
<i>ATP6V1A</i>	<i>SNTB1</i>	0.075	1.05E-03	0.127	2.78E-05	1.11E-01	0.704
<i>SULT2B1</i>	<i>LNX2</i>	0.017	4.74E-01	0.127	2.83E-05	5.89E-04	0.074
<i>TMOD1</i>	<i>SLIT3</i>	-0.023	3.22E-01	-0.126	2.87E-05	1.22E-03	0.112
<i>ILMN_1834017</i>	<i>DAAM2</i>	0.056	1.44E-02	0.126	2.91E-05	2.92E-02	0.468
<i>NPTX2</i>	<i>CBLN1</i>	0.035	1.24E-01	0.124	3.93E-05	5.60E-03	0.235
<i>ILMN_1880210</i>	<i>DUSP18</i>	-0.031	1.78E-01	-0.123	4.59E-05	4.05E-03	0.200
<i>RARRES2</i>	<i>PLA2G2A</i>	0.044	5.74E-02	0.123	4.63E-05	1.33E-02	0.344
<i>KLF15</i>	<i>SRPRB</i>	-0.050	3.04E-02	-0.122	5.68E-05	2.52E-02	0.443
<i>ACBD3</i>	<i>GMFB</i>	0.038	1.03E-01	0.121	6.13E-05	9.15E-03	0.296
<i>TMEM97</i>	<i>SULT2B1</i>	0.013	5.86E-01	0.120	7.13E-05	7.85E-04	0.087

Gene name	Gene name	PE cases		Controls		Difference between PE cases and controls	
		PCC	p value	FDR p value	PCC	p value	FDR p value
PLCG1	APOL3	0.060	9.77E-03	0.508	0.119	8.80E-05	0.040
MARCH2	TMX1	0.017	4.72E-01	0.970	0.118	9.07E-05	0.041
ZNF563	ZNF154	-0.004	8.68E-01	0.984	-0.117	1.12E-04	0.047

The table shows the partial correlation coefficients (PCC) and corresponding *p* value and FDR *p* value for the PE cases and controls as well as the *p* value and FDR *p* value for the difference between the PCC for the case and control group. Numbers in bold indicate that the PCC or the difference between the PCCs is significant (FDR *p* value ≤ 0.05)