Microarray Screening for Novel Preeclampsia Biomarker Candidates

Olav Lapaire a Simon Grill b Sebastien Lalevee b Varaprasad Kolla b Irene Hösli a
Sinuhe Hahn b

 a Department of Obstetrics and Gynecology, and b Laboratory for Prenatal Medicine, Department of Biomedicine/Department of Obstetrics and Gynecology, University Hospital Basel, Basel, Switzerland

Introduction: Our aim was to identify novel biomarker candidates for the near-term prediction of preeclampsia in a homogenous collective. In this study, we screened at the genome-wide level for gene expression in placental villous tissue from patients with severe preeclampsia in comparison to normal healthy pregnancies. Material and Methods: Total RNA was extracted from placental villous tissue from 9 preeclamptic patients and 7 normotensive controls after scheduled cesarean sections. After sample pooling, gene expression analysis was performed using six Affymetrix Human Gene 1.0 ST arrays, followed by quantitative RT-PCR and validation of selected markers in the serum of patients at the protein level. Results: In total, 896 significantly differentially expressed genes were identified (p ≤ 0.05). After restricting these to molecules present in the circulation, 9 upregulated and 5 downregulated genes were selected. Four of them (β-hCG, HTRA4, LHB1, all upregulated; and NOX4, downregulated) were validated by quantitative real-time RT-PCR. Finally, the maternal plasma protein levels of 2 of these genes (LHB and β-hCG) were confirmed to be significantly different between preeclampsia cases and controls. Discussion: We identified 14 potential new biomarker candidates for preeclampsia and validated 4 of them by quantitative RT-PCR and 2 of them with subsequent serum protein analyses. Further studies will assess the optimal marker combination for the imminent prediction of impending preeclampsia.
high-risk groups and have failed to prove efficacious in the prevention of preeclampsia in apparently healthy, nulliparous patients [5–7]. Therefore, widely applicable and affordable tests are needed to permit early diagnosis at symptomatic stages in order to identify and closely monitor the patients at risk and thus provide the optimal prenatal care for these women and their fetuses.

A number of promising candidate biomarkers for the early detection of preeclampsia are currently under investigation, of which biomarkers of angiogenesis are currently at the most advanced stage of development [8].

A major concern is that cases of preeclampsia are misdiagnosed or unrecognized at admission, leading to high levels of maternal/fetal morbidity and mortality. Therefore, it would be helpful to have a set of markers which would identify such high-risk cases of impending preeclampsia [9].

The use of microarrays is a powerful strategy to investigate in detail possible alterations in gene expression associated with the pathophysiology of preeclampsia at the genome-wide level. Although this approach has been used previously for the possible detection of new novel biomarker candidates [10–20], most of these studies were conducted in nonhomogenous collectives, which limited their usefulness or the conclusions that can be drawn from them.

In the present study, we performed a comprehensive analysis of the gene expression in placental villous tissue samples using a homogenous collective case-control study. Care was taken to use matched normal controls and severe preeclampsia cases which were all delivered by CS. Quantitative RT-PCR and serum protein analyses were used to validate potential novel biomarker candidates.

**Material and Methods**

**Study Population**

This prospective case-control study was approved by the local ethical committee. After the patients had given written informed consent, 10 ml of EDTA blood and 10 ml of blood without anticoagulant were taken immediately prior to delivery. All patients underwent either elective CS (controls, n = 7) or scheduled CS due to severe preeclampsia (cases, n = 9). Severe preeclampsia was defined as blood pressure ≥160/100 mm Hg, measured at least 6 h apart, in combination with proteinuria ≥2+ (dipstick) at least twice in 24 h. Patient data are summarized in table 1.

**Sample Collection**

Maternal plasma (EDTA) and serum were collected prior to starting the CS. Blood samples were centrifuged (2,500 g, 15 min), and serum was aliquoted and stored at –80°C until further analysis. Placenta specimens were processed in the delivery unit immediately after CS to avoid RNA degradation. The maternal decidua and amniotic membrane were removed, and pieces of villous tissue (0.5 × 0.5 × 0.5 cm) from different lobes were excised. After extensive washing (5 times) in sterile RNase-free cold PBS to remove remaining maternal blood, the sample was stored in RNAlater (Ambion, Cat. No. AM7020) according to the manufacturer’s instructions. Following the addition of 1.5 ml of RNAlater, the tubes were left overnight at 4°C then stored at –80°C.

**RNA Isolation**

Upon removal of RNAlater, tissue was homogenized in 1 ml of Trizol Reagent (Invitrogen, Cat. No. 15596-026) using a Mixer Mill MM 301 (Retsch) in the presence of 4 stainless steel beads (1 bead of 5 mm in diameter and 3 beads of 3 mm in diameter). Isolated total RNA was further purified on RNA Clean & Concentrator-5 columns (Zymo Research, Cat. No. R1016). Total RNA preparations were analyzed using an Agilent 2100 Bioanalyzer. For the samples checked, the mean ± SD RNA integrity number value was 8.6 ± 0.3 (range 8.1–9.2). A NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, USA) was used to determine RNA and DNA concentrations.

**cDNA Target Synthesis and GeneChip Hybridization**

For cDNA target synthesis, RNA samples were pooled. The cDNA target was synthesized, fragmented and labeled with biotin using the Whole Transcript Target Labeling Assay and Control Reagents (Affymetrix, Santa Clara, Calif., USA, Cat. No. 900652) starting with 270 ng of total RNA, according to the procedure described in the GeneChip Whole Transcript Sense Target Labeling Assay Manual, version 4 (Affymetrix). After a first cycle of cDNA synthesis followed by in vitro transcription, approximately 32 ± 6 μg of cRNA was obtained from each reaction. Then, 10 μg of each sample was further used to generate single-strand cDNA, giving rise to approximately 6.8 ± 0.4 μg of cDNA. Subsequently, 5.5 μg of single-stranded cDNA was fragmented into 40–70 nucleotides and monitored with the Bioanalyzer 2100 using the RNA Nano 6000 Chip. All synthesis reactions were carried out in 0.5-ml tubes using a PCR machine (TPProfessional, Biotometra, Göttingen, Germany). The hybridization cocktail (80 μl) containing fragmented biotin-labeled target DNA at a final concentration of 25 ng/μl was transferred into Affymetrix GeneChip Human 1.0 ST Arrays (Affymetrix, Cat. No. 901085) and incubated at 45°C on a rotator in a hybridization oven 640 (Affymetrix) for 17 h at 60 rpm. The arrays were washed and stained on a Fluidics Station 450 (Affymetrix) using the Hybridization Wash and Stain Kit (Affymetrix, Cat. No. 900720) in accordance with the Fluidics Procedure FS450_0007. The GeneChips were processed with an Affymetrix GeneChip® Scanner 3000 7G (Affymetrix). DAT image files of the microarrays were generated using an Affymetrix GeneChip Command Console (version 0.0.0.676, Affymetrix). All experiments were performed in accordance with the Minimum Information About a Microarray Experiment guidelines [21]. Data analysis was performed using Partek (Partek Incorporated, St. Louis, Mo., USA) for normalization and detection of differentially expressed genes and ingenuity (Ingenuity Systems, Redwood City, Calif., USA) for the identification of biomarker candidates.
Statistical Analysis

Raw gene expression data were analyzed using Partek (Partek). All data were log-transformed prior to analysis.

Normalization was carried out using quantile normalization, which forces all the slides to have the same intensity distribution. Spots on the arrays with expression values missing for >25% of the samples in the preeclampsia group or >25% in the control group were filtered out. After normalization, differential gene expression was evaluated using analysis of variance (ANOVA). A p value cutoff of 0.05 was applied to retrieve significantly deregulated genes. To allow the identification of gene products that are known to be present in a specific body fluid, in our case serum and plasma, the list of deregulated genes was further filtered using the IPA biomarker tool (Ingenuity Systems). The emerging candidates were then validated by quantitative real-time RT-PCR.

For descriptive statistics, Student’s t test was used to compare normally distributed data between groups. The Mann-Whitney U Test was used in the case of ordinal data. Fisher’s exact test was used in the case of categorical data. Statistical significance was assigned when a p value was less than 0.05. All statistical analyses were performed using SPSS, version 12 (SPSS Inc., Chicago, Ill., USA).

cDNA Synthesis and Quantitative Real-Time RT-PCR

cDNA was synthesized from 180 ng of total RNA according to the manufacturer’s instructions using the Promega Reverse Transcription System (No. A3500). We were able to design oligonucleotides for 12 out of 14 potential biomarkers using the program Primer3 [22], which were then synthesized by Microsynth AG, Switzerland. PCR reactions were performed in a total volume of 25 μl on an ABI Prism 7000 detection system (Applied Biosystems) using the default cycling conditions. For each reaction, 1 μl of cDNA template, 12.5 μl of 2× ABIsYBR green master mix (No. 4367659) and 300 nM of each sense and antisense primer were used.

Serum Analyses

Serum concentrations of β-human chorionic gonadotropin (β-hCG) were analyzed on a Cobas® e601 using the appropriate kit (No. 03271749-190). Luteinizing hormone β polypeptide (LHB) was quantified using an LHB ELISA from Genway Biotech (No. 40-056-20501) according to the manufacturer’s instructions.

Results

We compared gene expression profiles of approximately 28,000 genes using 9 villous tissue samples from women with severe preeclampsia and 7 normotensive controls, pooled into 4 and 2 Affymetrix Human 1.0 ST arrays, respectively.

After normalization, ANOVA comparative analysis between pooled controls and pooled preeclampsia cases resulted in 896 significantly differentially expressed genes (p ≤ 0.05), belonging to various pathways including riboflavin metabolism, leukocyte extravasation signaling, NF-κB signaling and chemokine signaling.

To achieve the goal to detect potential biomarkers, we filtered these 896 genes with the IPA biomarker tool to

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Controls (n = 7)</th>
<th>Preeclamptic patients (n = 9)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age, years</td>
<td>35.1 (33/37)</td>
<td>36.8 (22/43)</td>
<td>NS</td>
</tr>
<tr>
<td>Gestational age, weeks</td>
<td>38.6 (38/41)</td>
<td>34.6 (33/39)</td>
<td>NS</td>
</tr>
<tr>
<td>Prepregnancy body mass index</td>
<td>21.1 (18.5/37.1)</td>
<td>26.3 (20.0/33.9)a</td>
<td>0.072</td>
</tr>
<tr>
<td>Body mass index at delivery</td>
<td>26.4 (22.1/41.3)</td>
<td>32.4 (22.8/36.3)</td>
<td>NS</td>
</tr>
<tr>
<td>Gain in body mass index</td>
<td>4.5 (3.6/7.2)</td>
<td>4.8 (2.4/9.9)a</td>
<td>NS</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>123 (105/144)</td>
<td>179 (150/220)</td>
<td>0.001</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>71.0 (65/83)</td>
<td>105 (90/140)</td>
<td>0.001</td>
</tr>
<tr>
<td>Proteinuria (dipstick)</td>
<td>negative</td>
<td>+++</td>
<td>0.000</td>
</tr>
<tr>
<td>Multipara</td>
<td>0/7</td>
<td>2/9</td>
<td>NS</td>
</tr>
<tr>
<td>Nullipara</td>
<td>2/7</td>
<td>7/9</td>
<td>0.072</td>
</tr>
<tr>
<td>Fetal sex: male/female</td>
<td>3/7</td>
<td>3/11</td>
<td>NS</td>
</tr>
<tr>
<td>Infant birth weight, g</td>
<td>3,140 (2,970/4,300)</td>
<td>1,810 (1,300/3,420)b</td>
<td>0.002</td>
</tr>
<tr>
<td>IUGR (&lt;3rd percentile)</td>
<td>0/7</td>
<td>4/11b</td>
<td>NS</td>
</tr>
<tr>
<td>Preterm delivery (&lt;37 weeks of gestation)</td>
<td>0/7</td>
<td>7/9</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Values are expressed as medians (minimum/maximum), where applicable. IUGR = Intrauterine growth restriction; NS = not significant. a Data for 8 of 9 patients with preeclampsia. b Data for 11 children born from 9 patients with preeclampsia.
### Table 2. Selected potential biomarkers for the detection of preeclampsia

<table>
<thead>
<tr>
<th>Assignment/abbreviation/name/location</th>
<th>Fold change (array)</th>
<th>Significant fold change (qRT-PCR)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 NM_000737/CGB/chorionic gonadotropin, ( \beta ) polypeptide/19q13.32</td>
<td>4.34†</td>
<td>8†</td>
<td>glycoprotein hormone, maintenance of pregnancy</td>
</tr>
<tr>
<td>2 NM_153692/HTRA4/HtrA serine peptidase 4/8p11.23</td>
<td>4.03†</td>
<td>8.3†</td>
<td>degradation of misfolded secretory proteins</td>
</tr>
<tr>
<td>3 NM_001882/CRHB/corticotropin-releasing hormone-binding protein/5q11.2</td>
<td>2.01†</td>
<td>NA</td>
<td>corticotropin-releasing hormone-binding protein inactivates corticotropin-releasing hormone, which may prevent inappropriate pituitary-adrenal stimulation in pregnancy</td>
</tr>
<tr>
<td>4 NM_000894/LHB/luteinizing hormone ( \beta ) polypeptide/19q13.32</td>
<td>1.71†</td>
<td>13†</td>
<td>promotion of spermatogenesis and ovulation</td>
</tr>
<tr>
<td>5 NM_012413/QPCT/glutaminyl-peptide cyclotransferase (glutaminyl cyclase)/2p22.2</td>
<td>1.63</td>
<td>NS</td>
<td>transfers pyroglutamyl residues to activate different neuroendocrine peptides</td>
</tr>
<tr>
<td>6 NM_078481/CD97/CD97 molecule/19p13</td>
<td>1.61</td>
<td>NS</td>
<td>7-span transmembrane protein involved in cell adhesion and signaling processes early after leukocyte activation</td>
</tr>
<tr>
<td>7 NM_002429/MMP19/matrix metalloproteinase 19/12q14</td>
<td>1.61†</td>
<td>NS</td>
<td>breakdown of extracellular matrix, embryonic development, reproduction, tissue remodeling</td>
</tr>
<tr>
<td>8 NM_001464/ADAM2/ADAM metalloproteinase domain 2 (fertilin ( \beta ))/8p11.2</td>
<td>1.39†</td>
<td>NS</td>
<td>membrane-anchored protein involved in cell-cell and cell-matrix interactions, fertilization, muscle development and neurogenesis as well as a subunit of an integral sperm membrane glycoprotein called fertilin, which plays an important role in sperm-egg interactions</td>
</tr>
<tr>
<td>9 NM_005538/INHBC/inhibin, ( \beta ) C/12q13.1</td>
<td>1.26†</td>
<td>NS</td>
<td>belongs to the transforming growth factor-( \beta ) hormone superfamily and is involved in hormone secretion and growth and differentiation in various cell types</td>
</tr>
<tr>
<td>10 NM_002983/CCL3/chemokine (C-C motif) ligand 3/17q11–q21</td>
<td>1.16 †</td>
<td>NA</td>
<td>involved in the acute inflammatory state in the recruitment and activation of polymorphonuclear leukocytes</td>
</tr>
<tr>
<td>11 NM_016931/NOX4/nicotinamide adenosine dinucleotide phosphate oxidase 4/11q14.2–q21</td>
<td>1.69 †</td>
<td>6.5 †</td>
<td>acts as oxygen sensor</td>
</tr>
<tr>
<td>12 NM_001078/VCAM1/vascular cell adhesion molecule 1/1p32–p31</td>
<td>1.78</td>
<td>NS</td>
<td>mediates leukocyte-endothelial cell adhesion and signal transduction</td>
</tr>
<tr>
<td>13 NM_006732/FOSB/FBJ murine osteosarcoma viral oncogene homolog B/19q13.3</td>
<td>1.92</td>
<td>NS</td>
<td>leucine zipper protein, involved in the formation of activator protein-1 transcription complexes to induce proliferation, differentiation and transformation</td>
</tr>
<tr>
<td>14 NM_001337/CX3CR1/chemokine (C-X3-C motif) receptor 1/3p21.3</td>
<td>1.96</td>
<td>NS</td>
<td>involved in T cell homing and recruitment of immune cell populations</td>
</tr>
</tbody>
</table>

Quantitative RT-PCR (qRT-PCR) validations of 14 biomarker candidates. Threshold cycle values were normalized to GAPDH. Statistical significances were calculated using ANOVA. NA = Not applicable; NS = not significant.
further select for those candidates which could potentially be present in the circulation (serum or plasma). By this selection process, we obtained 14 potential biomarker candidates, of which 9 were upregulated and 5 were downregulated, as follows: β-hCG (4.34-fold), HTRA4 (4.03-fold), CRHB P (2.01-fold), LHB (1.71-fold), QPCT (1.63-fold), CD97 (1.61-fold), MMP19 (1.61-fold), ADAM2 (1.39-fold), INHBC (1.26-fold), CCL3 (1.16-fold), NOX4 (1.69-fold), VCAM1 (1.78-fold), FOSB (1.92-fold) and CX3CR1 (1.96-fold). The results of this screen are summarized in Table 2.

To validate these microarray results, real-time quantitative RT-PCR was performed, in which the expression levels of 12 of the 14 candidate genes (Table 2) could be successfully quantified. Two of the 14 (CCL3 and CRHB P) could not be evaluated. For these analyses, cDNA was synthesized from the same RNA preparations that were used for microarray examinations. Results were normalized using the housekeeping gene GAPDH (encoding glyceraldehyde 3-phosphate dehydrogenase), and fold changes were calculated using the ΔΔ threshold cycle method [23]. The results of those validations are shown in Table 2 as ratios of the threshold cycle values of the gene of interest and the threshold cycle value of GAPDH, respectively.

Among the 12 tested genes, we observed a significant concordance (p ≤ 0.05) with the results obtained by microarray analysis for 4 genes, namely β-hCG (upregulated), HTRA4 (upregulated), LHB (upregulated) and NOX4 (downregulated).

To investigate the eligibility of the validated candidates as biomarkers which could be present in the maternal circulation, two proteins, namely β-hCG and LHB, were subsequently examined in the maternal serum samples that were collected at the time of delivery. Figure 1 shows a box plot of the maternal serum β-hCG and LHB levels. The mean ± SD concentration of β-hCG in serum samples from women with preeclampsia was significantly higher than in the controls (68,900 ± 51,980 vs. 18,100 ± 15,330 IU/l; p = 0.03). The median concentration of LHB in preeclamptic serum samples was significantly higher than in the controls (0.739 ± 0.107 vs. 0.631 ± 0.022 IU/l; p = 0.05).

**Discussion**

The present study represents a comprehensive analysis of a microarray RNA expression profile of placental vil-lous tissue processed immediately after scheduled CS from patients with severe preeclampsia and healthy controls, with the use of RNA microarray technology and further validation with quantitative real-time RT-PCR and serum protein level determination. Nine upregulated and 5 downregulated genes were selected from 896 significantly differentially expressed genes, on the basis of possible presence in maternal serum, a vital requisite for putative biomarkers. In 12 out of these 14 instances, quantitative real-time RT-PCR was successfully per-
formed to confirm expression levels determined by microarray analysis, showing significant differences in 4 potential markers (β-hCG, HTRA4 and LHB, all upregulated; NOX4, downregulated). Furthermore, to verify the detectability of the gene products in maternal blood, we measured the protein concentration of 2 of these 4 candidates in serum samples that were collected together with the tissue samples prior to CS.

Similar to other studies using a microarray RNA expression profile of placental villous tissue of preeclamptic patients, we found differentially expressed genes such as β-hCG [24] or LHB, which could be the result of altered trophoblast physiology [25]. The elevated levels of these molecules, which are predominantly expressed in the syncytiotrophoblast, may be indicative of altered pathways that are related to uteroplacental vascular insufficiency or placental oxidative stress and reoxygenation [26].

Amongst the genes whose expression was significantly upregulated was HTRA4, a serine peptidase responsible for the hydrolysis of misfolded proteins. The mRNA level of this gene was elevated approximately 4-fold in placentae from women with preeclampsia when compared to normal controls. This is in accordance with a recently published study which exhibited similar results for HTRA1, a member of the family of HTRA proteins, which is a secreted multidomain protein with serine protease activity and which is involved in the physiological development of many organs as well as in the development of several pathologies [27, 28].

Amongst the genes whose expression was downregulated, it was interesting that we observed a significant decrease in the expression of NOX4, as this gene product is involved in blood pressure regulation and endothelial cell survival, vital features in the pathophysiology of preeclampsia [29]. This decrease in expression may be explained by preexisting high levels of superoxide in the preeclamptic placenta [30], due to increased production of superoxide within the placenta by abnormal mitochondria, increased activity of xanthine oxidase [31] and endothelial nitric oxide synthase [32].

Amongst the candidate biomarkers for which we could determine mRNA as well as protein expression levels was β-hCG. For this protein, we detected increased concentrations in the maternal circulation in patients with preeclampsia in comparison to controls. This could serve to indicate that the increased β-hCG concentrations may be an effect of the disease and not caused by gestational age differences of the subjects in our study. Similarly, serum levels of LHB were significantly elevated in the maternal circulation, which may be the result of altered trophoblast physiology.

The strength of this study was the homogenous case-control collective, in which care was taken to match patients with diagnosed severe preeclampsia and healthy normal controls. A further important feature was that the deliveries in both groups were performed by CS, thereby eliminating an important confounder between the study and control groups [33]. This permitted us to determine an altered gene expression profile in cases of preeclampsia when compared to matched healthy controls. By this means, we were able to detect 14 genes encoding putative secreted biomarkers present in maternal serum.

Our use of placentae immediately after delivery may assist with the development of biomarkers to detect pregnancies at risk for imminent preeclampsia. This is a major clinical concern, as many such cases are inadequately identified at admission [9].

Further studies are ongoing to investigate whether the expression of these genes is altered during gestation in those cases that subsequently develop preeclampsia and thereby to determine the predictive value of these candidate biomarker molecules.

Acknowledgements

We are grateful to all the pregnant women who participated in this study. We thank Philippe Demougin and Andreas Papasotiropoulos from the Affymetix Core Lab/Life Science Training Facility at the Biocenter in Basel, Switzerland, for technical support. Furthermore, we would like to thank Corinne Rusterholz for her contribution.

This study was funded by the EU Pregenesys Program (grant number 037244) and SNF (grant number 310030-129654).

References


Lapaire/Grill/Lalevee/Kolla/Hösli/Hahn
Screening for Novel Preeclampsia Biomarkers


