Placenta 34 (2013) 599–605

Contents lists available at SciVerse ScienceDirect

Placenta

journal homepage: www.elsevier.com/locate/placenta

Differentially expressed microRNAs and affected biological pathways revealed by modulated modularity clustering (MMC) analysis of human preeclamptic and IUGR placentas

L. Guo [a,b,c], S.Q. Tsai [e,f], N.E. Hardison [d], A.H. James [h], A.A. Motsinger-Reif [b,d], B. Thames [g], E.A. Stone [d], C. Deng [c], J.A. Piedrahita [a,b,*]

[a] Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA
[b] Center for Comparative Medicine and Translational Research, North Carolina State University, Raleigh, NC, USA
[c] Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction, Ministry of Education, Huazhong Agricultural University, Wuhan, PR China
[d] Department of Statistics, North Carolina State University, Raleigh, NC, USA
[e] Molecular Pathology Unit, Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, MA 02129, USA
[f] Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA
[g] Department of Obstetrics and Gynecology, Duke University School of Medicine, Durham, NC, USA
[h] Department of Obstetrics and Gynecology, University of Virginia Health Center, Charlottesville, VA 22908, USA

ARTICLE INFO

Article history:
Accepted 5 April 2013

Keywords:
Placenta
Preeclampsia
IUGR
MiR-194
MiR-149

ABSTRACT

Introduction: This study focuses on the implementation of modulated modularity clustering (MMC) a new cluster algorithm for the identification of molecular signatures of preeclampsia and intrauterine growth restriction (IUGR), and the identification of affected microRNAs
Methods: Eighty-six human placentas from normal (40), growth-restricted (27), and preeclamptic (19) term pregnancies were profiled using Illumina Human-6 Beadarrays. MMC was utilized to generate modules based on similarities in placental transcriptome. Gene Set Enrichment Analysis (GSEA) was used to predict affected microRNAs. Expression levels of these candidate microRNAs were investigated in seventy-one human term placentas as follows: control (29); IUGR (26); and preeclampsia (16).
Results: MMC identified two modules, one representing IUGR placentas and one representing preeclamptic placentas. 326 differentially expressed genes in the module representing IUGR and 889 differentially expressed genes in a module representing preeclampsia were identified. Functional analysis of molecular signatures associated with IUGR identified P13K/AKT, mTOR, p70S6K, apoptosis and IGF-1 signaling as being affected. Analysis of variance of GSEA-predicted microRNAs indicated that miR-194 was significantly down-regulated both in preeclampsia ($p = 0.0001$) and IUGR ($p = 0.0304$), and miR-149 was significantly down-regulated in preeclampsia ($p = 0.0168$).
Discussion: Implementation of MMC, allowed identification of genes disregulated in IUGR and preeclampsia. The reliability of MMC was validated by comparing to previous linear modeling analysis of preeclamptic placentas.
Conclusion: MMC allowed the elucidation of a molecular signature associated with preeclampsia and a subset of IUGR samples. This allowed the identification of genes, pathways, and microRNAs affected in these diseases.

1. Introduction

Preeclampsia (PE) is a pregnancy associated syndrome characterized by hypertension and proteinuria during pregnancy, which is a consequence of diverse pathophysiological processes involving impaired implantation, endothelial dysfunction, and systemic inflammation [1–4]. Intrauterine growth restriction (IUGR), of diverse causes, refers to the poor growth of a fetus that has not

Abbreviations: IUGR, intrauterine growth restriction; PE, preeclampsia; GSEA, Gene Set Enrichment Analysis; MMC, modulated modularity clustering.

* Corresponding author. Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, 4700 Hillsborough St., Raleigh, NC, USA. Tel.: +1 919 515 7407.
E-mail address: jorge_piedrahita@ncsu.edu (J.A. Piedrahita).

0143-4004/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.placenta.2013.04.007

reached its growth potential while in the mother's uterus during pregnancy [5]. This study describes the genome-wide gene expression analysis of a large ($n = 86$) set of human placentas in order to uncover expression patterns (or molecular signatures) associated with preeclampsia and IUGR. Previously, we successfully identified, using a linear model analysis, genes disregulated in preeclampsia [6]. However, a similar analysis of IUGR samples was less effective, likely due to the high heterogeneity of the IUGR samples. As a result, we used an alternative method of analysis referred to as modulated modularity clustering (MMC) [7] that identifies unique expression signatures in a heterogeneous sample population. MMC is analogous to k-means clustering [8] with the exception that the numbers of clusters or modules are independently identified by MMC, not arbitrarily selected by the investigator. Using MMC we were able to identify unique placental gene expression signatures for both preeclampsia and a subset of IUGR subjects and utilized those expression profiles to identify, using Gene Set Enrichment Analysis (GSEA) [9], microRNA candidates disregulated in IUGR and/or preeclampsia. In the case of PE we compared the MMC-generated results with our previously published linear model analysis of PE placentas [6].

2. Materials and methods

2.1. Study design

An initial study population, consisting of 86 Caucasian and African–American subjects collected during 2004–2008, was utilized for gene expression analysis. Initially a small subset of 14 subjects (batch #1) was analyzed using Illumina arrays to ensure the quality of the RNA before proceeding further. Once quality was confirmed the remaining 72 samples were analyzed (batch #2). Principle component analysis identified one sample as an outlier and that sample was discarded from the analysis. For the linear modeling batch #1 and batch #2 samples were included and the data was corrected for batch effect using a set of 8 technical replicates. For the MMC analysis only the second batch was utilized as a preliminary analysis indicated that batch effect negatively affected the performance of the MMC analysis. Thus, a subset population of 71 subjects in the following groups: (1) preeclampsia ($n = 16$); (2) IUGR ($n = 26$); and (3) control group ($n = 29$), were used for MMC and microRNA qRT-PCR validation.

2.2. Subjects and sample collection

Preeclampsia was diagnosed when both pregnancy-induced hypertension and proteinuria were present according to American College of Obstetricians and Gynecologists 2000 guidelines [10]. Pregnancy-induced hypertension was defined as a sustained (≥2 measures 6 h apart) blood pressure elevation (>140/90 mmHg) >20 weeks of gestation. Proteinuria was defined as a sustained (≥2 measures 4 h apart) presence of elevated protein in the urine (>30 mg/dL or >1+ on a urine dipstick). IUGR was defined when the estimated weight of a fetus was below the 10th percentile for its gestational age and whose abdominal circumference was below the 2.5th percentile. Subjects were enrolled at the Duke University Medical Center Obstetric Clinic starting August 1, 2003. The criteria for subject enrollment and procedure for sample collection and storage were described in our previous paper [6]. Summary characteristics of the studied population are presented in Table 1. The study was approved by the Duke University Medical Center Institutional Review Board (IRB 00016065).

2.3. RNA isolation from human placenta

Total RNA was isolated from term human placentas using the Totally RNA kit (Ambion). Small RNA used for microRNA qPCR validation, was isolated using the mirVana miRNA isolation kit (Ambion). Only samples with an OD_{260}:$OD_{280} \geq 2.0$ were used.

2.4. Real-time quantitative RT-PCR

Small RNA-containing total RNA was converted into cDNA using the miScript Reverse Transcription Kit (Qiagen). EvaGreen [11] based qRT-PCR was performed to profile miRNA levels in 71 placentas from healthy or case (PE or IUGR) complicated pregnancies. The fold change between the experimental sample and the calibration sample was calculated using the Pfaffl method [12] (See Supplemental method for detailed information).

2.5. Statistical analysis

Transcript data was log 2 transformed, and quantile normalized as described previously [6]. Principle components analysis [13] was performed to calculate the contribution of each of the factors to the measured transcriptional variation: classification (or module), gender, induction of labor, their pair-wise two-way interactions, and estimated gestational age, by using JMP Genomics 5.0 (SAS Institute, Cary, NC). Since no significant effect of induction of labor was detected in our previous study [6], we chose to use a model without induction of labor to perform gene-specific analysis of variance (ANOVA) using PROC MIXED in SAS (SAS Institute, Cary, NC): expression $= \mu +$ classification $+$ gender $+$ gender $\times$ classification $+$ batch $+ \varepsilon$, treating classification, gender, and batch as fixed effects. Custom hypothesis tests were constructed to test for differential expression between case (PE or IUGR) and control or between different modules predicted by MMC (with module in place of classification). Raw p-values were corrected for multiple comparisons via Benjamini–Hochberg FDR at $\alpha < 0.05$ (for pathway analysis) and Bonferroni at $\alpha < 0.05$ methods [14] as implemented in PROC MULTTEST in SAS (SAS Institute, Cary NC).

MiRNA-specific analysis of variance (ANOVA) was performed using PROC MIXED in SAS 9.2 (SAS Institute, Cary, NC), treating classification as a fixed effect. The difference of miRNA levels in three groups was also profiled by fitting relative expression unit to classification in JMP 9 (SAS Institute, Cary, NC). Student t test was performed to evaluate the module effect by using SAS 9.2 (SAS Institute, Cary, NC). Differences in demographics between the control and case groups were tested using one-way ANOVA and Chi-square test for continuous and discrete variables respectively, by using JMP 9 (SAS Institute, Cary, NC). Validation of microarray results using RT-PCR was previously reported [6].

2.6. Modulated modularity clustering and Gene Set Enrichment Analysis

Modulated modularity clustering (MMC) [7] was used to separate placentas on the basis of the gene expression profiles. This method compares the gene expression profiles of input samples, correlates them to each other, and creates modules based on an overall correlation index, with no information provided as to which samples were classified as control, IUGR or preeclampsia. The program does not know how many modules exist *a priori*, but determines how many different molecular signatures are there in the sample population. A normalized and quality filtered transcript data for 34,471 probes in 71 human placenta samples was submitted to the online software for clustering (http://mmc.gnets.ncsu.edu/). Validation of the MMC procedure was provided by comparison of previous linear model analysis of preeclamptic samples [6] with results obtained from MMC clustering.

Gene Set Enrichment Analysis (GSEA) [9] was performed on the ranked list (based on p-value) of differentially expressed genes to identify functionally enriched gene sets. Curated microRNA targets from the Molecular Signature Database (MSigDB, Borad Institute) were analyzed for enrichment.

Table 1
Demographic information of study population.

Variables[a]	Group			p-Value[b]		
	Control ($n = 40$)	PE ($n = 19$)	IUGR ($n = 27$)	Control vs PE	Control vs IUGR	PE vs IUGR
Fetal estimated gestational age (weeks)	37.7 ± 1.9	33.6 ± 3.7	35.0 ± 3.9	<0.0001	0.0006	0.1365
Female %	42.9	57.9	59.3	0.2761	0.1835	0.9263
Placental weight (g)	491.1 ± 141.3	448.2 ± 237.8	329.8 ± 114.7	0.3424	<0.0001	0.0166
Birth weight (g)	3263.7 ± 625.8	2278.4 ± 998.9	1936.8 ± 676.6	<0.0001	<0.0001	0.1244
Corrected percentile birth weight	49.6 ± 3.8	40.4 ± 5.7	6.7 ± 4.8	0.1831	<0.0001	<0.0001
Maternal parity	2.4 ± 1.7	0.4 ± 0.7	1.1 ± 1.6	<0.0001	0.0008	0.1464
Maternal weight (lb)	215.9 ± 57.6	209.2 ± 45.5	153.1 ± 35.0	0.6207	<0.0001	0.0003
Induced %	16.7	73.7	48.2	<0.0001	0.0049	0.0833

[a] Continuous data are presented as mean ± SD; categorical data as a percentage.
[b] Obtained by student t test and Chi-square test for continuous and categorical variables, respectively (JMP 9, SAS Institute, Cary, NC).