Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** Magee-Womens Research Institute and Foundation

2. **Reporting Period (start and end date of grant award period):** 01/01/2011 – 06/30/12

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Cheryl A. Richards, MBA

4. **Grant Contact Person's Telephone Number:** 412-641-8932

5. **Grant SAP Number:** 4100054859

6. **Project Number and Title of Research Project:** 1 **-** Analysis of small RNAs in the fetal placental maternal interface

7. **Start and End Date of Research Project:** 01/01/2011 – 06/30/2012

8. **Name of Principal Investigator for the Research Project:** Tianjiao Chu**,** PhD

9. **Research Project Expenses**.

   9(A) Please provide the amount of health research grant funds spent on this project for the entire duration of the grant, including any interest earned that was spent:

   $ 499,545

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project	Cost
Chu	Principal Investigator	35%	$32,380.01
Sadovsky	Co-investigator	10%	$15,976.10
Zhu	Bioinformatician	50%	$14,546.83
Goldin	Programmer	100%	$18,731.63
Hilko	Research Assistant	100%	$17,916.71
Yeniterzi	Senior Programmer	50%	$40,155.37

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project
None		

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

Type of Scientific Equipment	Value Derived	Cost
None		

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes_____ No___X_____

If yes, please indicate the source and amount of other funds:

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes __X_____ No_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If

you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

A. Title of research project on grant application	B. Funding agency (check those that apply)	C. Month and Year Submitted	D. Amount of funds requested:	E. Amount of funds to be awarded:
Small RNAs at the Placental Maternal Communication Interface	☒NIH ☐ Other federal (specify:_____) ☐ Nonfederal source (specify:_)	April 2011	$445,274	$445,274

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes_____ No___X_____

If yes, please describe your plans:

12. Future of Research Project. What are the future plans for this research project?

We have applied for and received a NIH R21 grant which was based on the preliminary findings of this project. With this new grant, we will continue our study of the trafficking and mechanism of miRNA transport in the maternal/placental/fetal interface, which will help us better understand the epigenomic pathways involved in human embryonic development.

13. New Investigator Training and Development. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes_____ No____X_____

If yes, how many students? Please specify in the tables below:

	Undergraduate	Masters	Pre-doc	Post-doc
Male				
Female				
Unknown				
Total				

	Undergraduate	Masters	Pre-doc	Post-doc
Hispanic				
Non-Hispanic				
Unknown				
Total				

	Undergraduate	Masters	Pre-doc	Post-doc
White				
Black				
Asian				
Other				
Unknown				
Total				

14. Recruitment of Out-of–State Researchers. Did you bring researchers into Pennsylvania to carry out this research project?

Yes____X____ No_____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

Suveyda Yeniterzi, MS, University of Southern California

15. Impact on Research Capacity and Quality. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes___X_____ No_____

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research.

This project allowed us to recruit two bioinformaticians, both with unique expertise in sequencing and microarray data analysis. The recruitment markedly enhanced the project, bolstered bioinformatics capacity at MWRIF, provided the much needed cutting edge computational support to the investigators in this and other projects, and enabled our investigators to take full advantage of cutting edge technologies, including next generation sequencing, in this and other research projects.

16. Collaboration, business and community involvement.

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes___X_____ No_____

If yes, please describe the collaborations:

Discussions about new ways for data analysis were initiated with the Institute for Systems Biology in Seattle. Dr. Sadovsky met and discussed with their directors on July 16, 2012.

16(B) Did the research project result in commercial development of any research products?

Yes_____ No__X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No___X_____

If yes, please describe involvement with community groups that resulted from the research project:

17. Progress in Achieving Research Goals, Objectives and Aims.

List the project goals, objectives and specific aims (as contained in the grant application's strategic plan). Summarize the progress made in achieving these goals, objectives and aims for the period that the project was funded (i.e., from project start date through end date). Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project

work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.

Introduction

This project centers on miRNA-based communication pathways between the human placenta and the fetal/maternal blood. MiRNAs regulating the expression of mRNAs and proteins are in distinct spatial patterns. In this project we posited that exquisite expression patterns of miRNAs among the maternal, placental and fetal compartments are germane for intact embryonic development, and sought to develop tools to interrogate the transport of miRNAs across these intrauterine compartments. Our initial goal was to profile the expression of miRNAs in the placenta and interfacing maternal and fetal blood during human pregnancy. Once we obtained the expression data of all miRNA species in the triad of human maternal blood, placenta, and fetal blood, we sought to develop statistical analyses to interrogate dynamics in miRNA trafficking among maternal blood, placenta, and fetal blood.

We initially chose to use the small RNA sequencing approach, because it is quantitative, sensitive to less abundant species, and exhibits low noise. It also does not require prior sequence knowledge, and allow the discovery of novel RNA species. We assumed that the dynamic range of modern high-throughput sequencing methodologies would make it possible to define even rare sequences. This might have been particularly germane for analysis of fetal miRNAs. Two reasons led us to slightly deviate from this approach: first, obtaining reliable libraries from the limiting amount of maternal and fetal plasma became a major hurdle, as described below, and second, most abundant species, which formed the basis or our computational analysis, were known sequences, rendering the goal of identifying unique sequences relatively minor.

In our original proposal we described our goal to deep-sequence all miRNAs from the placenta, maternal pre-delivery blood and fetal cord blood from 20 pregnant women, for a total of 60 libraries. For reasons described within our report we expanded our analyses to include maternal post-partum samples. Because of technical difficulties in obtaining reliable sequences we included a large number of microarray analyses of the relevant compartments, all described below. Overall, we collected specimens and extracted RNA from more than 100 participants. The RNA was extracted from placental biopsies, maternal blood before delivery (early in labor), maternal blood within 24 h after delivery, and fetal cord blood at delivery. Because of a number of technical issues (patient's refusal of an additional blood test, sample not collected and alike), post partum maternal blood was not collected from a fraction of the participants, resulting in 83

incomplete sets (all RNA samples except for post partum collection). There are also sets we will have to exclude because one or more of plasma samples was hemolized (which happened in 24 of the "complete" sets).

Our main challenge was the low level of miRNA in the plasma samples. We have examined several RNA extraction kits, designed for miRNA isolation from plasma. These kits included Norgen Biotech Plasma/Serum Circulating RNA Purification Kit (cat # 3000), Qiagen miRNeasy Mini Kit (cat #217004), Qiagen QIAamp Circulating Nucleic Acid Kit (cat # 55114), Macherey-Nagel NucleoSpin miRNA Plasma (cat #740981.250), Ambion mirVana miRNA Isolation Kit (cat #AM1560) and modified Triazol LS extraction with purification by Qiagen minicolumn. After extensive testing, we selected the Macherey-Nagel NucleoSpin miRNA Plasma kit for our process. Notably, the level of short RNA in our samples was prohibitive for using standard techniques (such as measurements by Nanodrop and Agilent Bioanalyzer) for estimation of RNA quantity and quality. These technical hurdles became a major challenge, resulting in inadequate miRNA libraries from our initial preparations, performed by a commercial laboratory in Durham, NC, Expression Analysis. After repeated failures to obtain reliable libraries, we contacted a different commercial laboratory, Ocean Ridge Bioscience, which uses a custom approach to library preparation. As described below, Ocean Ridge successfully constructed and sequenced several libraries, which provided useful information on miRNA expression levels at the desired compartments.

Sequencing data analysis

Our miRNA sequencing data consist of 20 libraries, constructed from placental tissue, pre-delivery maternal plasma, cord blood plasma, and post-delivery plasma samples from five patients, generated by Ocean Ridge Bioscience. We also obtained another 19 miRNA sequencing libraries generated by Expression Analysis, with 12 of these19 libraries consisting of the similar types of samples from 3 patients. This resulted in a total of 20 libraries. The fastq files for all miRNA sequencing libraries were processed as described in our first Annual Report, and detailed here. The processed miRNA sequencing libraries were then normalized by the library size to generate the reads per million (RPM) data. We noticed that the quality of the plasma libraries generated by Expression Analysis was not satisfactory. In particular, the size of some the plasma libraries is less than 30,000 reads, with 13,000 reads aligned to miRBase 18 miRNA database. Most of the placenta specific miRNAs that are known to be present in maternal plasma during pregnancy, including the C19MC miRNAs, are either absent or having only 1 or 2 copies in these libraries. Given that these miRNAs are among the most likely candidates transported between the fetal/maternal interface, those small libraries are not suitable for our study. Therefore, we decided to focus our analysis on the 19 out of the 20 miRNA libraries by Ocean Ridge (another library from Ocean Ridge, with only 10351 reads, was too small for our study). The information about these libraries is provided in Table 1. Using hierarchical clustering, we found that the placental samples form a tight group, while the plasma samples form another tight group (Figure 1). Within the plasma group, the maternal plasma, and the cord blood libraries were clustered separately.

Our original aim plan was to test the correlation of the expression levels for each miRNA between placental samples and pre-delivery plasma samples, and between placenta and cord

blood plasma samples. However, with data for only two patients, this approach was not feasible, because the correlation of any two two-dimensional vectors is either 1 or not defined. Instead, we used a modified model. Recall that our original model for the transport of miRNA between two compartments A and B assumed that:

Equation 1

$$X_{B,i,j} = k_{A,B,i}\, U_{A,i,j} + V_{B,i,j} + \varepsilon_{B,i,j}$$

where $X_{B,i,j}$ is the measured expression level of miRNA i in compartment B from the individual j, $U_{A,i,j}$ the amount of miRNA i in compartment A, $V_{B,i,j}$ the amount of miRNA i in compartment B not coming from compartment A, $k_{A,B,i} \geq 0$ the ratio of miRNA i in compartment B originating from compartment A to the total amount of miRNA i in compartment A, and called the rate of transport of i from A to B, and $\varepsilon_{B,i,j}$ is the measurement error for the expression levels of miRNA i in compartment B from the individual j. We note that for highly expressed miRNAs, the coefficient of variation of its measured expression level in a sequencing library is negligible. Therefore, for a miRNA that are highly expressed in compartment A, its expression level $U_{A,i,j}$ could be approximated by its measurement $X_{A,i,j}$. Equation 1 then could be simplified to:

Equation 2

$$X_{B,i,j} = k_{A,B,i}\, X_{A,i,j} + V_{B,i,j} + \varepsilon_{B,i,j}$$

If miRNA could be transported to compartment B from both compartment A and compartment C, then our model would be:

Equation 3

$$X_{B,i,j} = k_{A,B,i}\, X_{A,i,j} + k_{C,B,i}\, X_{C,i,j} + V_{B,i,j} + \varepsilon_{B,i,j}$$

In both models above, we would conclude that a miRNA is transported from compartment A (and C) to compartment B if the parameters $k_{A,B,i}$ (and $k_{C,B,i}$) are positive. With data from only a few patients, when we perform the regression using the model above, it is likely that the test for the coefficient $k_{A,B,i}$ (and $k_{C,B,i}$) will return a non-significant p value, which makes it difficult to evaluate whether the miRNA is transported from A (and C) to B. To address this problem, we decided to take the model selection approach, where we select the best sub-model of the model specified in Equation 2 and Equation 3 using the Akaike Information Criterion (Akaike 1974). Furthermore, we initially focused on a set of miRNAs for which the value of $V_{B,i,j}$, in Equation 2 is close to 0. As shown in our proposal, the smaller the variance of $V_{B,i,j}$, the higher the correlation between $X_{B,i,j}$ and $X_{A,i,j}$, hence the more likely the test would be significant. Therefore, for the study of miRNA transport from the placenta to the maternal/cord blood, we chose miRNAs expressed natively and abundantly in placenta, but not present in maternal/cord blood in situations when those compartments are not in contact with the placenta. Similarly, to study the transport of miRNAs from maternal/cord blood to placenta, we chose miRNAs that are present and abundant in maternal/cord blood in situations when those compartments are not in contact with the placenta.

Based on the above considerations, we examined the expression profile of the miRNAs in non-

pregnant adult plasma and cultured Primary human trophoblasts (PHT cells). In the first 6 months of this project, we already produced two miRNA libraries from cultured PHT cells under either normal or hypoxia conditions. We also imported from *Chen and et al. 2008* a miRNA sequencing library from pooled normal male and non-pregnant female plasma samples, and a library from pooled lung cancer patient sera. These data were annotated by miRBase 10.0 database, which was based on NCBI36 reference human genome. We mapped the miRBase 10.0 database and miRBase 18.0 database, and annotated the data from *Chen et al. (2008)* with miRBase 18.0 database. The similarity of these libraries with the 20 libraries we generated for this study can be viewed in the multidimensional scaling plot in

Figure 2.

We normalized the two PHT cells libraries and the serum library by library sizes, and obtained the reads per million (RPM) values for each miRNA. For the study of miRNA transport from the placenta to either the maternal blood or cord blood, we selected those miRNAs abundant in PHT cells and the placenta (RPM >= 1500) but nearly absent from the non-pregnant normal serum and the cancer serum (RPM <= 30). To ensure that these miRNAs are present in the maternal plasma, we further required that these miRNAs have moderate expression in pre-delivery maternal plasma (average RMP >= 100). For the study of miRNAs transport from either the maternal blood or fetal blood to the placenta, we selected those miRNAs that were abundant in non-pregnant normal serum (RPM >= 1500) but almost absent from PHT cells (RPM <= 50). In total we identify nine miRNAs for the study of miRNA transport from placenta to plasma, and five for the study of miRNA transport from plasma to placenta, as shown in Table 2. For convenience, in this report, we called the first 9 miRNAs the "placental miRNAs", and the other 5 miRNAs the "plasma miRNAs".

For each miRNA in Table 2, we further examined the Pearson's correlation between placenta and pre-delivery maternal plasma miRNAs, and the correlation between placenta and cord blood miRNAs, shown in Table 3. Note that because of the small sample size (4 to 5), none of the tests was significant. We then applied one-sided Wilcoxon signed rank test to the correlations, and found that for the nine placental miRNAs (candidates for placenta to maternal/fetal blood transport), both the correlation between placenta and pre-delivery maternal plasma and the correlation between placenta and cord blood were significantly above 0 (p values = 0.004 and 0.02 respectively). However, for the five plasma miRNAs (candidates for maternal/fetal blood to placenta transport), only the correlation between placenta and cord blood was significantly above 0 (p value = 0.03).

We compared the level of the nine placenta miRNAs in the pre-delivery maternal plasma to their level in the post-delivery maternal plasma, and tested the log2 ratio of miRNA expression before and after delivery using two-sided t test. As shown in Table 4, the expression levels for eight of these nine miRNAs were significantly decreased after delivery, which is consistent with the hypothesis that these nine miRNAs were transported from the placenta to maternal blood.

Applying the AIC score based model selection procedure to the sequencing data of the selected miRNA to compare the sets of all unsaturated models, we identified, for each miRNA and each direction of transport, the best unsaturated model, as shown in Table 5, Table 6, and

Table 7. Note that instead of compartments *A*, *B*, *C* in the general models, we used *P* to represent placenta, *M* maternal plasma, and *F* fetal plasma (cord blood).
The interpretation of our results for the transport of miRNA from placenta to maternal/fetal blood is relatively straightforward. For the transport of miRNA from placenta to pre-delivery maternal plasma, in Table 5, we found that for most miRNAs, the only non-zero parameter in the best model is $k_{P,M,i}$, which suggests that the selected miRNAs in maternal plasma are derived entirely from the placenta. The only miRNA where the model suggests no transport is miR-517c-3p, for which, not surprisingly, the correlation between pre-delivery maternal plasma and the placenta is almost 0. The results for the miRNA transport from the placenta to the cord blood are shown in Table 6. Similar to the results for placenta to maternal blood transport, for most miRNAs, the only non-zero parameter in the best model is $k_{P,F,i}$, which suggests that the selected miRNAs in cord blood come entirely from the placenta. The best models for two miRNAs, miR-517c-3p and miR-524-5p, have $k_{P,F,i}$, = 0, suggesting the absence of transport from the placenta to the cord blood, which is consistent with the observation that these miRNAs show weakly positive or negative correlation between placenta and cord blood.

The results for miRNA transport from maternal/fetal blood to placenta are more complicated. As shown in

Table 7, for most of the miRNAs, the best model contains a non-negative term $V_{P,i}$, which suggests the native expression of these miRNAs in placenta. This is certainly plausible given that placenta contains a mixture of cell types. We also note that in the model for miR-185-5p, the parameter $k_{M,P,i}$ is negative, suggesting a more complicated relationship between the expression of miR-185-5p in the maternal plasma and in the placenta.

Analysis by other platforms

With sequencing data from a small number of libraries, our analysis was limited. For example, when testing for each miRNA, the correlation of its expression in the placenta and in maternal/fetal blood, none of the tests yielded a significant p value. We therefore explored other platforms for estimation of miRNA transport. In particular, we investigated the use of Taqman RT-PCR cards, Ocean Ridge miRNA microarray, and Nanostring miRNA assay, and used data from these platforms to validate and supplement the results from sequencing data.

Taqman data analysis
We obtained expression data for miRNAs in the maternal plasma, cord blood, and placenta from 11 patients using Taqman miRNA RT-PCR cards (types A+B). When analyzing the Taqman data, we encountered two problems. The first was finding a suitable reference gene that could be used to normalize the measured miRNA expression levels. The manufacturer's suggested endogenous control, has-miR-654-3p, was undetectable in all plasma samples. One of the non-miRNA control, snRNA U6, was detected in both the placenta and plasma samples. However, the expression of U6 varied greatly between placenta and plasma. In the placenta, the Ct value of

U6 was about 1.5 less than the most abundant miRNA, while in the maternal and fetal plasma, the Ct value of U6 was about 12.9 higher than the most abundant miRNA. We therefore elected to use non-normalized Taqman data, given the lack of acceptable reference gene. The hierarchical clustering and multidimensional scaling plots of the Taqman data are presented in Figure 3 and Figure 4.

The second concern with the Taqman data was anomaly values for some miRNAs in certain samples. For example, the Ct values of hsa-miR-363-3p in plasma sample were all between 32 and 35, except for two maternal plasma samples. In those two maternal plasma samples, the Ct values of hsa-miR-363-3p were 25. We decided to apply a modified 1.5 IQR (Inter-Quartile Range) rule for outliers to identify those measurements that were likely to be outliers. More precisely, for each miRNA and each type of tissue, we calculated the 25% and 75% quantiles of the Ct values, as well as the difference between the two quartiles, which is the IQR. We then removed all entries that were 1.5 IQR or 2.25 cycles below the 25% quantile, or 1.5 IQR or 2.25 cycles above the 75% quartile, with the following exceptions: all entries within 1.5 cycles of the 25% or 75% quartiles were kept. We decided to study the previously selected 14 miRNAs in Table 2 using the processed Taqman data. Because some of the miRNAs were not represented on the Taqman cards, or were undetectable (mostly in plasma samples), we ended up with data for 10 miRNAs, represented by nine Taqman probes in maternal plasma, placenta, and cord blood from ten patients. Their correlation between placenta and maternal/fetal blood, and the log2 fold change of their levels after delivery are shown in Table 8 and Table 9. For each individual miRNA, the correlations were not significant, possibly due to the small sample size. However, considering the nine placenta miRNAs as a group, we found that their correlations between maternal/fetal blood and placenta were significantly above 0 (p values both equal to 0.03125, using Wilcoxon signed rank test). Moreover, for the 2 miRNAs where post-delivery maternal plasma data were available, we found that their levels in maternal plasma were significantly decreased after delivery.

Analyzing the Taqman data using the same method as we did for the sequencing data, we identified the best model, for each direction of miRNA transport between placenta and maternal/fetal plasma, as presented in Table 10, Table 11, and Table 12. The results are similar to those we obtained from the sequencing data. The main conclusion is that, for the majority of studied placenta miRNAs, their expression in maternal plasma/cord blood is derived almost entirely from the placenta. For plasma miRNAs, they were likely natively expressed in placenta, as suggested by the presence of the parameter $V_{F,i,j}$ in the best models for all plasma miRNAs. Moreover, for some miRNAs, the negative parameters in the best models suggest a more complicated relationship between their expression in maternal plasma and in placenta.

Array Data Analysis

We used Ocean Ridge miRNA microarrays to measure the miRNA levels in the maternal/fetal plasma and placenta samples from 37 patients, of which the data from 35 patients passed the quality control. The Ocean Ridge miRNA microarray platform has three probes for each target miRNA. We applied a modified two-step RMA normalization/summarization procedure to the placenta and plasma samples separately (Irizarry et al, 2003). A total of 147 samples, which passed the quality control, were divided into two groups, plasma group (109 samples) and placenta group (38 samples). The array data in each group were initially pre-processed by

quantile normalization. Next, the expression values of each miRNA were estimated by fitting the normalized intensities of three probes for the miRNA to a linear model using the robust median polish (MP) algorithm (Holder et al, 2001). The hierarchical clustering and multidimensional scaling plots of the microarray data are shown in Figure 5 and Figure 6.

We studied the previously selected miRNAs in Table 2 using the processed miRNA microarray data. Their correlation between placenta and maternal/fetal blood, and the log2 fold change of their level in maternal blood after delivery, are shown in Table 13 and Table 14. We noticed that the correlations between the miRNAs in the placenta and maternal/fetal blood are generally weaker (possibly because of the higher noise of the array data), although the correlation between placenta and maternal blood for the nine placenta miRNAs as a group were significantly above 0 (p value = 0.002 using one-sided Wilcoxon signed rank test). Furthermore, the level of the nine placenta miRNAs in maternal plasma were all significantly decreased after delivery.

Using the approach we deployed for the sequencing data, we identified the best models for each direction of miRNA transport between placenta and maternal/fetal plasma, as presented in Table 15, Table 16, and Table 17. Again, these models show strong similarity to the models we obtained from the sequencing data and the Taqman data. For the majority of the nine placenta miRNAs, the models from the microarray data suggest that their expression in maternal plasma and cord blood is entirely derived from the placenta. For the plasma miRNAs, the results based on the microarray data suggested that they were all natively expressed in placenta, with only one miRNA apparently transported from maternal plasma to placenta.

Nanostring Data Analysis

We further explored the Nanostring miRNA assays to examine miRNA levels in maternal plasma, placenta, and cord blood samples from 3 patients. Each Nanostring miRNA assay contains 800 probes targeting human miRNAs. We processed the data as suggested by the manufacturer, i.e., normalizing the data by the positive controls, and performing probe correction for those probes that need background subtraction. The hierarchical clustering and multidimensional scaling plots of the Nanostring data are presented in Figure 7 and Figure 8. Unfortunately, we found that several maternal plasma and cord blood samples showed very low signal. For example, in one maternal plasma sample, none of the miRNAs studied for plasma to placenta transport was above the background threshold. In light of the low signal for some of the plasma samples, we decided not to use the Nanostring data to study the miRNA transport in this project.

Conclusion

We analyzed plasma miRNA levels from pre-delivery maternal blood and fetal cord blood, as well as placenta miRNA samples, utilizing a variety of platforms, include Illumina sequencing, Taqman RT-PCR, and Ocean Ridge miRNA microarray. For a selection of 14 miRNAs, we studied patterns of miRNA transport between maternal blood, placenta, and fetal blood. In particular, we found that the nine placental miRNAs, which are abundant in PHT cells, absent in normal non-pregnant blood, and present in maternal/fetal blood, have a positive correlation between placenta and maternal/fetal blood, and tend to decrease significantly in maternal blood after delivery. Moreover, through a comprehensive information score based model selection

procedure, we found that the best model for the miRNAs transport suggests that these miRNAs in maternal/fetal blood are derived entirely from the placenta. Although our results for the plasma miRNAs, which are abundant in normal non-pregnant blood, almost absent in PHT cells, and present in placenta, are not as clear, our data suggest that the plasma miRNAs are likely expressed natively in placenta, and the simple transport model, as represented in Equation 3, may be not a good explanation of the expression of all these miRNAs in placenta.

Overall, we have been successful in developing most useful and relevant tools to interrogate the transport of miRNAs among the maternal-placental and fetal compartments, and provided essential information on the capabilities as well as limitations of today's technologies in meaningful analysis of plasma miRNAs. To our knowledge, the studies described here have not been previously performed. In the next phase of our studies we plan to interrogate additional sequencing data (work currently in progress), and decipher the impact of several placental disorders that manifest as fetal growth restriction and similar disorders related to placental dysfunctions. Long term, information gleaned from our data not only illuminates epigenomic pathways that influence human embryonic development, but may also suggest innovative, clinically-relevant biomarkers that define unique patterns of placental injury.

References

Akaike, Hirotugu (1974). "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* **19** (6): 716–723.

Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K., Guo, J., (...), Zhang, C.-Y. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research* 18 (10): 997-1006.

Irizarry, RA, Hobbs, B, Collin, F, Beazer-Barclay, YD, Antonellis, KJ, Scherf, U, Speed, TP (2003) Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data. *Biostatistics* .Vol. 4, Number 2: 249-264

Holder, D., Raubertas, R. F., Pikounis, V. B., Svetnik, V., and Soper, K. (2001). Statistical analysis of high density oligonucleotide arrays: a SAFER approach. In: *Proceedings of the ASA Annual Meeting, Atlanta, GA 2001*.

Table 1. Size of miRNA Sequencing Libraries

Patient ID	Sample	Size
122	Placenta	2142379
122	PreDelivery	1907871
122	CordBlood	1506999
122	PostDelivery	1835196
123	Placenta	3219145
123	PreDelivery	2436866
123	CordBlood	2277488
123	PostDelivery	3619539
127	Placenta	4855815
127	PreDelivery	3836054
127	CordBlood	5106835
127	PostDelivery	6248546
129	Placenta	4954475
129	PreDelivery	10351
129	CordBlood	1990521
129	PostDelivery	1007455
136	Placenta	4151283
136	PreDelivery	377012
136	CordBlood	802148
136	PostDelivery	1399432

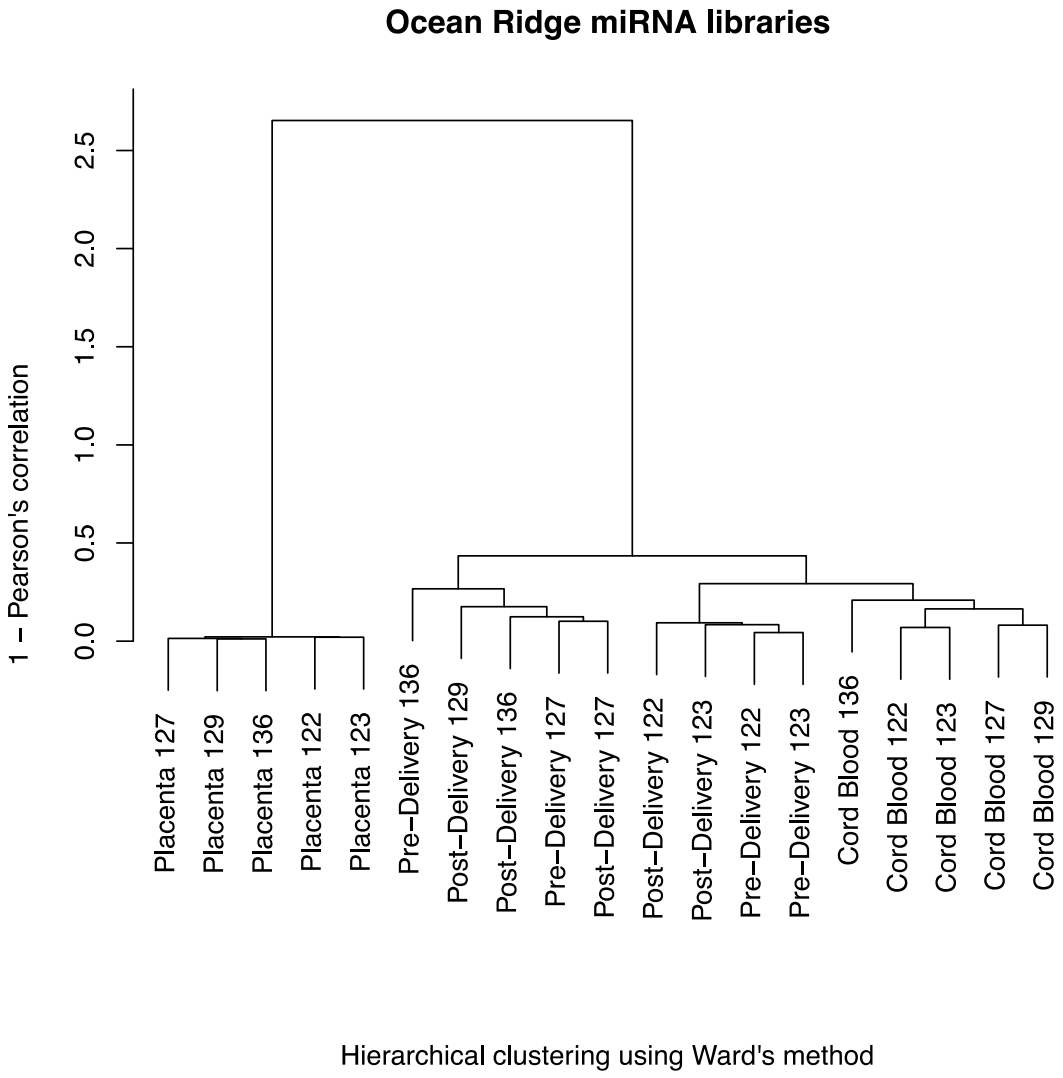

Figure 1. Hierarchical Clustering, Sequencing

Table 2. Selected miRNAs and Their Mean Expression (RPM), Sequencing

miRNA	Placenta	PreDelivery	CordBlood	PostDelivery
hsa-miR-1323	58461	3939	3345	143
hsa-miR-516b-5p	21424	330	584	35
hsa-miR-517a-3p	49292	395	235	29
hsa-miR-517b-3p	49292	395	235	29
hsa-miR-517c-3p	13528	119	66	12
hsa-miR-520a-3p	4045	157	83	5
hsa-miR-524-5p	5362	483	169	41
hsa-miR-525-5p	3239	340	194	18
hsa-miR-526b-5p	4286	124	65	7
hsa-miR-144-3p	12381	23193	31802	17773
hsa-miR-185-5p	276	4253	6602	3131
hsa-miR-363-3p	207	1893	1444	1217
hsa-miR-451a	9202	223630	136047	234267
hsa-miR-7-5p	152	409	924	606

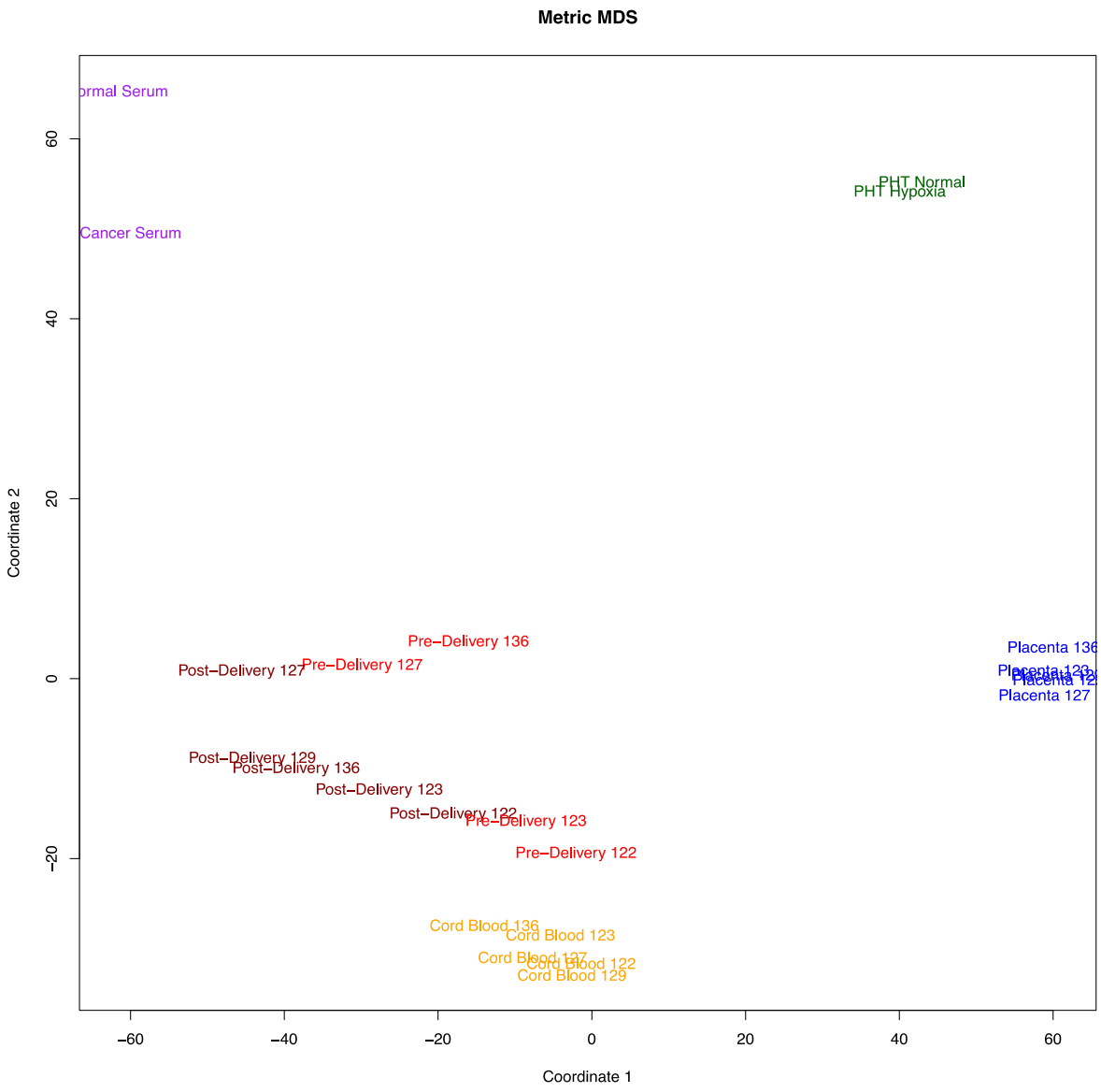

Figure 2. Multidimensional Scaling, Sequencing

Table 3. Correlation between Placenta and Maternal/Fetal Blood, Sequencing

	Placenta--PreDelivery Correlation	Placenta--CordBlood Correlation
hsa-miR-1323	0.539449211	0.607878406
hsa-miR-520a-3p	0.324152372	0.599214362
hsa-miR-517a-3p	0.564214857	0.460867625
hsa-miR-516b-5p	0.615872108	0.562551431
hsa-miR-517b-3p	0.564214857	0.460867625
hsa-miR-517c-3p	0.027033153	0.166786978
hsa-miR-524-5p	0.334381172	-0.510917069
hsa-miR-525-5p	0.701157874	0.711207708
hsa-miR-526b-5p	0.710354942	0.519860007
hsa-miR-144-3p	0.608842401	0.89830598
hsa-miR-185-5p	-0.630484399	0.299112482
hsa-miR-363-3p	0.401250356	0.334576756
hsa-miR-451a	0.284217849	0.419488432
hsa-miR-7-5p	-0.010385314	0.237978178

Table 4. Log2 Fold Change in Maternal Blood after Delivery, Sequencing

	Log2(preDlivery/postDelivery)	P value
hsa-miR-1323	4.531500	0.008813
hsa-miR-520a-3p	5.024450	0.029463
hsa-miR-517a-3p	3.640336	0.023903
hsa-miR-516b-5p	3.661570	0.063410
hsa-miR-517b-3p	3.640336	0.023903
hsa-miR-517c-3p	3.965785	0.045080
hsa-miR-524-5p	3.593982	0.003275
hsa-miR-525-5p	4.015166	0.008865
hsa-miR-526b-5p	3.531387	0.012597

Table 5. Placenta to Maternal Blood Transport, Sequencing

miRNA	Best model for $X_{M,i,j}$	Estimated $k_{P,M,i}$	p value
hsa-miR-1323	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.076422	0.135931
hsa-miR-516b-5p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.017494	0.099588
hsa-miR-517a-3p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.008516	0.116033
hsa-miR-517b-3p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.008516	0.116033
hsa-miR-517c-3p	$V_{M,i,j} + \square'_{M,i,}$	0	NA
hsa-miR-520a-3p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.039584	0.088062
hsa-miR-524-5p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.090447	0.006356
hsa-miR-525-5p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.116614	0.041382
hsa-miR-526b-5p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.183686	0.289645

Table 6. Placenta to Fetal Blood Transport, Sequencing

miRNA	Best model for $X_{F,i,j}$	Estimated $X_{F,i,j}$	p value
hsa-miR-1323	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.060750	0.041900
hsa-miR-516b-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.028331	0.027860
hsa-miR-517a-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.004772	0.000625
hsa-miR-517b-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.004772	0.000625
hsa-miR-517c-3p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-520a-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.020738	0.004873
hsa-miR-524-5p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-525-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.063914	0.032198
hsa-miR-526b-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.015972	0.068726

Table 7. Maternal/Fetal Blood to Placenta Transport, Sequencing

miRNA	Best model for $X_{P,i,j}$	Estimate $k_{M,P,i}$	Estimated $k_{F,P,i}$
hsa-miR-144-3p	$k_{F,P,i}\ X_{F,i,j} + V_{P,i,j} + \square'_{P,i,j}$	0	0.145311
hsa-miR-185-5p	$k_{M,P,i}\ X_{M,i,j} + k_{F,P,i}\ X_{F,i,j} + V_{P,i,j} + \square'_{P,i,j}$	-0.020076	0.032884
hsa-miR-363-3p	$V_{P,i,j} + \square'_{P,i,j}$	0	0
hsa-miR-451a	$k_{M,P,i}\ X_{M,i,j} + \square'_{P,i,j}$	0.044540	0
hsa-miR-7-5p	$V_{P,i,j} + \square'_{P,i,j}$	0	0

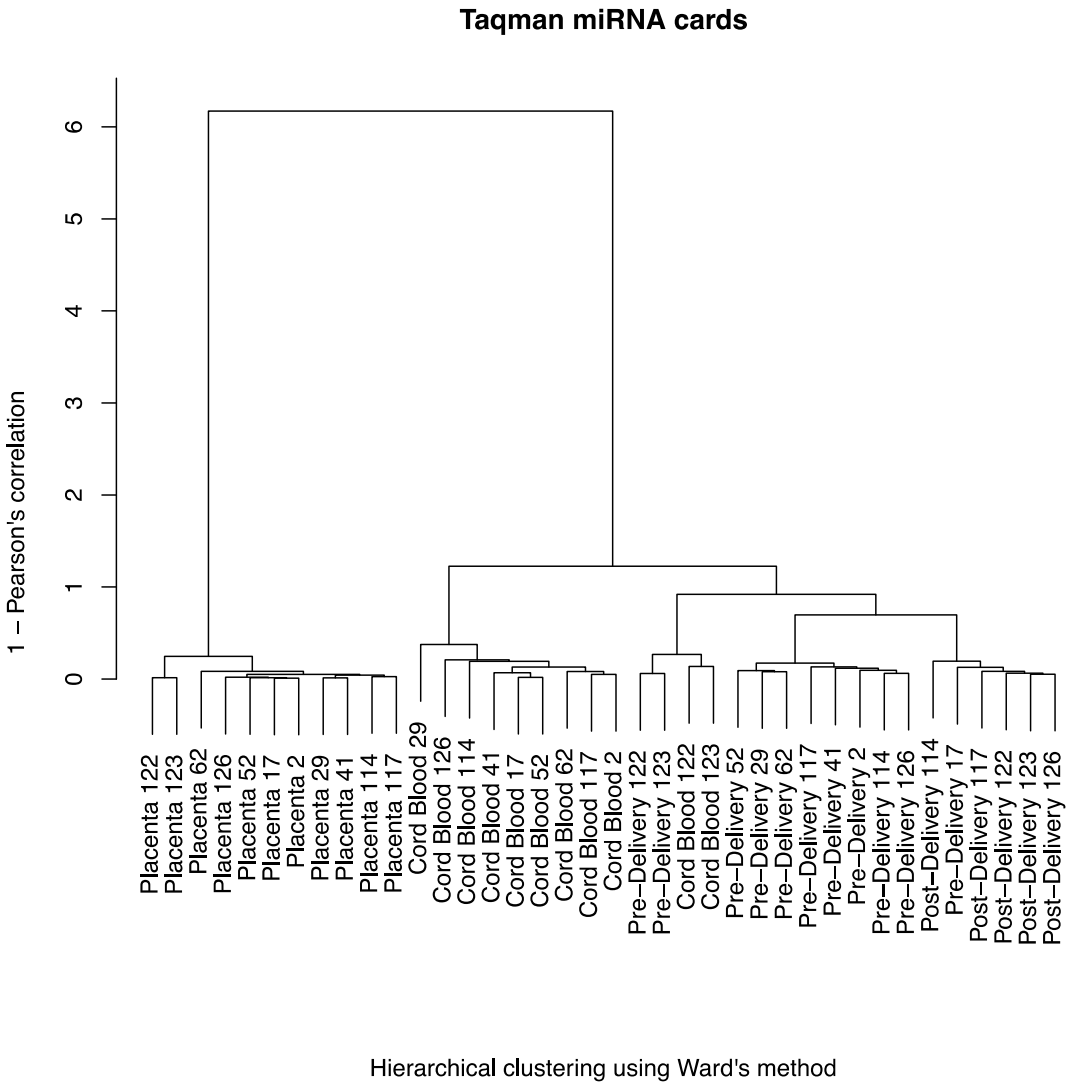

Figure 3. Hierachical Clustering, Taqman

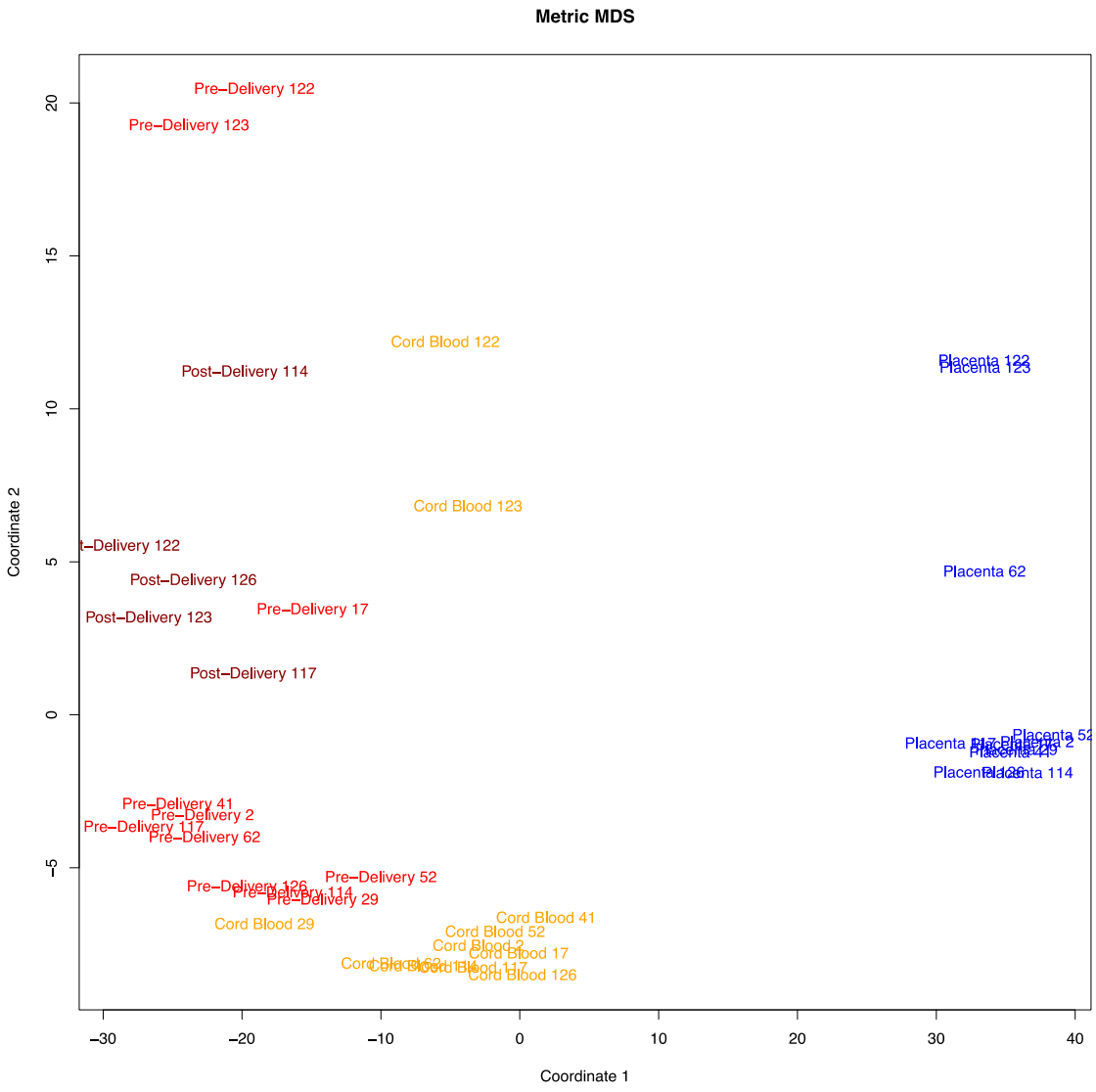

Figure 4. Multidimensional Scaling, Taqman

Table 8. Correlation between Placenta and Maternal/Fetal Blood, Taqman

	Placenta--PreDelivery Correlation	Placenta--CordBlood Correlation
hsa-miR-516b-5p	0.016306072	0.153201058
hsa-miR-517a-3p, hsa-miR-517b-3p	0.768760052	0.344712503
hsa-miR-517c-3p	0.458006888	0.168203231
hsa-miR-525-5p	0.59312366	0.588579984
hsa-miR-526b-5p	0.842344351	0.264952323
hsa-miR-185-5p	0.195500025	0.152391957
hsa-miR-363-3p	0.261055277	0.307419148
hsa-miR-451a	-0.200713802	-0.257254822
hsa-miR-144-3p	0.194850907	0.302591228

Table 9. Log2 Fold Change in Maternal Blood after Delivery, Taqman

	Log2(preDlivery/post Delivery)	P value
hsa-miR-516b-5p	NA	NA
hsa-miR-517a-3p, hsa-miR-517b-3p	3.916200	0.000608
hsa-miR-517c-3p	4.069800	0.002514
hsa-miR-525-5p	NA	NA
hsa-miR-526b-5p	NA	NA

Table 10. Placenta to Maternal Blood, Taqman

miRNA	Best model for $X_{M,i,j}$	Estimated $k_{P,M,i}$	p value
hsa-miR-516b-5p	$V_{M,i,j} + \square'_{M,i,j}$	0	NA
hsa-miR-517a-3p, hsa-miR-517b-3p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.000781	0.000681
hsa-miR-517c-3p	$V_{M,i,j} + \square'_{M,i,j}$	0	NA
hsa-miR-525-5p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.002034	0.002151
hsa-miR-526b-5p	$k_{P,M,i}\ X_{P,i,j} + \square'_{M,i,j}$	0.010706	0.008847

Table 11. Placenta to Cord Blood, Taqman

miRNA	Best model for $X_{F,i,j}$	Estimated $k_{P,F,i}$	p value
hsa-miR-516b-5p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-517a-3p, hsa-miR-517b-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.004329	0.006224
hsa-miR-517c-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.005896	0.011689
hsa-miR-525-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.010655	0.016560
hsa-miR-526b-5p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA

Table 12. Maternal/Fetal Blood to Placenta, Taqman

miRNA	Best model for $X_{P,i,j}$	Estimate $k_{M,P,i}$	Estimated $k_{F,P,i}$
hsa-miR-185-5p	$V_{P,i,j}$ + □$'_{P,i,j}$	0	0
hsa-miR-363-3p	$k_{M,P,i}\ X_{M,i,j}$ + $V_{P,i,j}$ + □$'_{P,i,j}$	-0.391335	0
hsa-miR-451a	$k_{F,P,i}\ X_{F,i,j}$ + $V_{P,i,j}$ + □$'_{P,i,j}$	0	0.133458
hsa-miR-144-3p	$k_{M,P,i}\ X_{M,i,j}$ + $V_{P,i,j}$ + □$'_{P,i,j}$	-0.056234	0

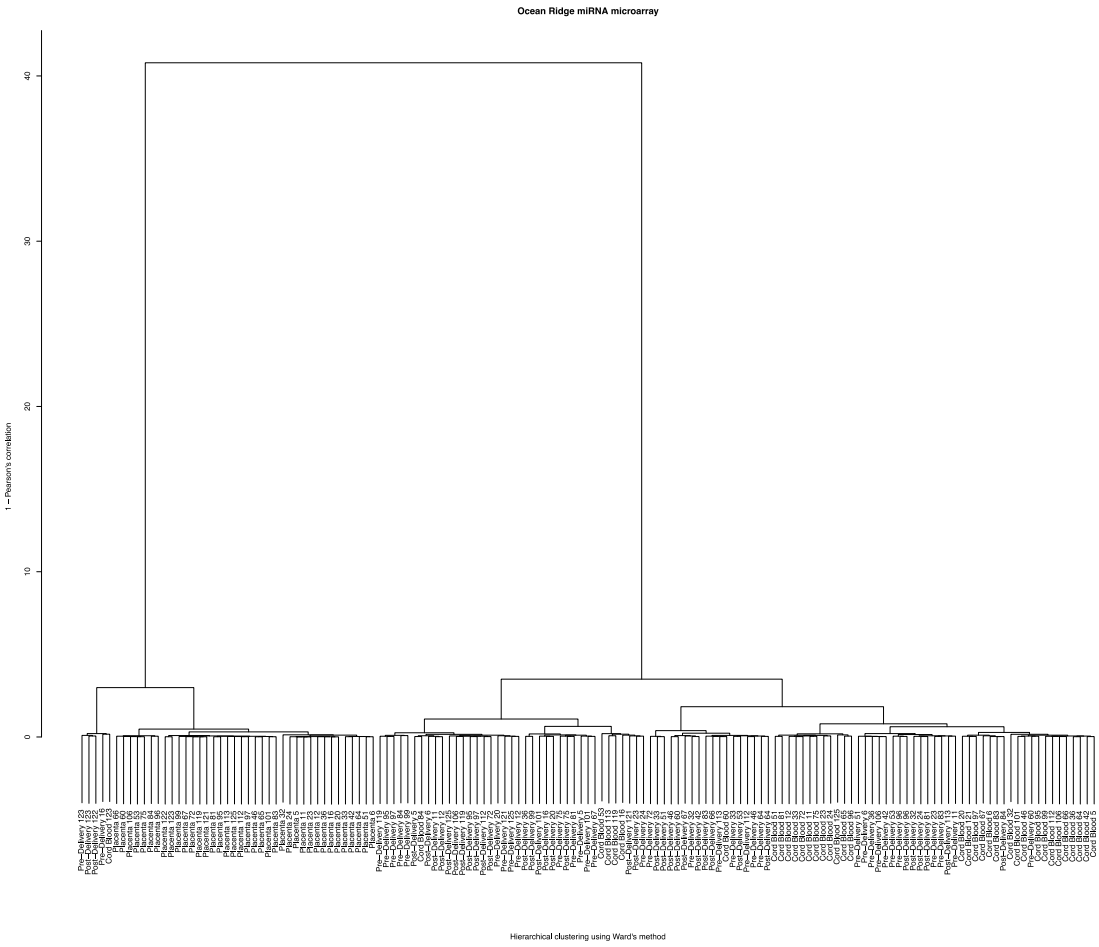

Figure 5. Hierarchical Clustering, Microarray

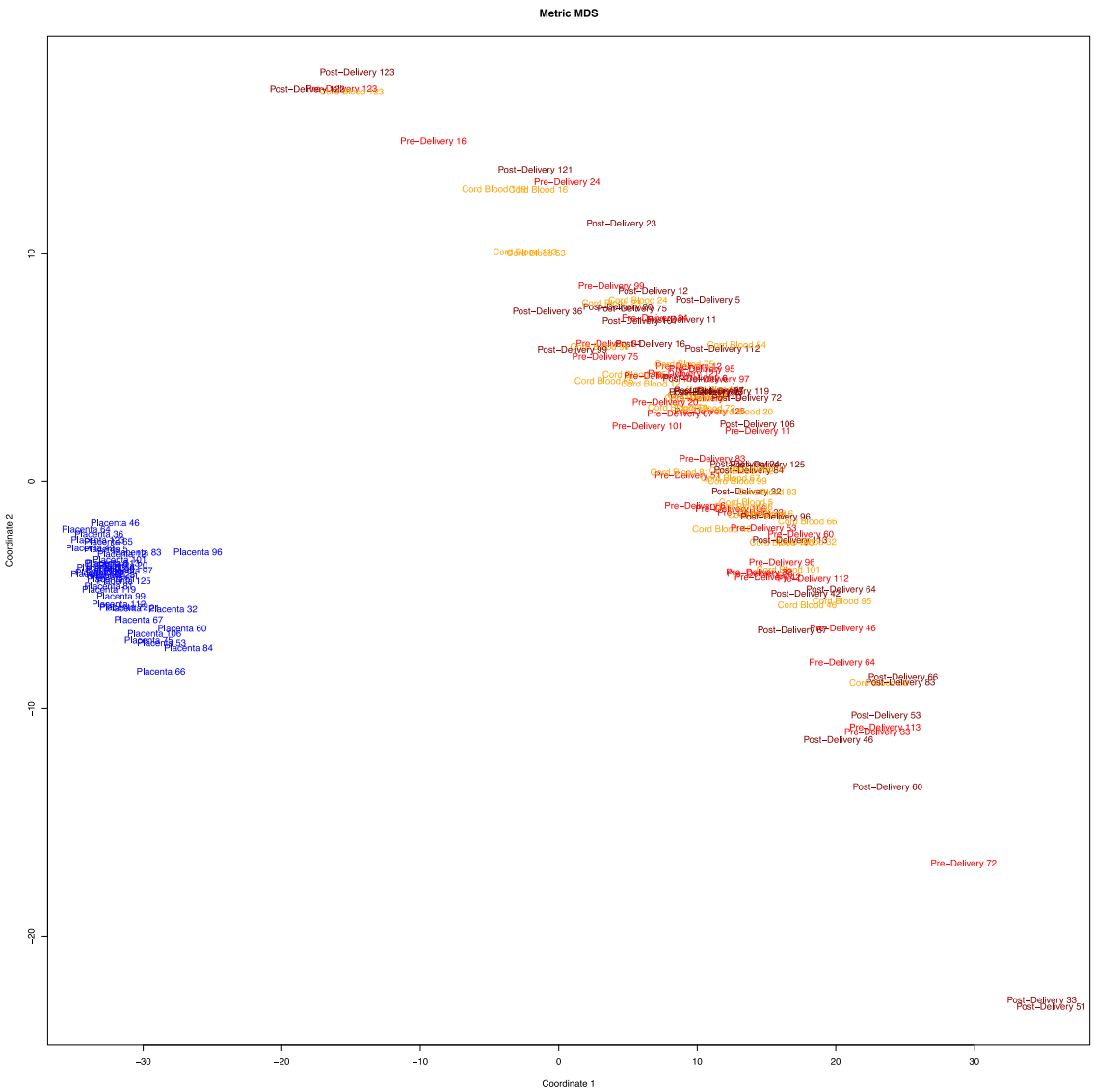

Figure 6. Multidimensional Scaling, Microarray

Table 13. Correlation between Placenta and Maternal/Fetal Blood, Microarray

	Placenta--PreDelivery Correlation	Placenta--CordBlood Correlation
hsa-miR-1323	0.157570189	-0.029162563
hsa-miR-516b-5p	0.259172658	0.012054593
hsa-miR-517a-3p	0.217424876	0.028102868
hsa-miR-517b-3p	0.177174881	0.009515001
hsa-miR-517c-3p	0.057236131	-0.204693664
hsa-miR-520a-3p	0.112596249	-0.191065049
hsa-miR-524-5p	0.144617848	0.154242837
hsa-miR-525-5p	0.131304417	0.152587799
hsa-miR-526b-5p	0.16575452	0.216923858
hsa-miR-144-3p	0.282932096	0.207292039
hsa-miR-185-5p	-0.046744423	0.168724428
hsa-miR-363-3p	-0.145872583	0.004012128
hsa-miR-451a	-0.053754624	-0.056743029
hsa-miR-7-5p	0.175728305	0.193483961

Table 14. Log2 Fold Change in Maternal Plasma after Delivery, Microarray

	Log2(preDlivery/postDelivery)	P value
hsa-miR-1323	2.10307234	6.46E-12
hsa-miR-516b-5p	1.387718289	1.05E-05
hsa-miR-517a-3p	2.314685006	9.65E-10
hsa-miR-517b-3p	1.609318868	3.58E-08
hsa-miR-517c-3p	0.900851156	7.49E-05
hsa-miR-520a-3p	0.369063636	3.52E-04
hsa-miR-524-5p	3.398331012	3.36E-14
hsa-miR-525-5p	1.768799349	7.43E-10
hsa-miR-526b-5p	2.000019891	3.96E-09

Table 15. Placenta to Maternal Blood Transport, Microarray

miRNA	Best model for $X_{M,i,j}$	Estimated $k_{P,M,i}$	p value
hsa-miR-1323	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.028210	0.000185
hsa-miR-516b-5p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.005477	0.042958
hsa-miR-517a-3p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.020433	0.004241
hsa-miR-517b-3p	$V_{M,i,j} + \square'_{M,i,j}$	0	NA
hsa-miR-517c-3p	$V_{M,i,j} + \square'_{M,i,j}$	0	NA
hsa-miR-520a-3p	$V_{M,i,j} + \square'_{M,i,j}$	0	NA
hsa-miR-524-5p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.019634	0.000653
hsa-miR-525-5p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.009980	0.000603
hsa-miR-526b-5p	$k_{P,M,i}\ X_{P,i,j} + V_{M,i,j} + \square'_{M,i,j}$	0.009622	0.003082

Table 16. Placenta to Fetal Blood Transport, Microarray

miRNA	Best model for $X_{F,i,j}$	Estimated $k_{P,F,i}$	p value
hsa-miR-1323	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.024787	0.000032
hsa-miR-516b-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.006792	0.000943
hsa-miR-517a-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.033878	0.001428
hsa-miR-517b-3p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.027618	0.000098
hsa-miR-517c-3p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-520a-3p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-524-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.011618	0.000045
hsa-miR-525-5p	$V_{F,i,j} + \square'_{F,i,j}$	0	NA
hsa-miR-526b-5p	$k_{P,F,i}\ X_{P,i,j} + \square'_{F,i,j}$	0.004144	3.16E-07

Table 17. Maternal/Fetal Blood to Placenta Transport, Microarray

miRNA	Best model for $X_{P,i,j}$	Estimate $k_{M,P,i}$	Estimated $k_{F,P,i}$
hsa-miR-144-3p	$V_{P,i,j} + \square'_{P,i,j}$	0	0
hsa-miR-185-5p	$V_{P,i,j} + \square'_{P,i,j}$	0	0
hsa-miR-363-3p	$V_{P,i,j} + \square'_{P,i,j}$	0	0
hsa-miR-451a	$V_{P,i,j} + \square'_{P,i,j}$	0	0
hsa-miR-7-5p	$k_{M,P,i}\ X_{M,i,j} + V_{P,i,j} + \square'_{P,i,j}$	0.495761	0

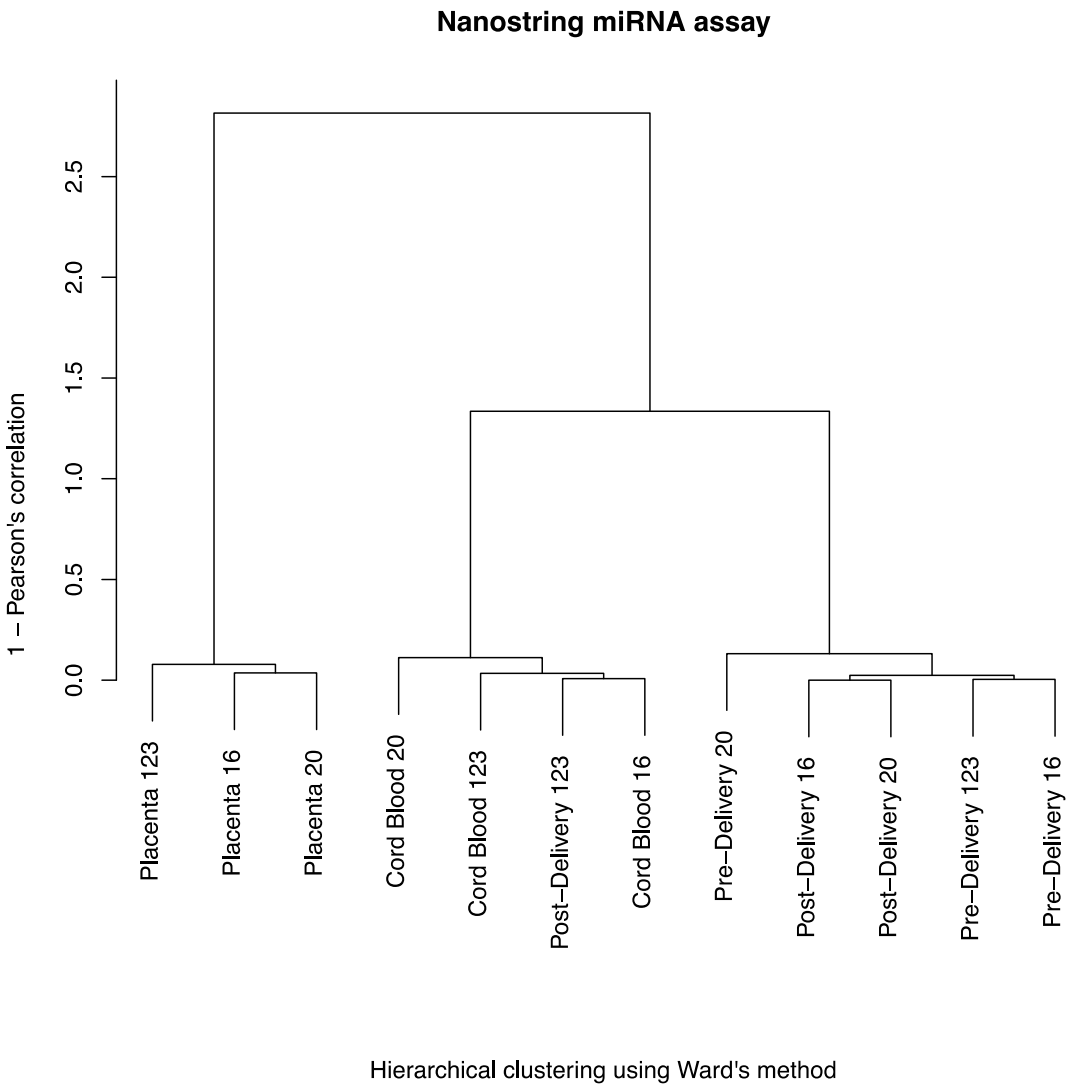

Figure 7. Hierarchical Clustering, Nanostring

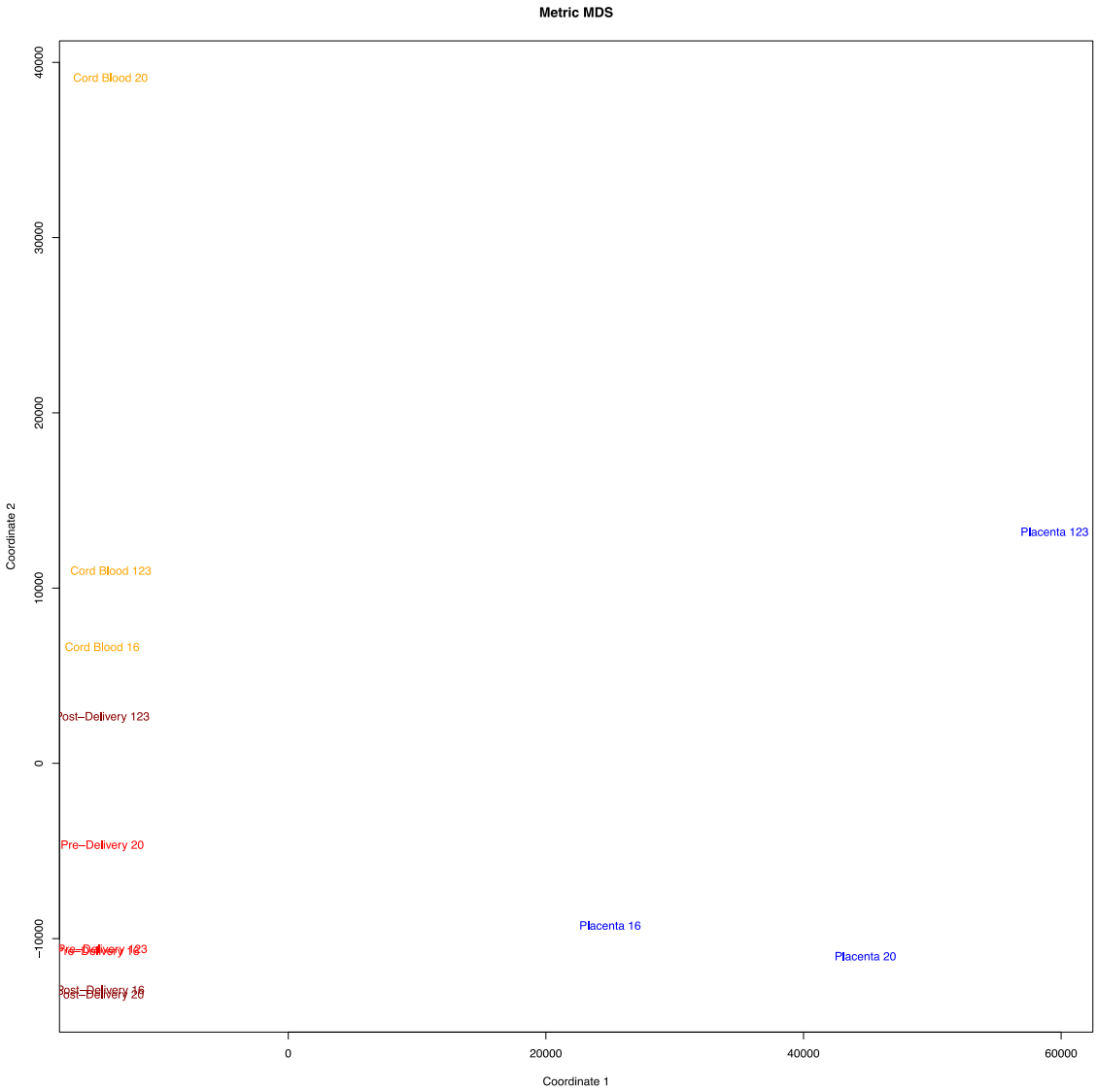

Figure 8. Multidimensional Scaling, Nanostring

18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X__No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X___No

If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

_____Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

_____Number of subjects originally targeted to be included in the study
_____Number of subjects enrolled in the study

Note: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females
_____Unknown

Ethnicity:
_____Latinos or Hispanics
_____Not Latinos or Hispanics

_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

19. Human Embryonic Stem Cell Research. Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
__X___ No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?
_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

20. Articles Submitted to Peer-Reviewed Publications.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, the number of the publication and

an abbreviated research project title. For example, if you submit two publications for PI Smith for the "Cognition and MRI in Older Adults" research project (Project 1), and two publications for PI Zhang for the "Lung Cancer" research project (Project 3), the filenames should be:

Project 1 – Smith – Publication 1 – Cognition and MRI
Project 1 – Smith – Publication 2 – Cognition and MRI
Project 3 – Zhang – Publication 1 – Lung Cancer
Project 3 – Zhang – Publication 2 – Lung Cancer

If the publication is not available electronically, provide 5 paper copies of the publication.

Note: The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

Title of Journal Article:	Authors:	Name of Peer-reviewed Publication:	Month and Year Submitted:	Publication Status (check appropriate box below):
1. None				☐Submitted ☐Accepted ☐Published

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes___X_____ No_____

If yes, please describe your plans:

We plan to submit a paper in the fall of 2012 to present our findings in this project. This manuscript is currently being written.

21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project. Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_____ No____X

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____

If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____
If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No____

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes___ No____

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes_____ No___X_____

If yes, please describe your plans:

24. Key Investigator Qualifications. Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

BIOGRAPHICAL SKETCH	
NAME Tianjiao Chu, PhD	POSITION TITLE Assistant Professor of OBGYN and Reproductive Sciences
eRA COMMONS USER NAME	

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
Shandong University, Jinan, China	BA	1989	Chinese Lang. & Lit
University of Toledo, Toledo, OH	MA	1997	Philosophy
Carnegie Mellon University, Pittsburgh, PA	MS	2001	Data Mining
Carnegie Mellon University, Pittsburgh, PA	PhD	2003	Logic & Computation

A. Personal statement

My research interest and experience have been mainly in the developing the statistical models and computational algorithms for the analysis of biological/medical datasets, especially the high throughput genomic and proteomic data. Recently I have been working on the developing of statistical procedures for non-invasive prenatal diagnosis of aneuploidy, as well as algorithms studying the relations among miRNA, mRNA, and proteins in human placenta.

Positions and employment

2003-2005 Research Scientist, Institute for Human and Machine Cognition, Univ. of West Florida
2005-2007 Research Scientist, JustSystems Evans Research, Pittsburgh, PA
2007-present Assistant Professor, Dept. of OBGYN and Reproductive Sciences, University of Pittsburgh

Honors

2008-2010 Program committee member of the 24th, 25th, and 26th Conferences on Uncertainty in Artificial Intelligence
2010 Prenatal Diagnosis 2009 Malcolm Ferguson-Smith Young Investigator Award
2010 Best paper on next generation sequencing in *Bioinformatics*

C. Selected publications – Refereed articles

1. **Chu T**, Glymour C, Scheines R, Spirtes P. A statistical problem for inference to regulatory structure from associations of gene expression measurement with microarrays. *Bioinformatics* 2003;19:1147-52. PMID: 12801876
2. Ning W, **Chu T**, Li CJ, Choi A, Peters D. Genome Wide Analysis of the Endothelial Transcriptome under Short-Term Chronic Hypoxia. *Physiol Genomics* 2004;18:70-8. PMID: 15100389
3. Peters DG, Kudla DM, Deloia JA, **Chu TJ**, Fairfull L, Edwards RP, Ferrell RE. Comparative gene expression analysis of ovarian carcinoma and normal ovarian epithelium by serial analysis of gene expression. *Cancer Epidemiol Biomarkers Prev.* 2005 Jul;14(7):1717-23.
4. Peters DG, Ning W, **Chu TJ**, Li CJ, Choi AMK. Comparative SAGE analysis of the response to hypoxia in human pulmonary and aortic endothelial cells. *Physiol Genomics* 2006;26:99-108. PMID: 16595741
5. **Chu T.** Limitations of statistical learning from gene expression data. *Interface 2004: Computing Science and Statistics* 2008;36:266-85.
6. **Chu TJ**, Peters DG. Serial analysis of the vascular endothelial transcriptome under static and shear stress conditions. *Physiol Genomics* 2008;34:185-92. PMID: 18505769
7. **Chu T**, Glymour C. Search for additive nonlinear time series causal models. *Journal of Machine Learning Research* 2008;9:967-91.
8. **Chu T**, Bunce K, Hogge WA, Peters DG. Statistical model for whole genome sequencing and its application to minimally invasive diagnosis of fetal genetic disease. *Bioinformatics* 2009;15;25:1244-50. PMID: 19307238
9. **Chu T**, Burke B, Bunce K, Surti U, Allen Hogge W, Peters DG. A microarray-based approach for the identification of epigenetic biomarkers for the noninvasive diagnosis of fetal disease. *Prenat Diagn* 2009 Nov;29(11):1020-30. PMID: 19650061

10. Mouillet JF, **Chu T**, Hubel C, Nelson DM, Parks WA, Sadovsky Y. MiR-205 Silences MED1 in Hypoxic Primary Human Trophoblasts. *FASEB* 2010;24(6):2030-9. PMID: 20065103
11. Sheng Y, Lin CC, Yue J, Sukhwani M, Shuttleworth JJ, **Chu T**, Orwig KE. Generation and characterization of a Tet-On (rtTA-M2) transgenic rat. *BMC Dev Biol.* 2010 Feb 16;10:17.
12. Mouillet JF, **Chu T**, Nelson DM, Mishima T, Sadovsky Y. The levels of hypoxia-regulated microRNAs in plasma of pregnant women with fetal growth restriction. *Placenta* 2010;31(9): 781-4. PMID:20667590
13. **Chu T**, Bunce K, Hogge WA, Peters DG. Statistical considerations for digital approaches to non-invasive fetal genotyping. *Bioinformatics* 2010;26(22):2863-6. PMID: 20870643
14. **Chu T**, Bunce K, Hogge WA, Peters DG. A novel approach toward the challenge of accurately quantifying fetal DNA in maternal plasma. *Prenat Diagn* 2010;30(12-13):1226-9. PMID:21072784
15. Ochalski ME, Shuttleworth JJ, **Chu T**, Orwig KE. Thalidomide treatment attenuates chemotherapy-induced gonadal toxicity. *Fertil Steril.* 2011 Feb;95(2):819-22.
16. Rajakumar A, **Chu T**, Handley DE, Bunce KD, Burke B, Hubel CA, Jeyabalan A, Peters DG. Maternal gene expression profiling during pregnancy and preeclampsia in human peripheral blood mononuclear cells. *Placenta*. 2011;32(1):70-8. PMID: 21075447
17. **Chu T**, Handley D, Bunce K, Surti U, Hogge WA, Peters DG. Structural and Regulatory Characterization of the Placental Epigenome at Its Maternal Interface. *PLoS One* 2011;6(2): e14723. doi:10.1371/journal.pone.0014723
18. Mouillet JF, **Chu T**, Sadovsky Y. Expression patterns of placental microRNAs. *Birth Defects Res A Clin Mol Teratol.* 2011 Aug;91(8):737-43. doi: 10.1002/bdra.20782. Epub 2011 Mar 21.
19. Budiu RA, Mantia-Smaldone G, Elishaev E, **Chu T**, Thaller J, McCabe K, Lenzner D, Edwards RP, Vlad AM. Soluble MUC1 and serum MUC1-specific antibodies are potential prognostic biomarkers for platinum-resistant ovarian cancer. *Cancer Immunol Immunother*. 2011 Jul;60(7):975-84. Epub 2011 Apr 2. PMID: 21461842
20. Peters D, **Chu T**, Yatsenko S, Hendrix N, Hogge WA, Surti U, Bunce K, Dunkel M, Shaw P, Rajkovic A. Non-Invasive Prenatal Diagnosis of a Fetal Microdeletion Syndrome. *NEJM*. Accepted
21. Oh SY, **Chu T**, Sadovsky Y. The timing and duration of hypoxia cluster changes gene expression in cultured primary human trophoblasts. *Placenta*. 2011 Oct 8. Epub
22. Hermann BP, Sukhwani M, Salati J, Sheng Y, **Chu T**, Orwig KE. Separating spermatogonia from cancer cells in contaminated prepubertal primate testis cell suspensions. *Human Reproduction*. 2011 Oct 20. Epub
23. Suzuki H, Ahn HW**, Chu T**, Bowden W, Gassei K, Orwig K, Rajkovic A. SOHLH1 and SOHLH2 coordinate spermatogonial differentiation. *Developmental Biology*, 2011 Oct 26. . Epub
24. Bunce K, **Chu T**, Surti U, Hogge WA, Peters DG. Discovery of epigenetic biomarkers for the non-invasive diagnosis of fetal disease. *Prenat Diagn*. 2012 Jan 9.
25. Himes KP, Handley D, **Chu T**, Burke B, Bunce K, Simhan HN, Peters DG. Comprehensive analysis of the transcriptional response of human decidual cells to lipopolysaccharide stimulation. *J Reprod Immunol*. 2012 Jan;93(1):17-27
26. Donker RB, Mouillet JF, **Chu T**, Hubel CA, Stolz DB, Morelli AE, Sadovsky Y. The expression profile of C19MC microRNAs in primary human trophoblast cells and exosomes. *Mol Hum Reprod*. 2012 Feb 29
27. Shalom-Barak T, Zhang X, Chu T, Schaiff WT, Reddy JK, Xu J, Sadovsky Y, Barak Y. Placental PPARgamma Regulates Spatiotemporally Diverse Genes And A Unique Metabolic Network. *Developmental Biology* [accepted]

C. Research support

Ongoing

NIH (PI: Donald Scott) 08/15/11-05/31/13
Epigenomic regulation of gene expression in diet induced obesity
NIH (PI: David Peters) 05/15/11-01/31/16
Non-invasive detection of fetal aneuploidy by next-generation DNA sequencing
NIH (PI: Yoel Sadovsky) 08/05/10-05/31/15
Placenta Injury and microRNA