

## Genetic Differences Between Ethnicities May Determine Breast Cancer Prognosis

The University of Miami's Lisa Baumbach and Maastricht University's Torik Ayoubi discuss how ethnicity-specific genetic changes may govern the risk of developing an aggressive form of breast cancer

By Stacey Ryder

**MIAMI, July 14, 2008**—Researchers at the University of Miami, led by Dr. Lisa Baumbach, have discovered hundreds of genetic differences between African-American, Hispanic and Caucasian breast cancer patients that may account for an increased incidence of a particularly aggressive form of the cancer in

African-American women. The discovery may pave the way for improved breast cancer treatments and prognoses based on ethnicity.

Baumbach and her team used the Almac Breast Cancer DSA™ research tool for their study. The Almac microarray, based on Affymetrix GeneChip®

technology, includes 60,000 transcripts and is designed specifically to investigate gene expression in breast cancer. Baumbach's team identified 600 transcripts which are differentially expressed among the eight Caucasian, 10 African-American and 10 Hispanic women in its initial sample.

A unique aspect of this study is that the team used formalin-fixed, paraffin-embedded (FFPE) tissue samples. In the past, FFPE tissues have been difficult to use because RNA extracted from these samples is usually degraded, but new technology has enabled researchers to obtain high-quality expression data from these tissues.

“Using paraffin-embedded tissue has opened up an unbelievable gold mine, not only for me, but for many other researchers in the field,” said Baumbach, an associate professor at the University of Miami.

Baumbach's group found that breast tissue samples from different ethnicities include groups of differentially expressed genes. Gene expression in breast tissue from African-American women differs from that in Caucasian and Hispanic women, just as gene expression in Hispanic women differs from both African-American and Caucasian women.

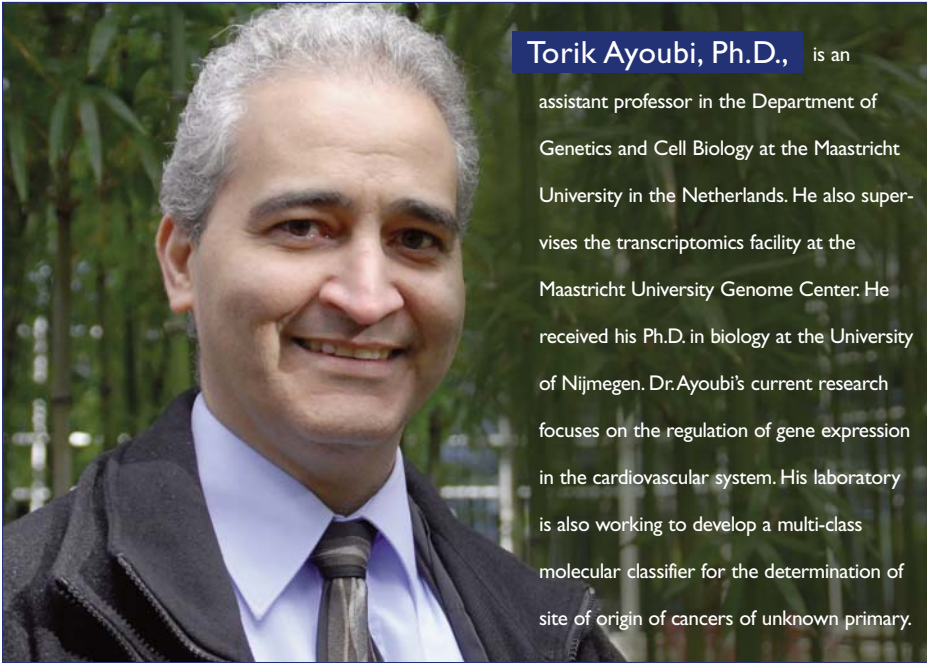
In their latest study, Baumbach and her team focused on women with the “triple-negative” form of breast cancer. These women are negative for the genes for estrogen receptor (ER), progesterone



**Lisa Baumbach, Ph.D.,**

is an associate professor in the Departments of Pediatrics and

Genetics at the University of Miami. She received her Ph.D. in biochemistry from the University of Florida. Baumbach's lab took over a project focusing on the genetic basis of African-American breast cancer in 2000. Recent studies from her laboratory focus on gene expression profiling and understanding potential genetic differences between breast cancer in African-American women and breast cancer in women of other ethnicities.



**Torik Ayoubi, Ph.D.**, is an assistant professor in the Department of Genetics and Cell Biology at the Maastricht University in the Netherlands. He also supervises the transcriptomics facility at the Maastricht University Genome Center. He received his Ph.D. in biology at the University of Nijmegen. Dr. Ayoubi's current research focuses on the regulation of gene expression in the cardiovascular system. His laboratory is also working to develop a multi-class molecular classifier for the determination of site of origin of cancers of unknown primary.

receptor (PR) and Her2/Neu, an epidermal growth factor receptor. This genetic combination suggests a particularly poor prognosis, because this form of cancer is unresponsive to the most effective receptor-targeted treatments.

Baumbach recently spoke with Dr. Torik Ayoubi, an assistant professor of genetics and cell biology at Maastricht University in the Netherlands, about her collaboration with Almac and the results her team achieved using FFPE tissues. The two discussed:

- The importance of examining differences in gene expression in breast cancer samples from different ethnicities.
- The value of being able to use FFPE samples on cancer diagnostic arrays.
- The current use of the Affymetrix® Genome-Wide Human SNP Array 6.0 to examine copy number changes, as well as future plans to expand the study and validate the current results.

#### Study design

**Ayoubi:** What was the rationale for your study, and what is the significance of identifying multiethnic differences in gene expression in breast cancer samples?

**Baumbach:** There is a lot of epidemiological evidence that African Americans may have a different type of

breast cancer than other ethnicities. They have a much more aggressive, earlier onset of breast cancer, even when you account for socioeconomic differences and other such factors. They also present at a later stage.

The vast majority of the breast cancers in African-American women occur in women less than 50 years of age, and there are many cases where the women are less than 30. Breast cancer is the leading cause of death in African-American women and its highest incidence is before age 50. Recent data has also shown that the basal-like phenotype of breast cancer is more common in African Americans.

All of this makes us think there might be gene expression differences among women of different ethnicities. We have done a number of studies looking at BRCA1 and BRCA2 mutations in African Americans. We found that they had fewer occurrences of deleterious germline mutations, but a high number of missense mutations. So, we decided that instead of trying to find additional genes, we would use genome expression profiling to identify transcriptional differences.

Recently, we added Hispanics to the study. Although not as much is known about breast cancer in Hispanics, there is early evidence that there may again be a difference in the age and stage of presentation.

**Ayoubi:** Your study focused on eight

to 10 individuals from each of the three ethnic groups. The “curse of multidimensionality” of microarray data predicts many false positives when performing tens of thousands of tests simultaneously. How do you deal with the multiple testing problem in such a complex context like ethnicity?

**Baumbach:** First of all, the key is the number of samples. Again, in this study we focused only on triple negatives to reduce at least some of the confounding variables.

Second, all the data is rigorously scrutinized. Although we get about 26,000 transcripts from the microarray, they go through several statistical analyses including a quality control analysis for false positives, a two-way ANOVA and a series of other statistical filters.

So, when we look at the differentially expressed transcripts, the numbers are much lower—about 6,000. We recently culled the list down to a data set of 600 genes in African Americans, Hispanics and Caucasians, which are differentially expressed between tumor and normal tissues within the three groups. We see some major differences.

For example, we recently found that ER alpha was reduced about five- to six-fold in African Americans and Hispanics versus Caucasians. Even in the ER-negative group, there are differences. This was substantiated by other work that was done by a group in the Netherlands. When they looked at the basal-like phenotype of breast cancer, they found there were five subclasses even within this single phenotype. It is very possible that we will also find these kinds of differences at the transcript level.

#### Almac's Breast Cancer DSA

**Ayoubi:** Why was it important for you to be able to use the FFPE samples?

**Baumbach:** It is well documented, as well as has been our experience, that many African-American breast cancer patients are seen at a later stage of presentation, so they almost all go straight to adjuvant therapy or chemotherapy. As a result, it is very difficult to acquire untreated frozen tissues. Being able to use the paraffin-embedded tissues from biopsies

or lumpectomies allowed us to look at primary tumor tissue that has not been exposed to radiation or chemotherapy.

**Ayoubi:** How does Almac's Breast Cancer DSA compare to other technologies available for this kind of research?

**Baumbach:** Almac started their business with the goal of developing disease-specific arrays that could be used primarily for paraffin-embedded tissue. For the Breast Cancer DSA, they completely sequenced about 400 samples of normal breast and tumor samples at the RNA level. From these sequences, they identified about 60,000 transcripts that are represented on the breast cancer microarray. This was done with heavy bias toward the 3' end of the RNA, in order to obtain coverage of even the small RNAs that might be partly degraded in the paraffin-embedded tissue.

A large number of the transcripts are unique RNAs that were previously undiscovered. The Breast Cancer DSA covers not only a number of the genes that are on the standard Affymetrix arrays, but also additional RNA sequences that are specific to breast cancer. Of the 60,000 transcripts, we read about 26,000 transcripts per sample from the microarray.

Using paraffin-embedded tissue has opened up an unbelievable gold mine not only for me, but for many other researchers in the field.

### Study results

**Ayoubi:** What steps did you take to ensure reproducibility in the results?

**Baumbach:** Working with Almac has helped us to avoid many of the problems we might have encountered on our own, because they have really optimized this technology.

In order to obtain quality preliminary data, we had Almac perform all of the preliminary RNA extraction and hybridization. We were very particular about how the FFPE sections were cut by our local pathologists and how these sections were shipped to Almac for further processing. In our experience, the best results with RNA extraction and hybridization are obtained by having only a 24-hour window between the time when the FFPE sections are cut and

Almac receives them and starts to process the sections for RNA extraction. We routinely send the FFPE sections on dry ice overnight to Almac. Adhering to this protocol has yielded consistently high-quality, high-quantity RNA yields.

We also try to process as many samples as possible at the same time. We have about 28 patients in our study. So, we send blocks of samples from 10 to 15 people.

**“Using paraffin-embedded tissue has opened up an unbelievable gold mine not only for me, but for many other researchers in the field.”**

Almac runs the samples in duplicate, but we try to include some overlapping samples so there is less of a chance that samples will be analyzed on different days.

We also have very specific questions about triple negatives, so we're going back to the literature and looking at the gene expression profiles that have been reported for triple-negative patients. We include those as positive controls in the validation studies.

**Ayoubi:** Did the dramatic differences between the profiles of Caucasian and African-American women with breast cancer surprise you?

**Baumbach:** We were very surprised to see the differences in normal tissue. If those are true normal differences, they are probably some of the most exciting results of the work. It is interesting to see very specific changes among these three ethnic groups. Some differences are specific to African Americans. They are not found in Hispanics or Caucasians. It is not necessarily surprising, but it is exciting because it may suggest that we can develop ethnicity-specific targeted therapies for these cancers.

**Ayoubi:** When you look in more detail at the genes that differ between ethnic groups, do you recognize particular pathways that can tell you something about carcinogenesis?

**Baumbach:** We are just starting work with the pathway analysis. We have already seen big differences in expression

levels of selected genes in both cell cycle and DNA repair pathways across ethnicities in tumor samples. For example, one of the markers for the basal-phenotype—the keratins—are also differentially regulated, with tremendous up-regulation in the breast tumors from the African-American patients only. Obviously, these results need to be substantiated by real-time PCR, but in the meantime, we are

trying to integrate this information into pathways that make a biological story. Within a month's time or so, we will have a much more complete understanding of the pathways involved.

### Moving forward

**Ayoubi:** What are your plans for future studies? Have you performed any kind of validation in an independent larger group of samples?

**Baumbach:** That is our goal. First, we are using real-time PCR to validate some of the biggest differences. We are doing copy number studies on the Affymetrix SNP Array 6.0 in parallel to examine how copy number changes relate to the expression differences. We want to increase the sample sizes to 30 or 35 samples per group. The grant we have pending with the Komen Foundation will fund that validation.

In addition, we are starting another study where we are obtaining paraffin-embedded samples from Kenya. Then we will be able to directly compare African samples to African-American samples. It will be very interesting to see what changes are shared between the two groups. We can then begin to sort out the true ethnic differences from the differences based on environmental influence.

**Ayoubi:** There are two main types of genetic polymorphisms that could explain the ethnic differences. One of them would be so-called cis-acting poly-

morphisms. That means that there could be SNPs in the genes themselves that cause different expression. The second type would be trans-acting polymorphisms, meaning that some transcription factors that regulate these genes might have polymorphisms. Do you favor either of these mechanisms? Do you have plans to tackle them using an integrated approach, combining expression and SNP analysis, in the future?

**Baumbach:** We are doing that right now. We are extracting DNA from the same samples that were examined for RNA expression. We are running the copy number variation experiment on the Affymetrix SNP Array 6.0. That data is going to be very important when we integrate the results of the different studies. It will help us understand whether the transcriptional differences are due to copy number variation in those genes or the result of epigenetic regulation at the promoter level.

I am not favoring either explanation right now. I want to see what the results show. Integrating the data will be quite a challenge, but it is important. We will not be publishing our data until we have both data sets complete.

**Ayoubi:** How would you proceed to identify a causal role for these genes in the etiology of breast cancer or in the way they might affect the response to particular treatment regimes?

**Baumbach:** We need to complete the pathway analysis. Actual biological pathways are crucial to this story. It is not enough to have genes that are differentially regulated. That is what we are focusing on.

Once we finish the pathway analysis, we will proceed to investigate other areas, such as proteomics. We will design an in vitro model in order to validate our findings.

We are at the beginning of what could be a very important and very long-term study. But we will have to go back to the basic biology of the cell, and maybe even to primary breast tissue, to answer some of these questions.

### Clinical and diagnostic implications

**Ayoubi:** Do you think your findings can be extrapolated to other types of cancer?

**Baumbach:** That is a very good question. Prostate cancer is at almost-epidemic proportions in African-American men and it is the leading cause of death for them. Maybe there is also a genetic difference in the basis of prostate cancer in African-American men versus Caucasian men. If our work is substantiated, it will show for the very first time that there are genetic differences in cancers between different ethnic groups. That may set the stage for the exploration of ethnicity-specific differences in other cancers.

# AFFYMETRIX MICROARRAY BULLETIN

## Editorial Staff

Rachel Steger, *Founding Editor*

rachel\_steger@affymetrix.com

Bronwyn Barnett, *Managing Editor*

bronwyn\_barnett@affymetrix.com

Sven Brown, *Senior Editor*

Stephen Antonopoulos, *Photo Editor*

Stacey Ryder, *Associate Editor*

Sabrina Banwait, *Associate Editor*

Marva Maida, *Graphic Designer*

Jim Butler, *Graphic Designer*

**Ayoubi:** What might be the direct clinical or diagnostic implications of this work?

**Baumbach:** I think if we can validate these potential gene signatures in the different ethnicities, we might find some new therapeutic avenues for the triple-negative breast cancer. That would be very important because it is largely unresponsive to the current therapies.

As we collect additional samples, we will also be correlating them with therapeutic outcomes. That may help us develop an ethnic-specific prognostic signature—this is one of our ultimate goals.

## FOR MORE INFORMATION

### Contacts

- Lisa L. Baumbach-Reardon, Ph.D.

Associate Professor

Miller School of Medicine

University of Miami

6021 Mailman Center for Child Development

1601 N.W. 12th Avenue

Miami, FL 33136

Email: lbaumbach@med.miami.edu

- Torik Ayoubi, Ph.D.

Assistant Professor

Genome Center Maastricht

Department of Genetics and Cell Biology,

section of Clinical Genomics

Maastricht University

Universiteitssingel 50,

P.O. Box 616, #16

6200 MD, Maastricht, The Netherlands

Email: torik.ayoubi@gen.unimaas.nl

website: www.gen.unimaas.nl/popgen

### Companies

- Affymetrix, Inc. – www.affymetrix.com

- Almac Group, Diagnostics –

www.almacgroup.com/diagnostics/

### Organizations

- University of Miami – www.miami.edu

- University of Maastricht – www.unimaas.nl/

- Komen Foundation – www.komen.org

### Further Reading

- Olopade O., *et al.* Breast Cancer Genetics in African-Americans. *Cancer Supplement* 97:236-245 (2003).

- Baumbach L., *et al.* Multiethnic Comparisons of Genome-wide Alterations in Breast Cancer Using Paraffin-Embedded Samples. *American Society of Human Genetics Annual Meeting Abstract* #442 (2007).

- Hosey AL., *et al.* A molecular basis for Estrogen Receptor- $\alpha$  deficiency in BRCA1-linked breast cancer. *Journal of the National Cancer Institute* 99(22):1683-94 (2007).

- Jacobs S., *et al.* Genome-wide, high-resolution detection of copy number, loss of heterozygosity and genotypes from formalin-fixed, paraffin-embedded tumor tissue using microarrays. *Cancer Research* 67(6):2544-51 (2007).