Genomic Fingerprinting Estrogen Receptor-Negative Breast Cancer
Project Update October 2003

Introduction

The purpose of Project 2 is to provide Center of Excellence investigators with a comprehensive expression profile of estrogen receptor (ER)-negative breast cancers by performing expression array analysis of RNA extracted from patient tumor samples. The expression array platform chosen for this study is the Affymetrix U133 GeneChip, providing expression information on more than 14, 500 annotated human genes, and 33,000 human transcripts. After performing the array analysis, expression data sets from these tumors will be posted on the Center of Excellence web site, and made available to Center investigators. During the funding cycle of this grant, the entire data set will be made publicly available to the scientific community.

In secondary aims, Project 2 will work with Center investigators to explore expression signatures of activated growth pathways in ER-negative cancer. The strategies will be varied, but the principle is to search for signatures, "fingerprints", of activated pathways by deriving gene lists from model systems in which pathways can be experimentally activated. The fingerprints obtained from array analysis of human tissues will be compared to in-situ experiments looking at markers of pathway activation (phosphorylation of substrates, over expression of critical proteins, etc). In this way, global expression profiles will be correlated to more directed immunophenotyping of our clinical samples.

The goal in this work is to provide a more detailed classification of ER-negative breast cancer and to molecularly identify homogeneous groups of tumors within the larger group of receptor-negative cancers. We hope by resolving the heterogeneity if ER-negative breast cancers, identifying homogenous groups of tumors and determining the cognate molecular pathways for these groups, the stage will be set for tailored and more effective treatments.

1. Provide an expression profile for ER-negative cancers using Affymetrix U133 GeneChips.

This work was done with cancer specimens collected by the SPORE in Breast Cancer at the Brigham and Women's Hospital, using funds from the SPORE and from private sources within the Women's Cancers Program of the Dana-Farber Cancer Institute. Affymetrix U133 GeneChips have been acquired for this work using private funds from the Women's Cancers Program and from the Brigham and Women's Hospital. The Department of Defense Center of Excellence provided labor and other supplies to accomplish this goal.

The frozen tissue repository in the Breast Cancer SPORE was searched and a cohort of ER-negative cancers identified. In total, 54 new receptor-negative frozen cancers were found suitable for extraction of RNA and profiling. Adding to these new

cancers, we will re-hybridize RNA from 19 previously analyzed receptor-negative cancers (analyzed on U95A arrays). Therefore, a total of 73 ER-negative cancer specimens will be arrayed on U133 arrays. The following table describes the new group of 54 specimens in more detail (Table 1).

Table 1. Characteristics of 54 ER-negative cancers

Characteristic	Number	Percent of Group
Histologic Grade		
III	48	88%
II	6	
HER2 Status		
Positive	14	26%
Negative	36	67%
Unknown	4	
T Stage		
T1b	5	9%
T1c	21	39%
T2	27	50%
T3	1	
Lymph Node Status		
Positive	18	33%
Negative	32	59%
Unknown	4	
Age		
<40	6	11%
41-50	20	37%
51-85	28	52%

RNA has been extracted from 35 of these 54 tumors and in every case more than 5_g of high-quality RNA is available. Affymetrix U133 A and B chips have been purchased and are waiting in the Affymetrix Microarray Core at the Dana-Farber Cancer Institute for hybridization. We expect this portion of the Project to be completed by February or March 2004. DNA for HuSNP (single nucleotide polymorphism genotyping arrays from Affymetrix) will be prepared from these same tumors. In this case, it is necessary to micro dissect tumor elements from each section of cancer, and to have a source of normal germ line DNA available. For the normal DNA, we will rely on either peripheral blood lymphocytes, obtained in the Breast SPORE Tissue Repository, or on micro dissected lymph nodes from either the sentinel node biopsy or the axillary dissection. Finally, paraffin blocks from these tumors will be pulled and Tissue Microarrays (TMAs) prepared.

Data from the SNP arrays will be analyzed for Loss of Heterozygosity (LOH) at informative loci. We will use the DNA Microchip Analyzer software (*dChip*) written for Affymetrix expression and SNP arrays by Dr. Wing Wong. d*Chip* software is open-source and available at www.dchip.org. Genotype data will be posted on the Center of

Excellence web site for Center investigators, and will become openly available to all researchers during the current funding cycle of our Center.

2. Classify ER-negative cancers into discrete groups, based upon expression profiles, growth factor pathways, or other genomic data.

During the first year of this Center, we have completed expression analysis in a cohort of 89 cancers, 19 of which were ER-negative. Unsupervised clustering of global expression profiling, using different array technologies, produces two groups of breast cancers in first-degree separation. These are predominantly ER-positive and ER-negative breast cancers. In our work, we reproduced this pattern of clustering, and identified two groups of ER-negative cancers in the second-degree of separation after unsupervised clustering. These two groups of ER-negative cancers were composed of predominantly HER2-positive tumors in the first group (the HER2 group) and a second group of HER2-negative, ER-negative cancers. This second group has been found by others, and termed the "basal-like" group (1,2). It is a high-grade disease, HER2 and ER-negative, and frequently displays highly expressed p53 (usually indicative of a p53 gene mutation).

We micro-dissected tumor and normal cells from 42 cancers in the group of 89 used for expression arrays. Affymetrix 1.5K SNP arrays were used with modified *dChip* software to determine LOH events. In addition, we calculated the frequency of LOH events as a proportion of informative genetic loci on the chip, and expressed this value as FLOH (frequency of LOH). FLOH may be a measure of genetic instability and illegitimate mitotic recombination. A manuscript has been submitted (see below) and will be published in Cancer Research next year.

This work has identified frequent and relatively unique LOH events in the "basal-like" group of cancers, identified by the expression arrays. Interestingly, the HER2 group of cancers did not harbor chromosomal LOH events that were unique or significantly more common than seen in the other groups. Furthermore, the HER2 group had FLOH values not dissimilar than those seen in high-grade ER-positive cancers. The following Table displays the characteristics of the expression array-determined groups, and shows the frequency of LOH events in each group.

Table 1. Summary for tumor grades, ER/HER-2 status, P53 mutations, and the levels of LOH in 34 tumors allotyped by SNP array analysis and subclassified by expression clusters in a larger cohort of 89 tumors.

	Cluster I		**Cluster II**	
Subgroups	IA (n=15)	IB (n=7)	IIA (n=4)	IIB (n=8)
Tumor Grade				
I	20[a]	0	25	25
II	27	0	25	13
III	53	100	50	62
ER+	53	0	100	100
HER2+	60	0	0	0
P53	47	57	0	0
Mean FLOH [b]	14	37	6	14

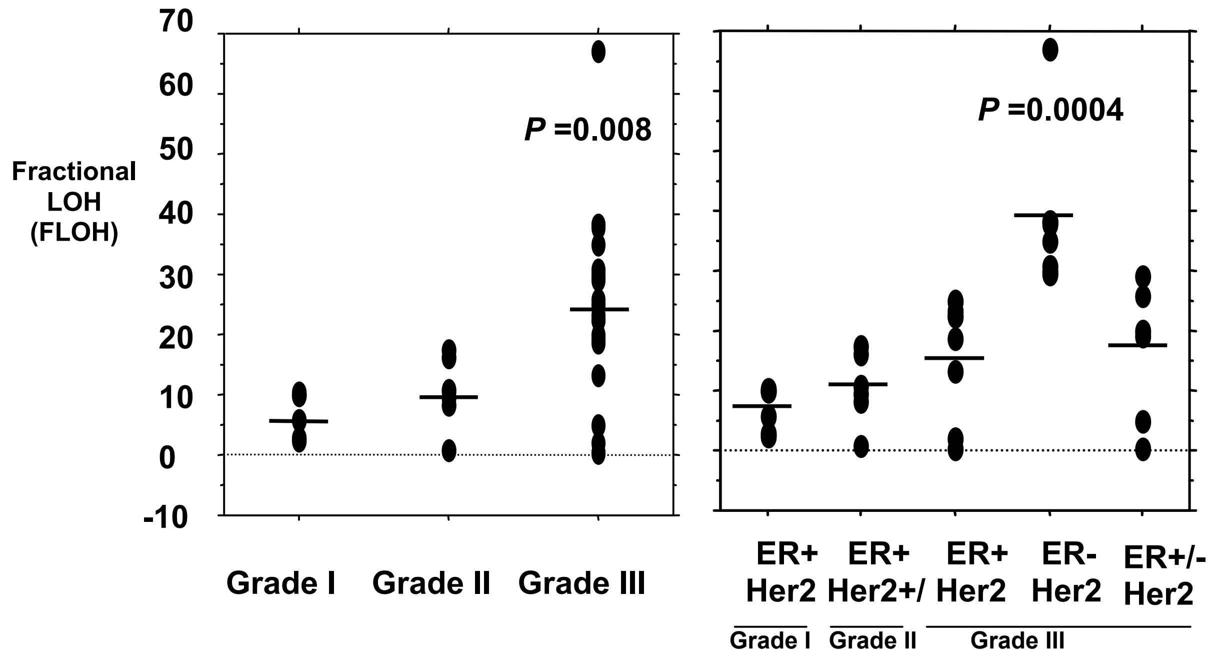
Cluster I are the ER-negative cancers; sub-cluster IA are the HER2-positive cancers and sub-cluster IIB are the “basal-like” cancers.

[a] Percentage of tumors

[b] Fraction (%) of LOH in all informative (heterozygous) loci for each tumor.

FLOH values tend to increase with increasing grade of breast cancer (shown in Figure 1, below). However, breaking the groups down further reveals the highest level of FLOH in the group of cancers that are both HER2-negative and ER-negative. We hypothesize that a major division of ER-negative cancers may be between those that are genetically unstable, with high levels of mitotic recombination and LOH events, and those that are caused by growth factor pathway abnormalities, such as the HER2/neu-driven cancers. This hypothesis will be tested when more cancers have been arrayed by expression and SNP analysis.

Figure 1. Frequency of LOH in classes of breast cancers



3. Subdivide ER-negative cancers into more homogeneous groups using expression profiles and immunophenotypes indicative of particular growth factor pathways.

If we are correct, and ER-negative cancers can be subdivided by illegitimate activation of growth pathways, and perhaps abrogation of apoptotic pathways, then our task is to identify those pathways. A variety of strategies can be used to approach this goal. One way is to derive signatures of pathway activation using experimental, in-vitro systems, derive gene lists associated with the activation of a particular pathway, and use this list to "probe" an expression database from clinical specimens. Another way is to investigate ER-negative cancers using antibodies directed against phosphorylated intermediates in growth pathways as markers of pathway activation. We intend to take both approaches in this Center of Excellence.

To accomplish this, we must first gather our cohort of ER-negative cancers, and perform the expression profiling described in Goal #1. Also, tissue microarrays need to be constructed from these same cancers for the immunohistochemical approach. This work is in progress, and we will report progress in later communications.

Research Accomplishments in the Past Year

Preliminary work has subdivided ER-negative cancers into those that are HER2-positive and into a group of genetically unstable cancers, termed "basal-like" tumors.
Signature chromosomal loss associates with the basal-like group, and may be a useful marker of this sub-group
Seventy-three ER-negative breast cancers have been identified, and we are nearly through with RNA extraction from all of these tumors. Hybridization to Affymetrix expression arrays will be done soon.
Construction of tissue microarrays and DNA extraction from these tumors will start shortly.

Manuscript submission: Zhigang C. Wang, Ming Lin, Lee-Jen Wei, Cheng Li, Alexander Miron, Gabriella Lodeiro, Lyndsay Harris, Sridhar Ramaswamy, David M. Tanenbaum, Matthew Meyerson, James D. Iglehart, Andrea Richardson. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Research*, accepted, 2004 publication

Research Opportunities/Additional Funding: This work represents a collaboration of multiple individuals, and is funded by multiple sources. As noted earlier, the tissue collection was accomplished by funding from the SPORE in Breast Cancer at Harvard (NCI) and by private sources (The Women's Cancers Program at the Dana-Farber and sources at the Brigham and Women's Hospital). The microarrays (expression and SNP arrays) will be purchased using private funding at the Brigham and Women's Hospital.

The taxonomy of breast cancer, both ER-negative and ER-positive cancer, is central to tailored approaches to diagnosis and therapy. Furthermore, proper identification of cancer phenotypes is crucial to gene discovery. Therefore, this research on ER-negative cancer is central to several Programs in the Dana-Farber/Harvard Cancer Center. These include the SPORE in Breast Cancer (NCI), the Pre-Operative Therapies Program (Breast Cancer Research Foundation), Classification of Breast Cancer using Single Nucleotide Polymorphisms (Komen Foundation, pending), Analyzing the Genome, Transcriptome and Proteome of Breast Cancer (Breast Cancer Research Foundation), and others.

Conclusions

Central to the work in Project 2 is the hypothesis that breast cancer is a heterogeneous disease, and different cancers are in fact different diseases. Proper identification of the many diseases we call "breast cancer" is critical to choice of therapy, evaluation of diagnostics, and to the basic science of gene discovery and signaling

pathway identification. By lumping all these varied diseases under one label, we are impeding progress on a number of fronts. In this Project, we are concentrating on subdividing ER-negative breast cancers into their component "diseases" and on discovering clinically useful markers of these different diseases.

References

1. Perou C, et al. Molecular portraits of breast cancer. Nature 406:747-752, 2000.
2. Sorlie T, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci, USA 100:8418-8423, 2003.