

# #3008 VALIDATION OF MULTIPLEX GENOMIC MARKERS FOR PREDICTING BREAST CANCER RECURRENCE IN A FISH ASSAY FORMAT.

Cole Harris, MA 1, Lisa Davis, PhD 1, Patti Doherty, RN 1, Peter Hrabec, PhD 1, Lei Tang, PhD 1, Therese Bocklage, MD 2, Katie Doeden, MD 2, Bryan Hall, MS 2, Ian Rabinowitz, MD 2, Thomas Williams, MD 2 and John Hozier, PhD 2. 1 Exagen Diagnostics, Inc., Albuquerque, New Mexico, United States, and 2University of New Mexico, Albuquerque, New Mexico.

## Introduction

Predicting risk of recurrence in breast cancer patients is currently limited, resulting in the possibility of unnecessary adjuvant chemotherapy for some women and difficulty identifying those who could benefit from more aggressive treatment. With the aim of developing a clinically useful and practical prognostic test, we considered potential test implementation technologies as well as the availability of data sources for marker discovery. Fluorescent in situ hybridization (FISH) based assays are in wide use in cancer diagnostics, thus the introduction of a new prognostic FISH assay would likely be quickly adopted.

## Background

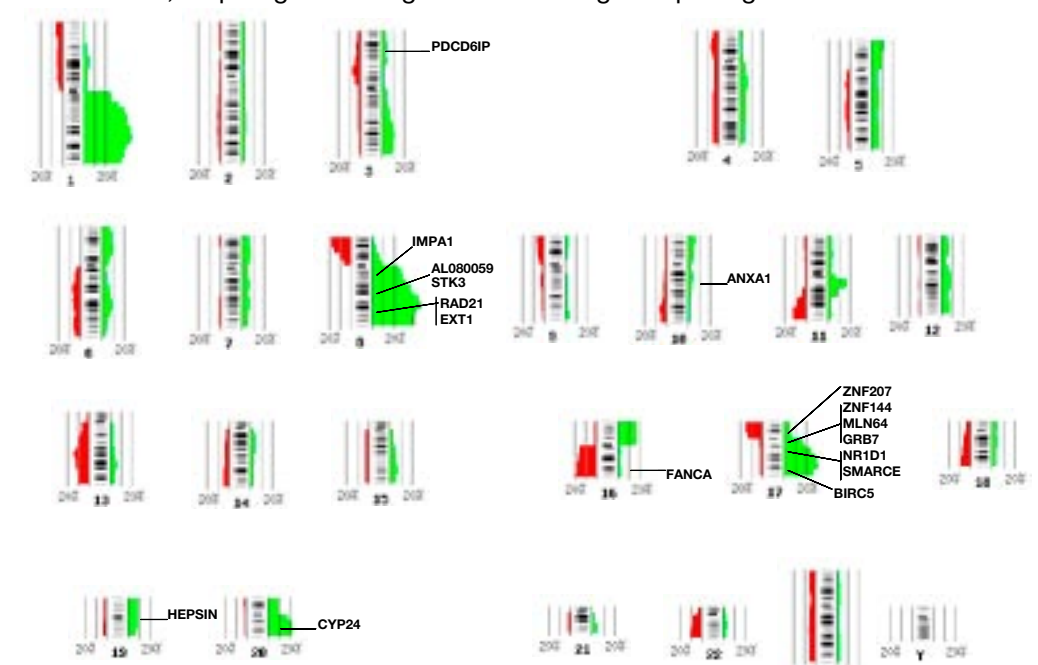
Towards this end, 16 genes that are prognostic for breast cancer recurrence were discovered by concurrent global mining of genome-wide RNA microarray data and high resolution CGH microarray data. A 17th gene that was differentially prognostic among hormone positive and negative tumors was also discovered. (See Harris, SABCS 2004 Poster #3007, for details on the discovery of the genes used in this study). Measuring DNA copy numbers of these genes in discrete genome regions offers the potential for development of prognostic assays, using a procedure we refer to as "patterns of genomic amplification-FISH" or PGA-FISH™.

## Objective

For the current study, our objectives were two fold: (1) to validate the prognostic power of the 17 genes identified in the global search described above, and (2) to discover new markers comprised of subsets of those 17 genes, that have greater prognostic power than the individual 17 genes alone. Towards this end, we measured the DNA copy number of each of the 17 genes in a series of archived breast cancers, using FISH as the assay. One of the genes was measured in duplicate. We then identified patterns of DNA copy numbers that correlate with recurrence from among those 17 genes.

The locations of the 17 genomic regions assessed in this study are shown in the panel below, superimposed on a genome ideogram depicting chromosome regions commonly gained (shown in green) and lost (shown in red) in invasive ductal carcinomas. The regions of chromosome gains and losses shown are collated from published CGH data on breast tumors (www.progenetix.com).

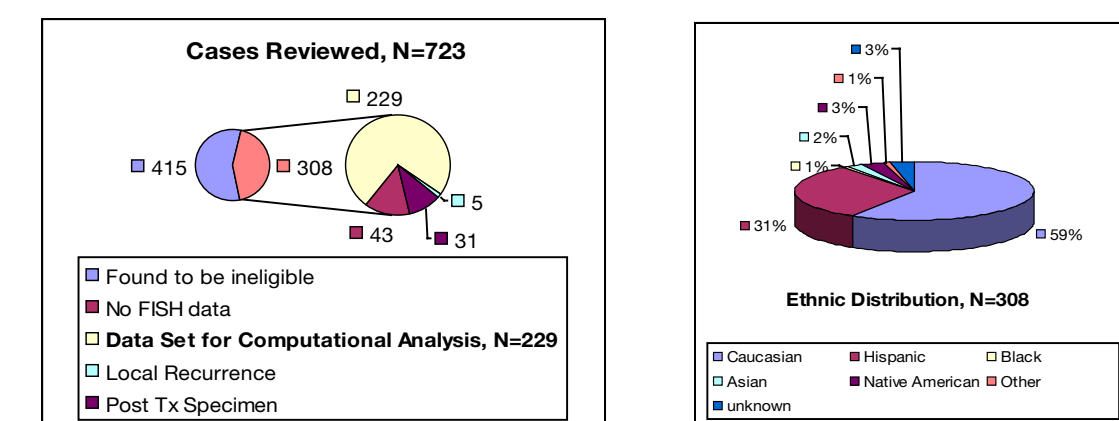
Genes identified as predictors of recurrence by global mining of expression and CGH data, map to genome regions that undergo frequent gains and losses.



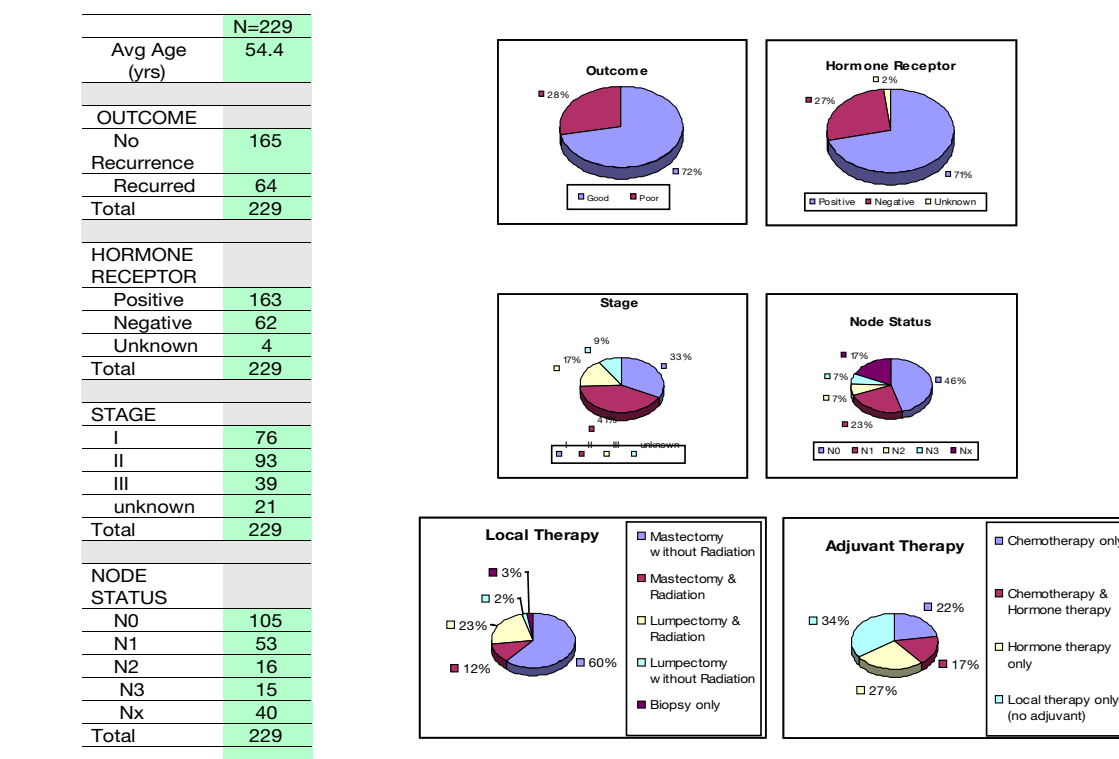
Progenetic M. Baada  
www.progenetix.com

## Study Population

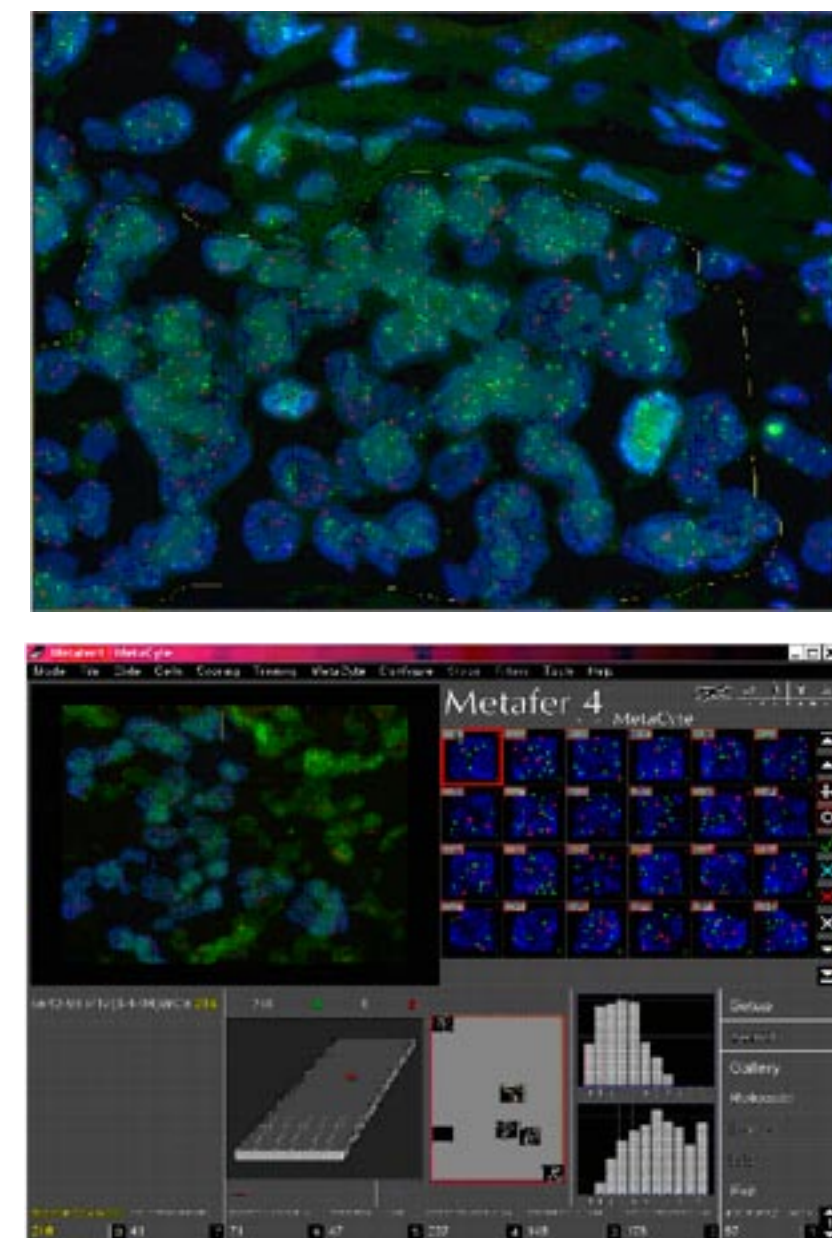
Medical records from 723 cases of breast cancer diagnosed at the University of New Mexico Hospital between 1986 and 1999 were reviewed and 308 eligible cases were identified. Eligibility criteria included a diagnosis of Stage I, II or III invasive ductal carcinoma, availability of original clinical and pathological records, availability of archived tumor specimen, and at least four years of clinical follow-up. Mean follow-up was 8.9 years. Recurrence was defined as clinical evidence of metastasis or death from breast cancer. The study population consisted of 98 Stage I (33%), 118 Stage II(38%), and 61 Stage III (20%) patients. Treatments included mastectomy (59%), mastectomy/radiation (14%), lumpectomy (4%) and lumpectomy/radiation (21%). (Six patients underwent biopsy only biopsy only.) Patients received adjuvant hormone treatment (25%), chemotherapy (29%), chemotherapy and hormone treatment (15%) or no adjuvant treatment (22%). Tumors were classified as hormone receptor positive (HR+) if positive for either (or both) estrogen or progesterone receptors. Two hundred and nine (209) cases (68%) were HR+ and 83 (27%) were HR-. (Hormone status was undetermined in 16 cases, 5%). The disposition of the 723 reviewed cases is shown on the left and the ethnic makeup of the 308 eligible cases is shown on the right.



Outcome, Hormone Status, Stage, Node Status, and Therapy  
N= 229

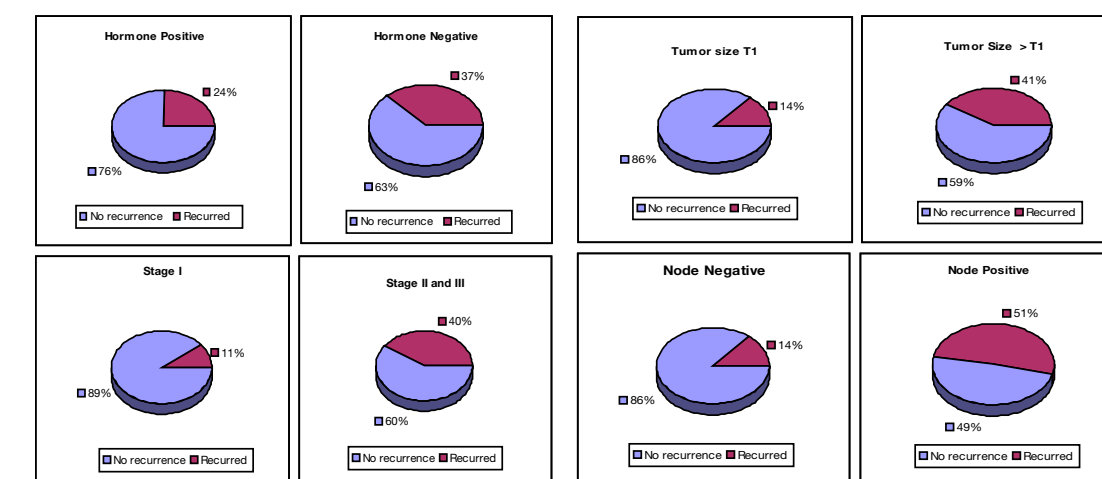


## Study Methods – FISH Data Collection Virtual Microdissection



Pathology reports and microscopic slides were reviewed and diagnostic blocks selected for sectioning for FISH studies. FISH data were collected on 265 cases. The hybridizations were performed on serial sections and the personnel performing the hybridizations and signal collection were blinded to both the probe identities and clinical histories. The genomic probes used in this study were selected from the human "32K" BAC Re-Array (Children's Hospital Oakland Research Institute, http://bacpac.chori.org/). BAC clone inserts were verified by PCR and the chromosome localization verified by FISH hybridization to mitotic chromosomes. The 17 BAC probes were hybridized in pairs (labeled with Spectrum Red or Spectrum Green fluorochrome), and hybridizations followed standard protocols for sample deparaffinization, hybridization, and wash. Hybridization signals were collected, imaged, stored, and analyzed using MetaCyte automated FISH analytical hardware and software (MetaSystems, Altusscheim, Germany). Generally, twenty 40x fields of view with good signal quality were operator-selected for each probe pair, and then captured in Metacyte. Regions not containing carcinoma cells (e.g. stromal or infiltrating lymphocytic cells) were excluded from further analysis using a process we refer to as "Virtual Microdissection" (left panel). Typically FISH signals from at least one hundred cells (usually 200) from at least 2 fields of view (usually 5-6) were analyzed for each gene probe pair. A few specimens (>5%) yielded smaller numbers of analyzable cells. Signal count data were collected in a "tiling" pattern designed by MetaSystems to minimize the effects of non-uniform distribution of nuclei in thin sections. On the right, panels a) and b) demonstrate the relationship between the field of view and the tiles. The right figure also demonstrate the relationship between the microscope slide (c) and the selected fields of view with their tiling pathways (d). Data are reported as a histogram of the numbers of signals per tile (e). Data were then transferred to Exagen in flat file format for computational analysis.

	N=229	Avg. age (yrs)	Hormone Status		Stage		pT		Node Status		Ethnic Background		
			HR+	HR-	I	II and III	T1	T>1	N0	N>0	Caucasian	Not Cauc	Hispanic
No Recurrence	165	56.5	121	41	68	79	87	66	90	41	99	12	47
Recurred	64	50.9	39	24	8	53	14	45	15	43	39	4	21



- Factors found to be significant predictors of recurrence include: younger age ( $P = 4.3 \times 10^{-4}$ ), pathological stage greater than pT1 ( $P = 1.0 \times 10^{-5}$ ), overall stage greater than I ( $P = 6.0 \times 10^{-6}$ ) and node status greater than N0 ( $P = 4.2 \times 10^{-8}$ ).
- Neither hormone status ( $P = 0.3208$ ) nor ethnic origin (Caucasian vs. other, Hispanic vs. others, and Non-Caucasian vs others,  $P = > 0.75$ ) were found to be significant predictors of recurrence.
- Among the node negative group, neither age ( $P = 0.4485$ ) nor tumor size ( $P = 0.1996$ ) were found to be predictors of recurrence.

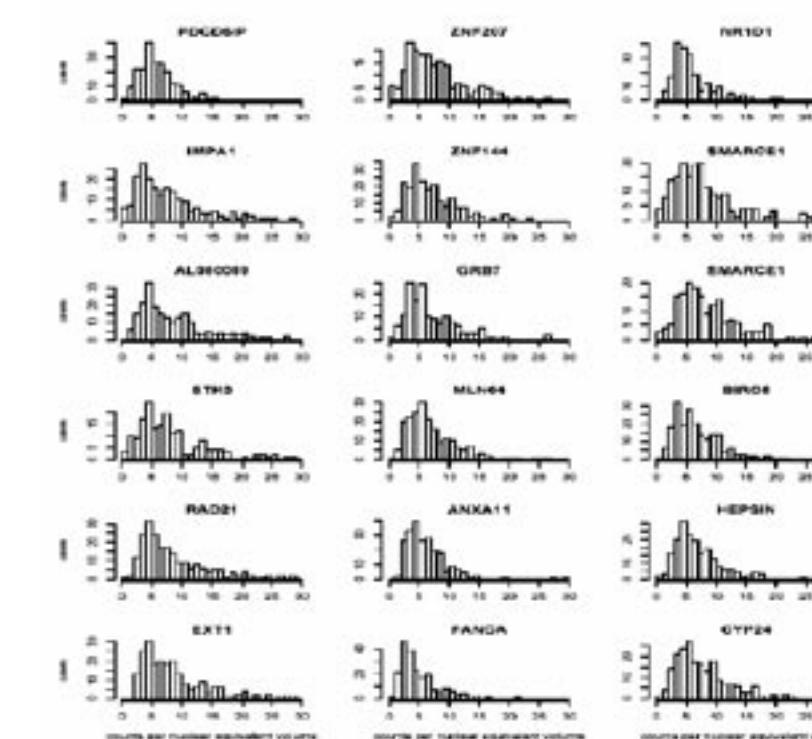
## Methods and Results - Data Analysis

**Raw Data.** Raw data (DNA copy number per tile) were normalized to copy number per nuclear equivalent volume (NEV) by dividing the observed FISH signals per nuclear DAPI-stained cross-sectional area by a computed cross-sectional area per nuclear equivalent volume (Pahlplatz et al, 1995). Of the 308 cases eligible for the study, FISH data were obtained from 265 specimens. Thirty-one (31) cases had been collected after neo-adjuvant therapy, and five specimens were associated with local recurrences. These 36 cases were set aside for further study and are excluded from the analysis reported here, leaving 229 cases for analysis. Mean numbers of DNA copy number per nuclear equivalent volume (NEV) across all specimens are summarized in the histogram below.

**Univariate Analysis.** For the 18 probes, univariate analysis revealed that 11 are significantly associated with distant recurrence at  $p < .05$  (Wilcoxon test). P values for the individual genes are reported in the table below. This result can be compared with the number of expected false positives for no association between copy number and recurrence. For 18 probes, on average less than one probe would be determined significant at  $p = .05$ . Thus these results validate the techniques employed in the original selection of the probes. However, we do not find any individual probe sufficiently prognostic to be clinically viable. Combinations of probes are required.

**Patterns of DNA copy number.** Patterns of copy number measurements for subsets of the 17 unique genomic regions correlating with distant recurrence were evaluated with respect to the performance of each subset's prognostic index. The prognostic index

maps a score calculated from the copy numbers of the predictive pattern to risk of recurrence. The search for predictive patterns was conducted in a randomly chosen training set in each analysis. The remaining samples were withheld as a blinded test set. The prognostic index scores were grouped in low, moderate and high risk categories, and negative and positive predictive values were calculated. Prognostic patterns were studied in Hormone Positive (HR+) and Hormone Negative (HR-) cases. The best prognostic pattern in each subgroup was further tested on the subset of cases that was lymph node negative (HR+, N-). Combinations of three genes were found to be optimal in detecting Hormone Receptor Positive (HR+ marker) changes that were indicative of good/poor prognosis. A separate combination of three genes was optimal in detecting Hormone Receptor Negative (HR- marker) changes that were indicative of good/poor prognosis.



The mean numbers of signals per nuclear equivalent volume, for each probe, across all specimens, are shown.

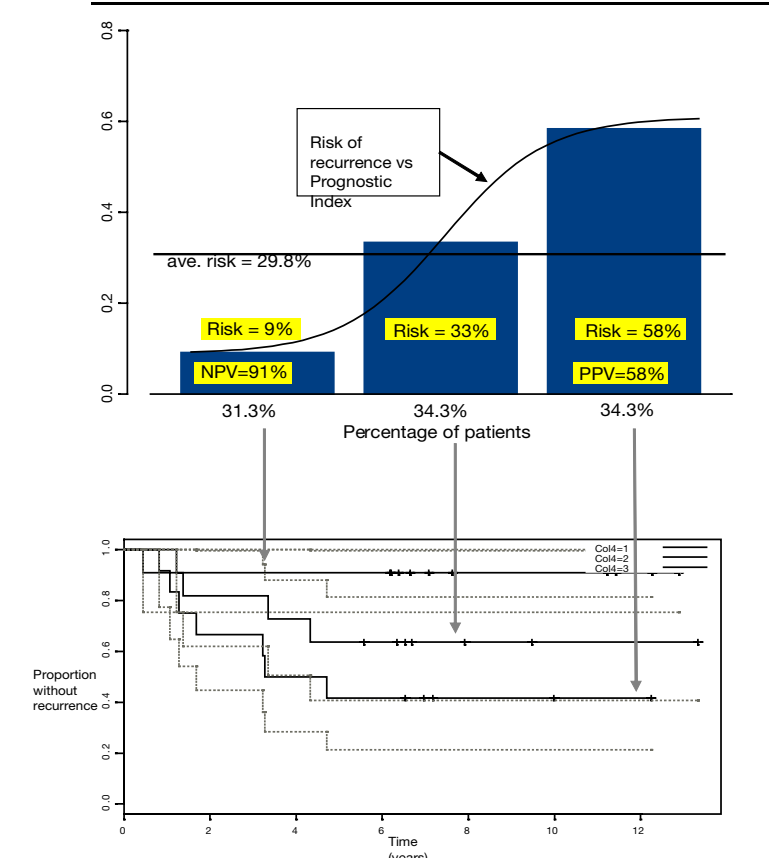
## PGA FISH Univariate Analysis

Probe	location	p (Wilcoxon)
CYP24	20q13.2	0.001
EXT1	8q24.11	0.002
NR1D1	17q21.1	0.003
MLN64	17q12	0.003
FANCA	16q24.3	0.004
BIRC5	17q25.3	0.007
ZNF144	17q12	0.008
RAD21	8q24.11	0.012
GRB7	17q12	0.016
HEPSIN	19q13.12	0.02
ZNF207	17q11.2	0.03
STK3	8q22.2	0.051
IMP1	8q21.13	0.082
AL080059	8q22.1	0.073
SMARCE1	17q21.2	0.075
ANXA11	10q22.3	0.086
SMARCE1 (dup)	17q21.2	0.153
PDCDEIP	3p23	0.526

If no markers were associated with recurrence, on average less than one false positive would be expected at  $p < .05$ .

11 markers associated with recurrence at  $p < .05$

## HR- Marker in Hormone Negative (HR-) test set



## HR- Marker in Hormone Negative (HR-) test set

	Training Data (N=13)		
	Low Risk	Moderate Risk	High Risk
No Recurrence	10	8	4
Recurred	1	1	1

Test Data (N=35)  
p=.04 Fisher's Exact Test

	Low Risk	Moderate Risk	High Risk
	No Recurrence	10	8
Recurred	1	4	7

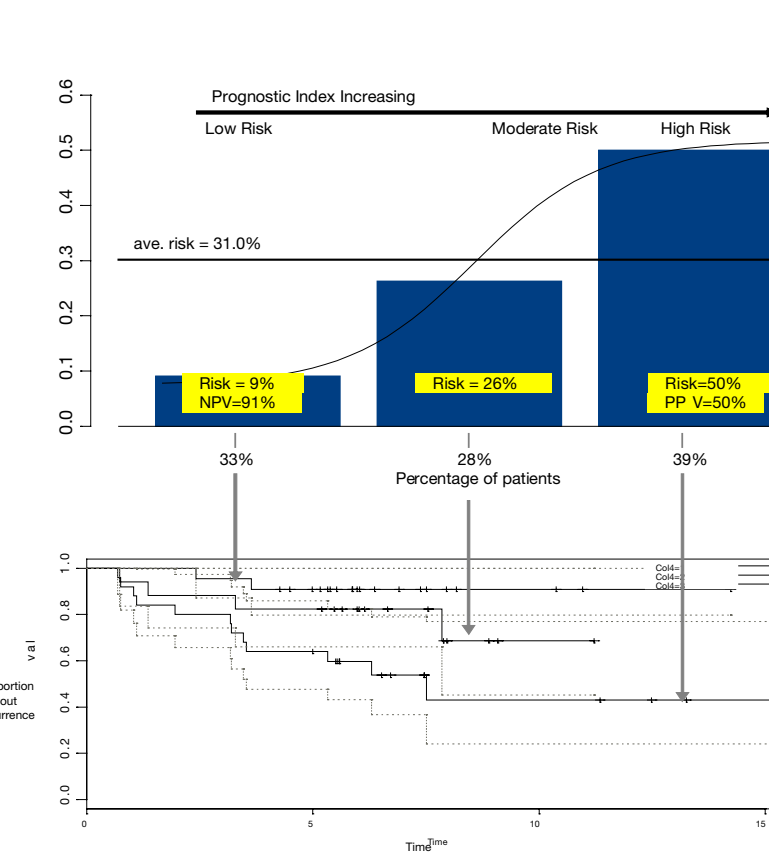
## HR- Marker on Node Negative (N0) test set

	Training Data (N=4)		
	Low Risk	Moderate Risk	High Risk
No Recurrence	1	2	0
Recurred	0	0	1

Test Data (N=16)

	Low Risk	Moderate Risk	High Risk
	No Recurrence	6	5
Recurred	0	1	1

## HR+ Marker in Hormone Positive in (HR+) test set

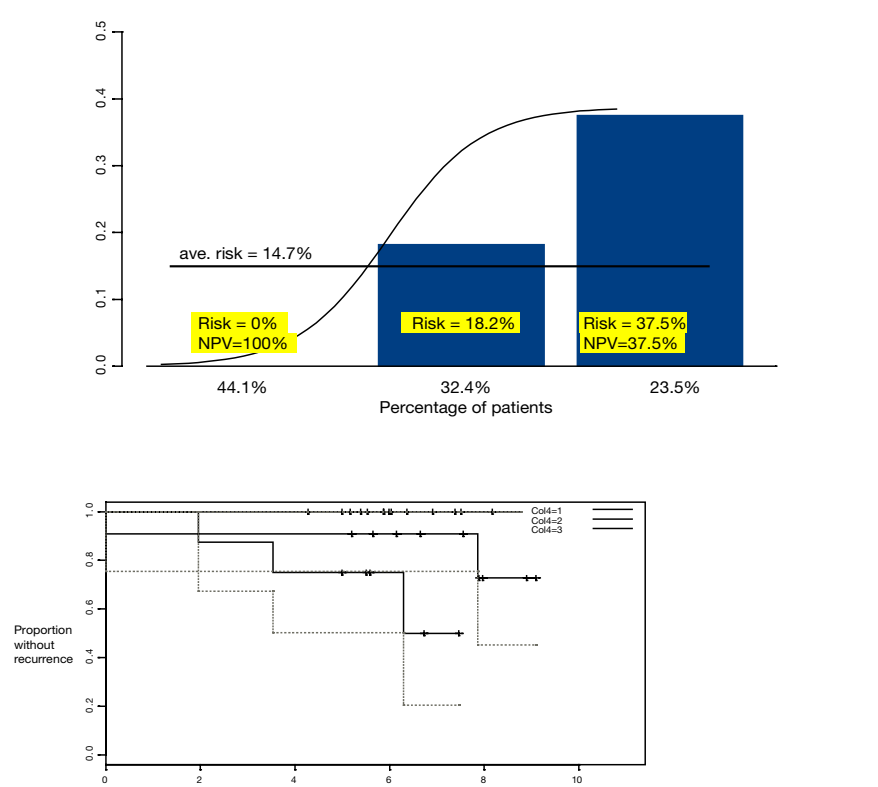


	Training Data (N=58)		
	Low Risk	Moderate Risk	High Risk
No Recurrence	17	16	7
Recurred	2	3	13

Test Data (N= 67)  
p=.0035 Fisher's Exact Test

	Low Risk	Moderate Risk	High Risk
	No Recurrence	20	14
Recurred	2	5	13

## Hormone Positive Node Negative (HR+, N0) test set



## Conclusions

- The FISH method described here, "patterns of genomic amplification FISH", or PGA FISH, is a valid assay for assessing gene copy number in archived tissue samples.
- PGA FISH allows assessment of gene copy number exclusively in carcinoma cells rather than stromal or inflammatory cells.
- 13 of the 17 genes are validated to either univariately predict recurrence or to participate in multiplex patterns that are prognostic.
- We found a three-gene marker in a training set that was prognostic for low risk of recurrence in women with hormone receptor negative (HR-) cancers. The marker's negative predictive value was 91% in independent test sets.
- We found a second 3-gene marker in a training set that is prognostic for low risk of recurrence in women with hormone receptor positive (HR+) cancers. The marker's negative predictive value was also 91% in independent test sets.
- In women with hormone receptor positive, lymph node negative (HR+, N0) cancers, the second marker had a NPV of 100% in a test set (0 of 15 women).
- The prognostic value of these genomic markers will be studied in larger numbers of women with hormone receptor positive, node negative (HR+, N0), and hormone receptor negative, node negative (HR-, N0) cancers.

## References

- Pahlplatz et al, Cytometry 1995 20(3):193  
Pollack et al. PNAS 2002 99(20):12963  
Sorlie et al. PNAS 2001 98(19):10869  
Van