Gene Expression Profiling in Neoadjuvant Clinical Trials

Lajos Pusztai M.D., D.Phil.

Associate Professor of Medicine

Department of Breast Medical Oncology

THE UNIVERSITY OF TEXAS
MD ANDERSON
CANCER CENTER

Tasks

- Potential value of gene expression profiling
- Conceptual and technical limitations
- Baseline assessment versus serial biopsies
- Sample size calculations

The value of gene expression profiling

- **Hypothesis testing**
  - **"New predictors can be discovered from human data"**
  - **"Predictors defined *in vitro* will also predict *in vivo*"**
- **Molecular Data Base Building**
  - **Matures over time**
  - **The larger the better**
  - **Uniform data acquisition is ideal**

Molecular data base ~ Clinical data base

Would you ever consider not recording patient age, race, tumor size, nodal status in a clinical research data base?

Why do we accept not measuring Topo II, Ki 67, AKT, p53, HER2, ER, PR, c-myc and ….12, 000 other genes?

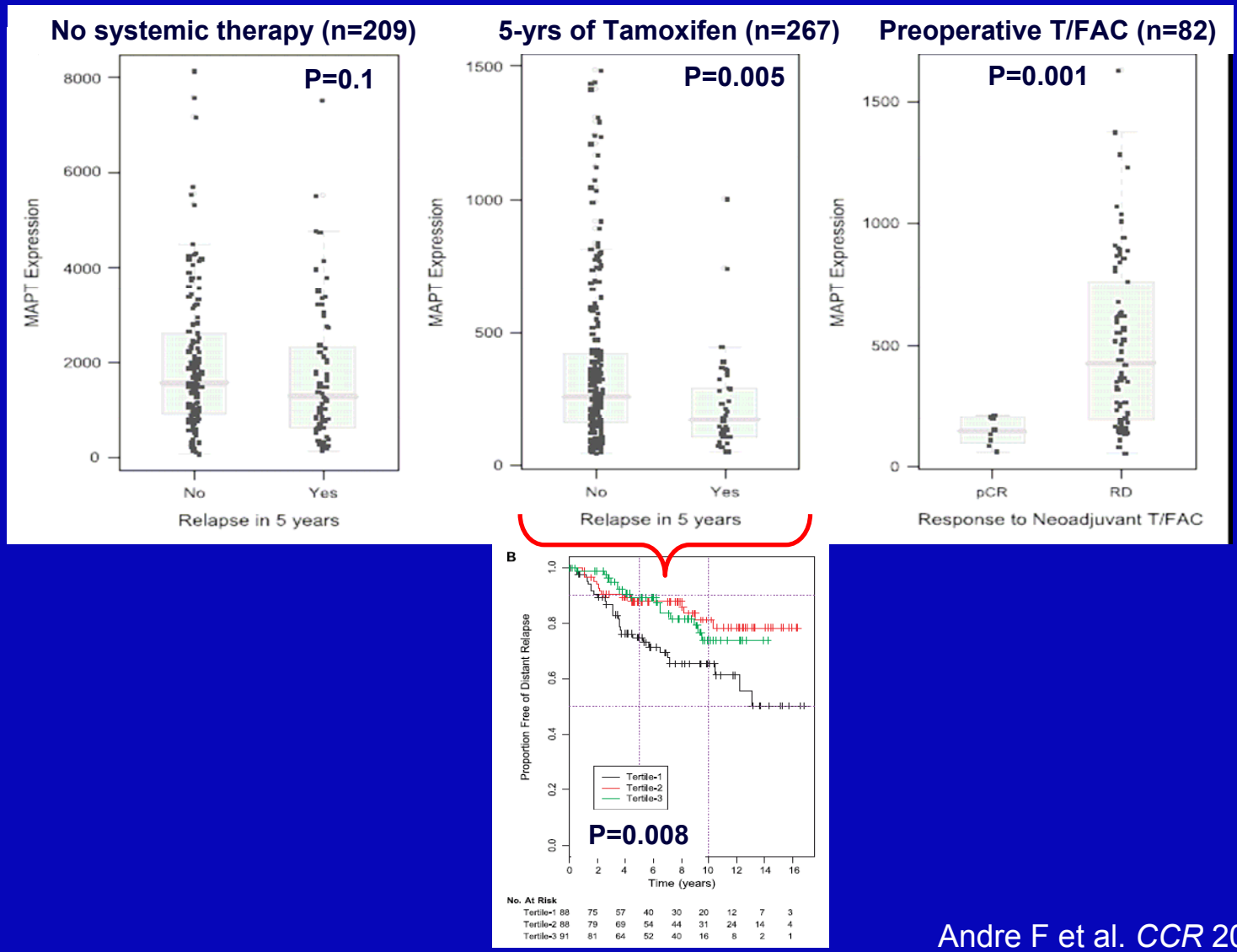
Microarray Quality Control Consortium.
Nature Biotechnology, Vol 24, 2006 (Sept 8).

The value of PUBLIC and uniform molecular data bases

- Is Marker X that is associated with good survival in Clinical Trial A based on a retrospective correlative study
  - a predictor of prognosis ?
  - a predictor of sensitivity to therapy ?
    - Which therapy chemotherapy or hormonal therapy ?

Wang J et al. (Affy U133A) *Lancet* 2005: 365; 671	Surgery alone	N=286
Loi S et al. (Affy U133A) *J Clin Oncol* 2007 (in press)	Tamoxifen x 5 yr	N=267
Hess K et al. (Affy U133A) *J Clin Oncol* 2006: 24; 4236	Neoadjuvant T/FAC	N=133

Example 1.
Is Tau prognostic or predictive to Tamoxifen or to T/FAC in ER-positive breast cancer?



Andre F et al. *CCR* 2007 (in press)

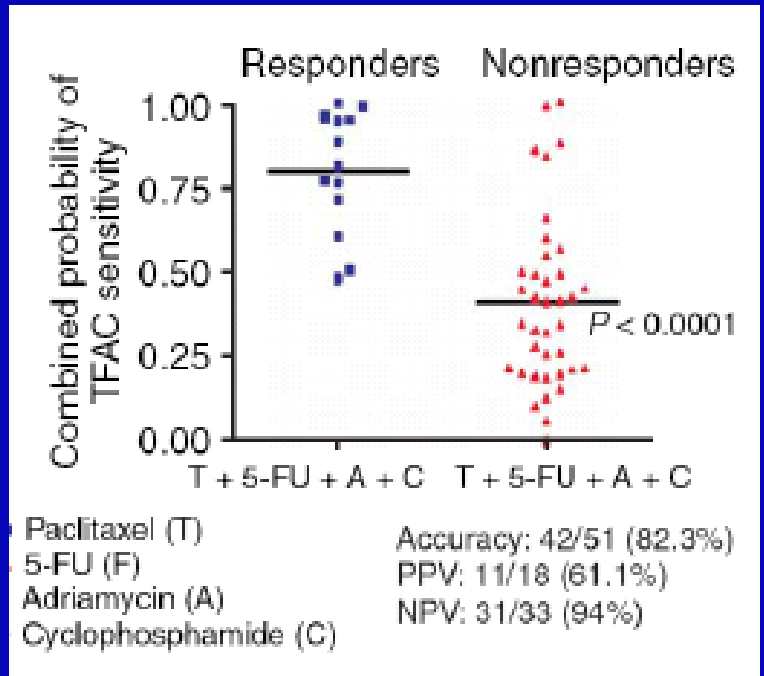
Example 2.
Does an in vitro defined chemotherapy sensitivity signature predict response in human breast cancer?

Define predictive signatures to paclitaxel, 5-FU, doxorubicin, and cyclophosphamide in cell lines In vitro

↓

Test predictor in human breast cancer neoadjuvant gene expression data

→



A Potti et al, Nature Med, Oct 22, 2006 on-line

Gene expression profiling as a tool to develop new predictive signatures from human data.

- **The Hypothesis: "Novel single genes or combination of multiple genes will yield powerful new predictors of response to therapy"**

Table 2 'Proof-of-principle' gene expression profiling studies in breast cancer to identify predictors of response to preoperative chemotherapy.

Reference	Patients and sampling	Chemotherapy	Gene set (*n*) and platform	Predictive accuracy
Chang *et al.*[36]	*n* = 24 + 6 (CNB)	Docetaxel	92 genes (Affymetrix HgU95)	88%[a]
Ayers *et al.*[37]	*n* = 24 + 18 (FNA)	T/FAC	74 genes (Millennium cDNA array)	78%[b]
Gianni *et al.*[38]	*n* = 89 (FFPE/CNB)	AT / weekly T	86 genes (Genomic Health RT–PCR)	Not done
Yoshimoto *et al.*[39]	*n* = 75 (CNB)	Weekly T	23 genes (cDNA & Agilent)	100% [a]
Sotiriou *et al.*[30]	*n* = 10 (CNB)	AC	37 genes (NCI cDNA array)	Not done
Hannemann *et al.*[40]	*n* = 49 (CNB)	AC versus AD	30 genes (NKI cDNA array)	Not done
Buchholz *et al.*[41]	*n* = 21 (CNB)	Mixed (AT, T, FAC)	241 genes (Millennium cDNA array)	Not done

The table summarizes the currently available published results on pharmacogenomic predictors of response to preoperative chemotherapy for breast cancer. Abbreviations: CNB, core needle biopsy; FNA, fine needle aspiration; FFPE, formaldehyde-fixed paraffin-embedded tissue; T, paclitaxel; FAC, 5-fluorouracil doxorubicin and cyclophosphamide; AT, doxorubicin and paclitaxel; AC, doxorubicin and cyclophosphamide; AD, doxorubicin and docetaxel. [a]Overall cross validation accuracy; [b]Overall accuracy in independent validation set.

Pusztai & Gianni. Prediction of response to preoperative chemotherapy in operable breast cancer. Nature Clin Practice Onc 1:44-50, 2004

The Challenges

- Low response rate
- Small sample size
- Multiple comparisons

- Confounding effect of the association between genes and clinical-pathological variables

Simon R. Development and evaluation of therapeutically relevant predictive classifiers using gene expression profiling. JNCI 98:1169,2006

ER-status and grade are associated with large scale gene expression patterns

- ER-negative and ER-positive cancers differ in the expression of several thousands of genes
  - (Gruvberger S, et al. Cancer Research; 61:5979, 2001
  - Pusztai L, et al., Clinical Cancer Res 9:2406, 2003.)
- Low grade tumors differ from high grade tumors in several hundreds of genes.
  - (Sotiriou C, et al . JNCI; 98:262, 2006)

ER-status and grade are ALSO associated with response to chemotherapies

Unadjusted comparison of responders with non-responders will yield gene lists dominated by ER- and grade-related genes.

What happens if one adjusts for ER and grade

Unadjusted for ER, grade
N=132 (33 pCR / 99 RD)

FDR	Number of Probe Sets (pCR vs RD)
0.00001	5
0.0001	27
0.001	112
0.01	408
0.05	1138

Cases matched by ER, grade
N=50 (25 pCR / 25 RD)

FDR	Number of Probe Sets (pCR vs RD)
0.00001	0
0.0001	0
0.001	0
0.01	0
0.05	0

This does not mean that there are no differentially expressed genes after adjustment for ER and grade, but that this particular discovery approach is high risk for false discovery.

List of differentially expressed genes from FAC- treated, grade- and ER-matched cases FDR=0.5 !

rank	p.values	pCR.mean	RD.mean	Gene
1	9.66E-06	2.242	2.168	VAX2
2	2.08E-05	1.912	1.989	SPATA6
3	2.44E-05	2.631	2.291	**TOP2A**
4	4.94E-05	2.259	2.422	LRBA
5	5.53E-05	3.205	3.084	KPNB1
6	0.000126995	2.695	2.415	**TOP2A**
7	0.000245899	1.986	1.918	TYMS
8	0.000275285	2.579	2.670	DCTD
9	0.000298475	2.573	2.645	LOC57146
10	0.000321767	2.923	2.668	TYMS

Half of these are spurious associations, the other half may be real discovery

Strong associations between ER, grade and large scale gene expression patterns bias pharmacogenomic discovery towards finding "general chemotherapy sensitivity" signatures.

In fact, these predictors may be, to large extent, the molecular equivalents of a combined ER-grade score.

Molecular predictors even if closely associated with ER and grade can still outperform (or compliment) pathologic variable based prediction models

Table 3. Performance metrics of the genomic and clinical predictors in the validation set (n=51)

A.	Clinical variables (age, ER, nuclear grade)	DLDA-30 probe sets
Accuracy:	0.78 (0.65-0.89)	0.76 (0.62-0.87)
Sensitivity:	0.61 (0.32-0.86)	0.92 (0.64-1.0)
Specificity:	0.84 (0.69-0.94)	0.71 (0.54-0.85)
PPV:	0.57 (0.29-0.82)	0.52 (0.3-0.7)
NPV:	0.86 (0.71-0.95)	0.96 (0.82-1.0)

KR Hess, et al J Clin Oncol 24:4236-4244, 2006.

Grade cannot be targeted with drugs, the molecular mechanisms that determine grade might serve as new therapeutic targets !

Statistical simple size calculations for discovery studies

- There are many different ways to estimate sample size for multi-gene predictor discovery…(none of them guaranties success !)
  - Sample size depends on (i) background heterogeneity, (ii) magnitude of expression difference, and (iii) event rate.
    - Power the study to identify individual differentially expressed genes
    - Adjust sample size based on interim look (i.e. fit learning curves to predictors)

Table 1. Sample size required to discover differentially expressed genes on the basis of standardized effect size and response rate

Response rate (%)	Standardized effect size		
	1.0	0.75	0.5
10	166	295	663
25	80	142	319
50	60	106	239

The table shows the numbers of patients required to detect differentially expressed genes between responders and non-responders with a two-sided α error rate of 1% and β error rate of 10%.

Learning curves for Genomic Predictors in different data sets

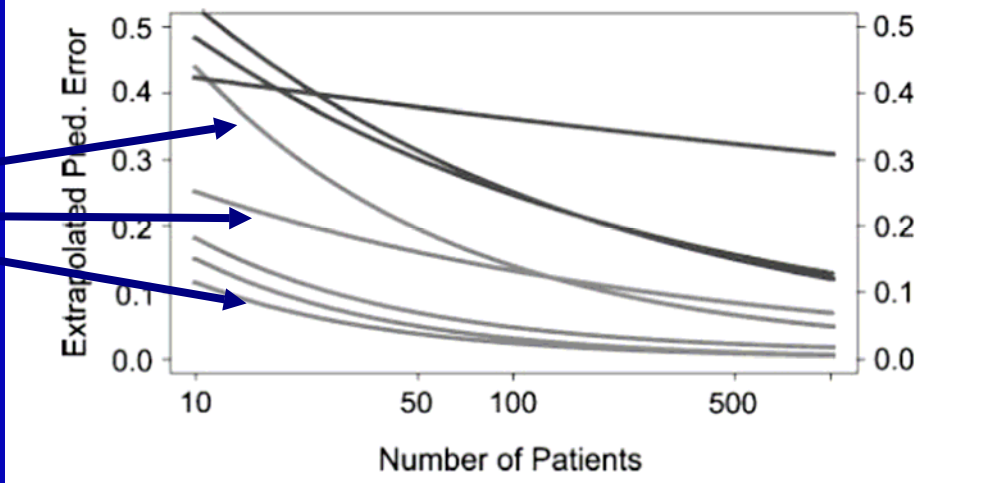


Figure 1. Simulated cross validation classification error rates of support vector machine learning algorithm as a function of training sample size.

Pusztai L, Hess KR. Clinical trial design for microarray predictive marker discovery and assessment. Ann Oncol 15(12):1731-1737, 2004

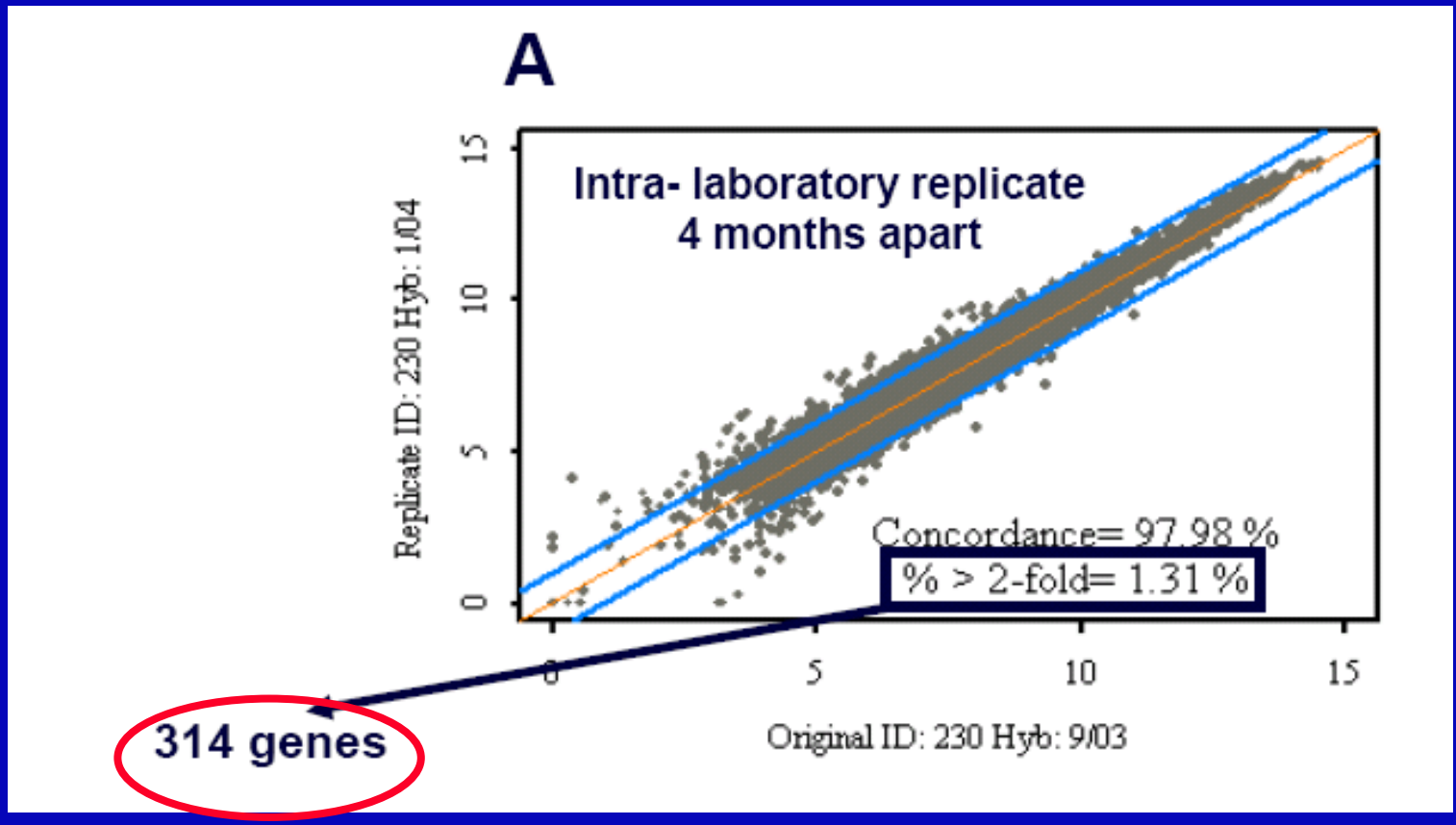
Validation design

- **Marker-positive patients have higher pCR rate than the average rate in unselected patients.**
  - **Lower bound of 95% CI of PPV of the test is above upper bound of 95% CI of pCR rate in unselected patients**
  - **Sensitivity is also important**
- **Precision of estimate**
  - **How large the study needs to be in order to conclude that we are certain about the predictive accuracy (whatever the accuracy is!)**
    - **The point estimate of outcome needs to have acceptable standard deviation**

Serial biopsies to examine transcriptional response over time after therapy.

- The Hypothesis: ”Transcriptional changes in response to therapy will be more predictive of outcome than base-line gene expression pattern”
- The Additional Problems:
  - Optimal timing is unknown
  - Time variation due to non-compliance
  - Missing data (“optional procedure”)
  - Changes in a few dozen important genes may be blurred by technical noise.

Technical noise

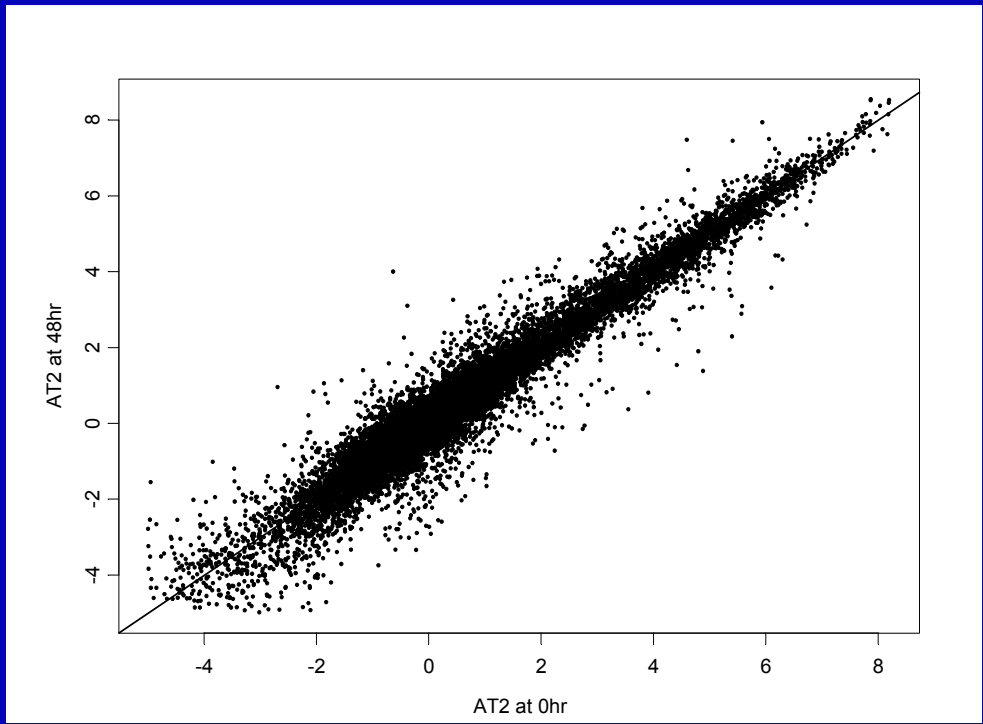


Microarrays are somewhat risky strategy to detect less than large scale trasncriptional changes in SMALL sample sets

MAQC consortium.. *Nature Biotechnology 24:1151-1161, 2006.*

K Anderson et al. *Clin Cancer Res 12:1721, 2006*

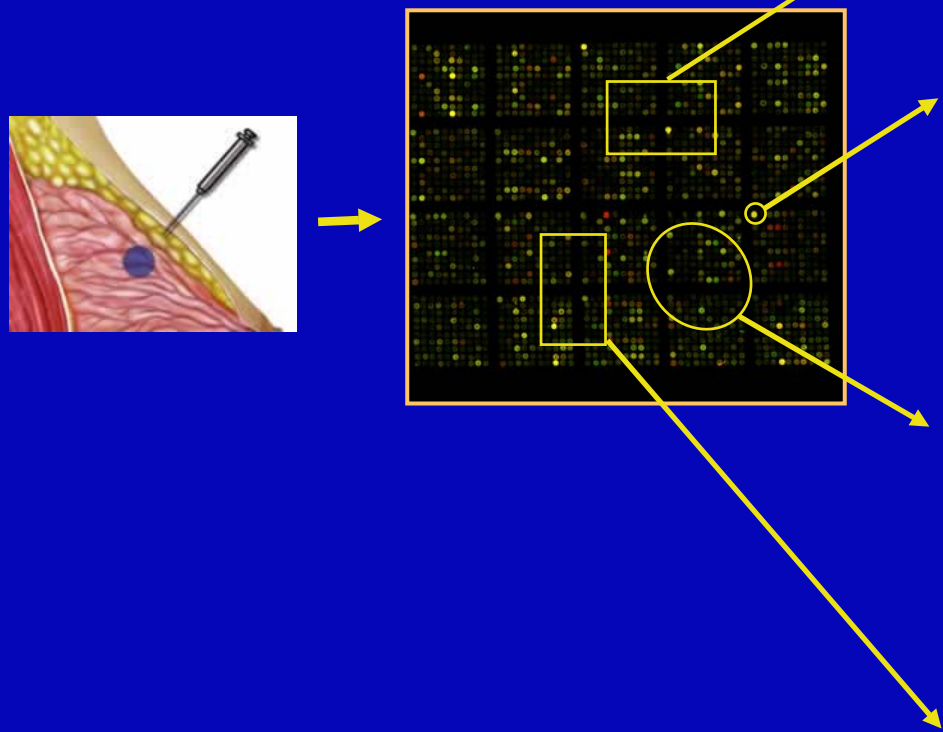
Gene expression changes in serial core biopsy AT 0h vs. AT 48h



"…Surprisingly, different genes changed after chemotherapy, in each patient no single gene showed consistent expression change in all 5 patients…"

Buchholz TA., et al, Cancer J 8:461-468, 2002.

What genomic tests are currently available for clinical testing ?



Estimate prognosis
70-gene MammaPrint
76-gene predictor

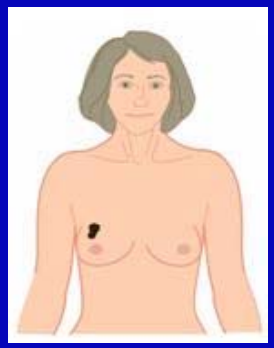
Measure ER and HER2
(mRNA based!)

Predict endocrine sensitivity among ER-positive patients
Oncotype DX, Luminal A type, MDACC-200 gene index

Predict chemo-sensitivity
Onoctype DX, MDACC-30 gene, Proliferation signature, GGI

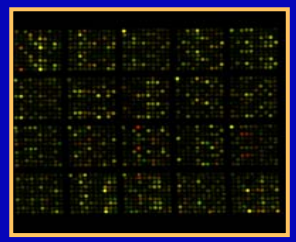
Can we find clinical value in a series of imperfect molecular predictors ?

MDACC 2006-0543: Neoadjuvant Molecular Triaging Protocol for Stage I-III breast cancer



Day 1

FNA



DLDA-30 & SET

Day 7

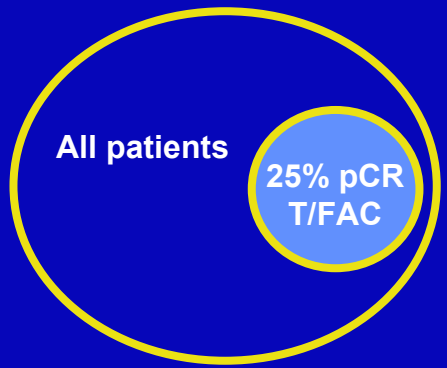
HER2 + → Trastuzumab +Chemo

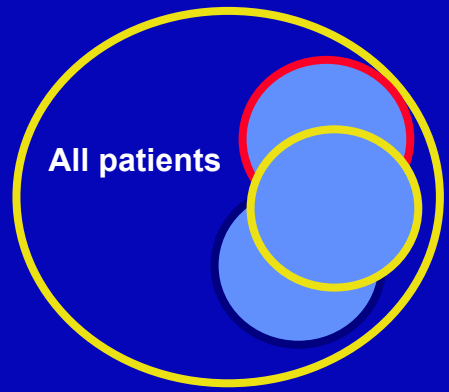
T/FAC sensitive → pre-op Taxol/FAC

Hormone sensitive → pre-op AI

Not sensitive → New Drug TRIAL (targeted drugs based on pathway signatures)

The objective is to determine if triaging patients into treatment groups can increase pathologic CR rates for the whole population.





pCR with endocrine Tx

pCR with T-FEC

pCR with new agent

Simultaneous endocrine and chemotherapy sensitivity prediction results in 126 cases

Neoadjuvant chemotherapy

Predicted Endocrine Sensitivity by ER genomic score	Predicted T/FAC Response by gene signature		
	RD	pCR	%
Low	54	52	84%
Intermediate	11	1	10%
High	8	0	6%
%	58%	42%	

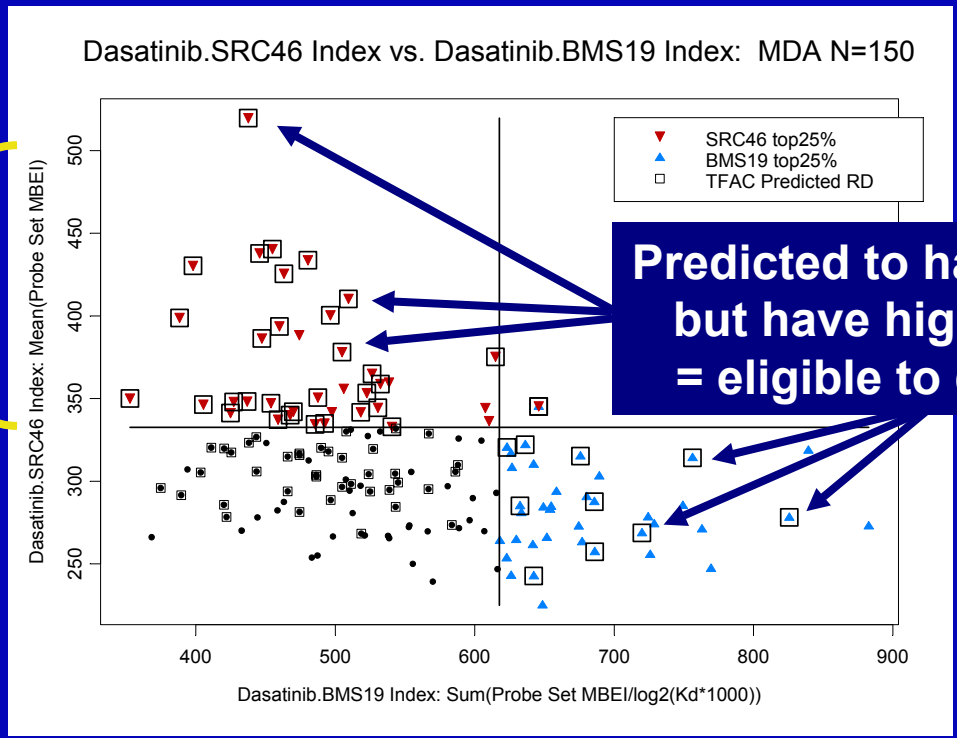
Neoadjuvant endocrine therapy

Pathway tailored neoadjuvant Phase II trial

We applied 3 distinct predictors to each of 150 FNAs

1. MDACC 30-gene pCR predictor
2. BMS dasatinib cell line predictor (F Huang et al., Cancer Res, 2007)
3. 19-gene dasatinib target index

Top 25% BMS Prediction score



Predicted to have residual disease but have high dasatinib scores = eligible to dasatinib + chemo

Top 25% Target Index

Conclusions 1

- Genes are not independent variables, there are large scale, coordinated expression patterns that are associated with clinical variables
  - This helps the discovery of general chemotherapy sensitivity signatures.
  - It hinders the discovery of regimen specific markers.
    - "Chance favors the bold"
- Microarrays give a reliable and reproducible snapshot of global gene expression status of breast cancer.
  - This allows building of large molecular data-bases that will be an invaluable resource for discovery and hypothesis testing !
    - Raw data (!) must be made public and uniform platform is desirable!

Conclusions 2

- Time is right to start to test the clinical value of response predictors (single gene or multi-gene…) prospectively.
  - Oncotype DX, MDACC-30 gene predictor, Topoisomerase II, Luminal A types, ….
- Individually moderately accurate predictors may be assembled into a clinically relevant diagnostic strategy.