Cancer Therapy: Clinical

Clinical Cancer Research

Gene Expression, Molecular Class Changes, and Pathway Analysis after Neoadjuvant Systemic Therapy for Breast Cancer

Ana M. Gonzalez-Angulo[1,2], Takayuki Iwamoto[1], Shuying Liu[1], Huiqin Chen[1], Kim-Anh Do[3], Gabriel N. Hortobagyi[1], Gordon B. Mills[2], Funda Meric-Bernstam[4], W. Fraser Symmans[5], and Lajos Pusztai[1]

Abstract

Purpose: To examine gene expression differences between pre– and post–neoadjuvant systemic therapy (NST) specimens of breast cancers and identify biologic changers that may lead to new therapeutic insights.

Methods: Gene expression data from prechemotherapy fine needle aspiration specimens were compared with resected residual cancers in 21 patients after 4 to 6 months of NST. We removed stroma-associated genes to minimize confounding effects. PAM50 was used to assign molecular class. Paired *t* test and gene set analysis were used to identify differentially expressed genes and pathways.

Results: The ER and HER2 status based on mRNA expression remained stable in all but two cases, and there were no changes in proliferation metrics (Ki67 and proliferating cell nuclear antigen expression). Molecular class changed in 8 cases (33.3%), usually to normal-like class, which was associated with low residual cancer cell cellularity. The expression of 200 to 600 probe sets changed between baseline and post-NST samples. In basal-like cancers, pathways driven by increased expression of phosphoinositide 3-kinase, small G proteins, and calmodulin-dependent protein kinase II and energy metabolism were enriched, whereas immune cell–derived and the sonic hedgehog pathways were depleted in residual cancer. In non–basal-like breast cancers, notch signaling and energy metabolism (e.g., fatty acid synthesis) were enriched and sonic hedgehog signaling and immune-related pathways were depleted in residual cancer. There was no increase in epithelial–mesenchymal transition or cancer stem cell signatures.

Conclusions: Our data indicate that energy metabolism related processes are upregulated and immune-related signals are depleted in residual cancers. Targeting these biologic processes may represent promising adjuvant treatment strategies for patients with residual cancer. *Clin Cancer Res; 18(4); 1109–19. ©2012 AACR.*

Introduction

Neoadjuvant systemic therapy (NST) provides an opportunity to directly assess the chemotherapy sensitivity of breast cancer and also benefits patients with tumor down staging and lesser surgery (1). Residual cancer volume after NST carries important prognostic information, and attaining a pathologic compete response (pCR: absence of invasive cancer in the breast and lymph nodes) heralds excellent long-term survival (2–5). Conversely, patients with residual cancer have variable prognosis. Some patients, particularly those with estrogen receptor (ER)-positive cancer, have good survival despite residual cancer, others, particularly those with triple receptor negative cancer, have poor prognosis if they have residual disease (3–5). NST with anthracycline and/or taxane-containing regimens results in pCR rates between 6% and 30%, depending on tumor grade and receptor status (6–8). Because of the clear association between residual cancer burden and survival, NST is an *in vivo* screen for efficacy of neoadjuvant chemotherapies. However, despite the well recognized poor prognosis of ER- and HER2-negative patients with residual cancer after neoadjuvant chemotherapy, no effective adjuvant therapy exists for this population.

One approach to increase treatment efficacy is to develop regimens that induce higher rates of pCR; the combination of trastuzumab and chemotherapy represents a successful example of this approach (9). Another strategy could be to develop biologically targeted post-NST, adjuvant therapies that exploit the vulnerabilities of the

Authors' Affiliation: Departments of [1]Breast Medical Oncology, [2]Systems Biology, [3]Biostatistics, [4]Surgical Oncology, and [5]Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Corresponding Author: Ana M. Gonzalez-Angulo, Department of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Unit 1354, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-2817; Fax: 713-794-4385; E-mail: agonzalez@mdanderson.org

doi: 10.1158/1078-0432.CCR-11-2762

Translational Relevance

Neoadjuvant systemic therapy (NST) provides an opportunity to directly assess the therapy response in breast cancer. Residual cancer volume after NST carries important prognostic information. By examining gene expression differences between pre- and post-NST specimens of breast cancers, we could identify biologic changers that may lead to new therapeutic insights for a biologically resistant disease. By using gene expression data from prechemotherapy fine needle aspiration specimens and comparing them to surgically resected residual cancers, we found that expression of 200 to 600 probe sets changed between baseline and postchemotherapy samples and that energy metabolism–related processes are upregulated and immune-related signals are depleted in residual cancers. Targeting these biologic processes may represent promising adjuvant treatment strategies for patients with residual cancer.

residual cancer. Such approach will first require detailed molecular characterization of residual cancer and definition of targetable molecular pathways. The purpose of this analysis was to compare matched gene expression profiles of pretreatment cancer with residual cancer after NST. We assessed changes in molecular class defined by the PAM50 classifier; changes in ER, Ki67 [proliferating cell nuclear antigen (PCNA)], and HER2 status based on mRNA expression and also identified individual genes and gene sets that were differentially expressed between the matched specimens to define molecular pathways that seem to be enriched in post-NST residual cancer.

Materials and Methods

Patients and tumor tissues

Prechemotherapy fine needle aspirates (FNA) were collected in the context of a prospective biomarker discovery program conducted at MD Anderson Cancer Center. Matching postchemotherapy frozen surgical specimens were retrieved from the institutional tumor bank, patients with pCR were excluded from the search. We identified 25 cases with matching specimens, 21 of these patients received NST. The remaining 4 cases with matching FNA biopsy and surgical tissue without NST were used to estimate methodologic variability in repeat gene expression data from the same cancer. Patient characteristics for the 21 patients who received NST are in Supplementary Table S1. The amount of residual cancer was determined by routine pathology examination of lumpectomy or mastectomy specimens. In addition, we also carried out molecular class prediction on paired FNA and core needle biopsies obtained during one biopsy session, without any intervening chemotherapy from 37 cancers to assess the sensitivity of the classification method to tissue sampling (10). This research was approved by the Institutional Review Board, and all patients signed informed consent to allow biomarker studies to be carried out on their specimens.

RNA extraction and gene expression profiling

RNA was extracted from the FNA and surgical tumor samples using the RNAeasy Kit (Qiagen). The amount and quality of RNA were assessed with DU-640 UV Spectrophotometer (Beckman Coulter), and they were considered adequate for further analysis if the optical density$_{260/280}$ ratio was 1.8 or more and the total RNA yield was 1 μg or more. cRNA generation, second-strand cDNA synthesis, and gene expression profiling with Affymetrix HG-U133A gene chips were done as described previously (10–12). The 25 matching pre- post-NST gene expression data is available under GEO (Gene Expression Omnibus) accession number GSE 32072 and the 37 matching FNA and core biopsy data are available under accession number GSE 32518.

Data analysis

All gene expression data was normalized with MAS5 algorithm, mean centered to 600 and $\log_2$ transformed before further analysis. We removed from further analysis probe sets with average expression values $\leq$ the lowest 25% to reduce noise from low expressed probe sets and also excluded stroma-related genes ($n = 1{,}618$ probe sets; Supplementary Table S2) that were defined in a previous publication (10). Stromal genes were removed from the analysis to minimize false discovery due to the different tissue composition of the pretreatment FNA and postchemotherapy surgically resected tissues (13). Paired *t* test was used to identify differentially expressed genes between cohorts, and false discovery rates (FDR) were calculated by the Benjamini and Hochberg permutation method (14). Differentially expressed genes were defined from the entire data set as well as from basal-like ($n = 9$) and non–basal-like ($n = 12$) cancers separately. The function of the differentially expressed genes was mapped to biologic pathways using the Ingenuity Pathway Analysis software (IPA, http://www.ingenuity.com/). We also examined changes in the mRNA expression values of ESR1 ("205225_at") and ERBB2/HER2 ("212022_s_at") receptors and the proliferation markers Ki67 ("216836_s_at") and PCNA (201202_at). An ESR1 mRNA probe set expression value of more than 10.18 was considered as ER positive, and HER2 probe set expression value more than 12.54 was considered HER2 positive, on the basis of a previously established thresholds (10, 15). Molecular class assignment was done by using the PAM50 method, briefly, we calculated Spearman's rank correlation between each sample and each subtype centroid and assigned the class of the most highly correlated centroid to each sample (16). Six (*ANLN*, *CDCA1*, *CXXC5*, *GPR160*, *TMEM45B*, and *UBE2T*) of the 50 genes that are included in the PAM50 predictor could not be mapped to the Affymetrix HGU133A chip, the remaining 44 genes were used for molecular class prediction. PAM50 class concordance rate was calculated using the number of matched pairs that were assigned to the same subtypes

divided by the total number of matched pairs. We also used gene set analysis (GSA) to assess changes in a priori defined gene set before and after chemotherapy (17). In this analysis we included 235 canonical biologic pathways from IPA and 2 additional gene sets (Mego and colleagues, $n = 254$ genes, Creighton and colleagues, $n = 230$ genes) previously reported to be associated with residual disease after NST in breast cancer (Supplementary Table S3) (18, 19). All analysis was done using BRB Array Tools v 4.1.0 (http://linus.nci.nih.gov/BRB-ArrayTools.html) and R software v 2.7.2 (http://www.r-project.org; ref. 20). Significance was estimated with permutation test ($n = 1{,}000$). The null hypothesis was that the average degree of differential expression of members of a given gene set between the pre- and post-NST samples was the same as expected from a random set of genes of similar size.

Results

Changes in estrogen receptor and HER2 expression and proliferation measures

First, we assessed changes in individual molecules with therapeutic and prognostic relevance. Pre- and posttreatment ER and HER2 levels were highly correlated, the Spearman's rank correlation coefficients were $R = 0.706$ and 0.836, respectively. The ER status before and after NST changed from ER-negative to ER-positive on one case and from ER-positive to ER-negative in another case (Fig. 1A). HER2 status also changed in 2 cases in a similar manner (Fig. 1B). Interestingly, among the HER2-negative ($n = 18$) cancers, ERBB2 mRNA levels were slightly, but significantly, higher in the post-NST samples ($P = 0.011$, unadjusted for multiple comparisons). However, this was not sufficient to alter HER2 status in any but one case, indicating relative stability and robustness of ER and HER2 status. The pre- and posttreatment expression levels of proliferation markers Ki67 and PCNA were moderately correlated, Spearman's rank correlation coefficients were $R = 0.301$ and 0.395, respectively (Fig. 1C and D). There were no significant differences in proliferation measures between pre- and post-NST samples across all cases or when ER-positive and ER-negative cases were examined separately (data not shown). The ER, HER2, and Ki67 or PCNA expression levels did not change significantly in any of the 4 control specimens that received no NST.

Changes in molecular class before and after chemotherapy

Next, we examined the stability of molecular class in the 21 pre- and post-NST specimens. Table 1 shows the class assignment of each sample including the correlation values to each reference class centroid (the larger the correlation coefficient, the greater similarity the sample has to a given centroid and class is determined by the largest coefficient). There were 8 cases in which the molecular subtype by PAM50 changed between pre- and postchemotherapy samples, corresponding to an overall concordance rate of 62%. Most of the time (6 of 8 samples), the change was to normal-like cancer in the post-NST sample; this was accompanied by an increase in the correlation coefficient for the normal-like centroid and decrease in the correlation to the original class centroid. In 3 cases the molecular class changed from basal-like to normal, in 2 cases from luminal A to normal, and in one case, each from HER2 enriched to normal. The residual cancer cellularity in many of these cases was low, 10%, 20%, and 30% (in 2 cases it remained relatively high 50% and 60%; Table 1). One case changed from HER2 enriched to luminal B and another from normal to basal-like. In 4 cases the changes in the correlation metrics were quite large, indicating substantial change in the gene expression pattern of the 44 genes that were used for classification. In the remaining cases, relatively small-scale changes in correlation values altered the class assignment that suggests less than ideal robustness of the class assignment method under these circumstances. When PAM50 was applied to the baseline FNA and the surgical specimen of the 4 control cases, there were no changes in the molecular subtype of the pairs. To further assess the effect of tissue sampling on molecular class assignment, we also applied the PAM50 predictor to 37 matching FNA and core biopsies that were collected simultaneously before any chemotherapy (10). Molecular class changed for the same tumor depending on sampling method in 6 cases, corresponding to an overall concordance rate of 83% (Supplementary Table S4). These observations indicated that alterations in tissue cellularity due to chemotherapy, or due to sampling method, can introduce substantial variability to molecular class prediction.

Enriched and depleted molecular pathways in residual cancers

To identify gene expression changes in residual cancer compared with pretreatment samples, we carried out paired sample *t* test on 15,062 probe for all cases ($n = 21$) and also separately for the basal-like ($n = 9$) and non–basal-like ($n = 12$) cancers. When all cancers were analyzed together, 532 probe sets were differentially expressed between pre- and posttreatment samples at FDR ≤ 0.005 (corresponding to $P \leq 0.0003$), 271 were overexpressed, and 261 were underexpressed in residual cancer relative to pretreatment samples (Supplementary Table S5). When only basal-like cancers were considered, 77 probe sets were overexpressed and 115 were underexpressed. In non–basal-like cancers the corresponding numbers were 149 and 135, respectively (Supplementary Table S5). To find out what molecular pathways may be affected by these transcriptional changes, we mapped the over- and underexpressed genes into biologic pathways by IPA. Table 2 lists the 10 canonical pathways that were up- or downregulated in the residual cancers. Results are presented for all cancers as well as for the basal-like and non–basal-like subsets. When cancers were analyzed without molecular stratification, the top 10 upregulated pathways in residual cancer included ER signaling and nuclear receptor signaling (e.g., PPAR signaling, PPARα/RXRα activation, Nur77 signaling in T lymphocytes, and aryl hydrocarbon receptor signaling), as well as the

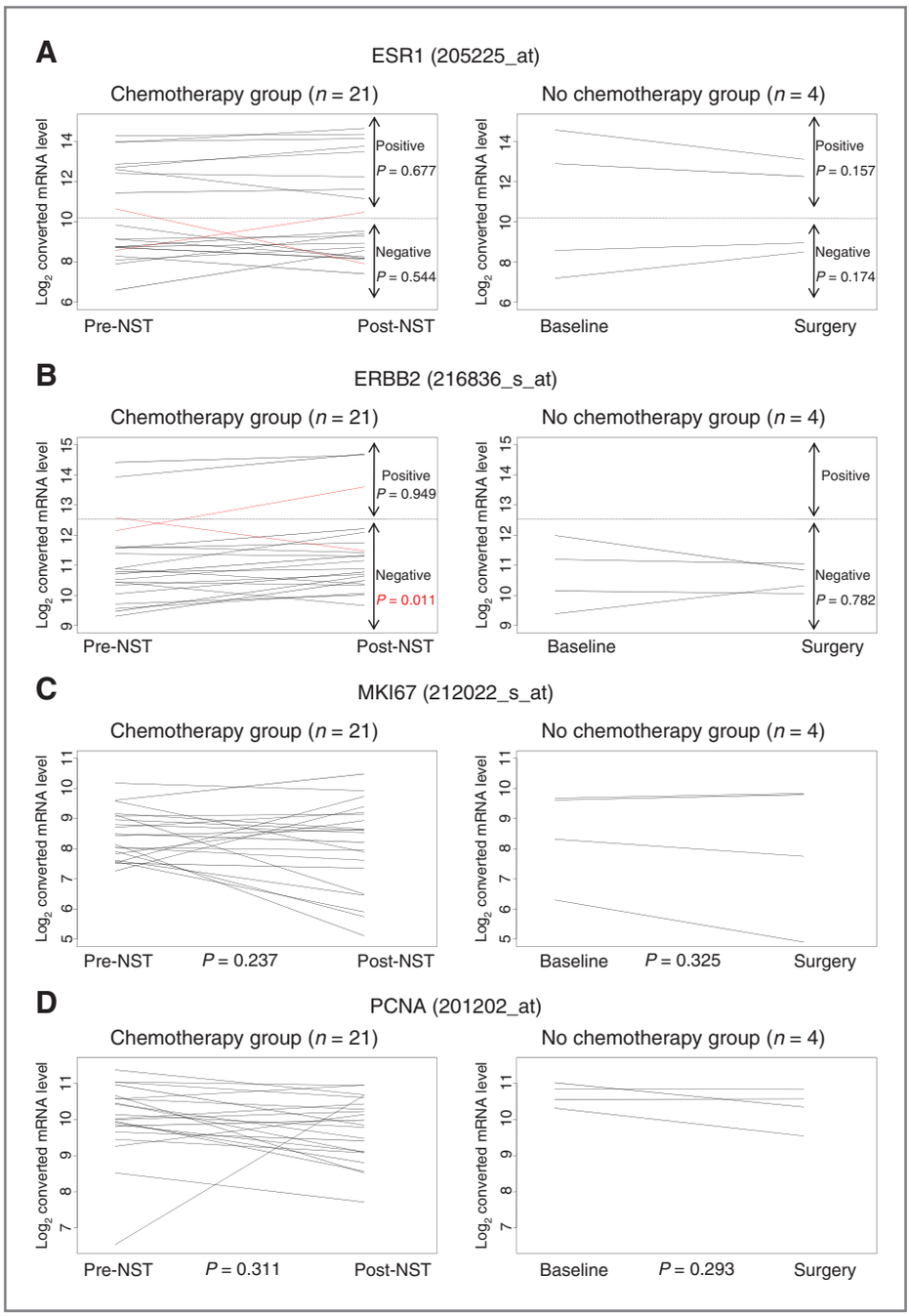

Figure 1. Gene expression levels of ESR1 (A), ERBB2 (B), MiK67(C), and PCNA (D) in pre- and postneoadjuvant chemotherapy samples ($n = 21$). *P* values were calculated from paired *t* test comparing pre- and posttreatment groups and are not adjusted for multiple comparisons.

Notch and WNT signaling pathways. Pathways involved in hedgehog signaling, PI3K/AKT, ATM, and cell-cycle control were depleted in the residual cancers. It is likely that the different molecular subtypes of breast cancer may use different survival pathways or rely on distinct mechanisms of drug resistance and, therefore, we also conducted the same pathway analysis basal-like and non–basal-like breast cancers separately. Metabolic pathways (glutamate metabolism, chondroitin sulfate biosynthesis, and propanoate metabolism) were the most consistently upregulated in basal-like residual cancers; the other upregulated pathways were united by the key functional roles of increased phosphoinositide 3-kinase (PI3K), small G-protein, and calcium/calmodulin-dependent protein kinase II alpha (CAMK2) expression. Notch, hypoxia, and TNF and ILK signaling were upregulated in non–basal-like cancers. Sonic hedgehog signaling was consistently depleted in residual cancers in both molecular types, and the depletion of

Table 1. PAM50 prediction results of the 21 pre- and postchemotherapy cases

		Prechemo (FNA)					Postchemo (surgical excision)						
		Correlation coefficient to each centriod						Correlation coefficient to each centriod					% residual tumor cellularity
Sample ID	Subtype	Basal	Her2	LumA	LumB	Normal	Subtype	Basal	Her2	LumA	LumB	Normal	
703	Normal	0.006	−0.07	0.156	−0.329	0.302	Normal	0.01	0.027	0.221	−0.351	0.282	40
647	Basal	0.376	−0.141	−0.093	−0.356	0.291	Basal	0.566	0.183	−0.428	−0.052	−0.106	40
604	LumA	−0.525	−0.376	0.679	0.027	0.153	LumA	−0.484	−0.252	0.585	0.069	−0.026	60
595	Her2	0.062	0.569	−0.278	0.083	−0.205	Her2	0.197	0.511	−0.521	0.151	−0.353	60
569	LumA	0.021	−0.196	0.219	−0.169	0.113	LumA	−0.549	−0.34	0.689	−0.101	0.195	50
566	LumA	−0.243	−0.312	0.523	−0.227	0.385	Normal	−0.13	−0.261	0.285	−0.222	0.309	**30**
570	Basal	0.624	−0.111	−0.294	−0.464	0.292	Basal	0.461	−0.159	−0.135	−0.381	0.286	30
521	Basal	0.681	0.002	−0.442	−0.274	0.155	Basal	0.615	0.159	−0.52	−0.145	0.041	70
439	Basal	0.219	0.212	−0.066	−0.264	0.151	Normal	0.16	0.195	−0.122	−0.355	0.216	**20**
380	Her2	0.119	0.141	−0.311	0.088	−0.173	LumB	−0.259	0.321	−0.156	0.533	−0.41	60
364	LumA	−0.221	−0.312	0.482	−0.208	0.326	LumA	−0.484	0.023	0.501	0.055	0.026	50
312	Basal	0.539	−0.047	−0.223	−0.485	0.358	Normal	0.363	−0.162	0.113	−0.574	0.481	**50**
208	LumA	−0.501	−0.386	0.677	−0.085	0.137	LumA	−0.542	−0.19	0.639	0.053	0.116	60
572	Basal	0.642	0.129	−0.4	−0.254	0.098	Normal	0.191	−0.014	0.035	−0.434	0.349	**10**
151	Basal	0.406	0.177	−0.284	−0.228	0.107	Basal	0.491	0.204	−0.27	−0.365	0.135	Not available
269	LumA	−0.265	−0.208	0.42	0.11	−0.105	LumA	−0.165	−0.4	0.647	−0.511	0.589	Not available
307	Her2	0.163	0.31	−0.057	−0.34	0.181	Normal	0.122	0.099	0.111	−0.562	0.409	**60**
476	LumA	−0.348	−0.268	0.646	−0.307	0.462	Normal	−0.107	−0.327	0.47	−0.384	0.539	**Not available**
517	Normal	0.143	0.075	0.043	−0.284	0.255	Basal	0.803	0.018	−0.533	−0.317	0.152	Not available
667	Basal	0.498	0.199	−0.591	−0.001	−0.222	Basal	0.506	0.146	−0.448	−0.11	−0.039	Not available
688	Basal	0.615	0.078	−0.389	−0.367	0.136	Basal	0.586	0.089	−0.424	−0.362	0.1	Not available

NOTE: Cases with discordant class assignment before and after chemotherapy are highlighted in gray. Percent cellularity of residual cancer for cases that converted to normal-like status after chemotherapy is bold.

Table 2. Top 10 biologic pathways from differentially expressed genes

	Overexpression in post-NST		Underexpression in post-NST	
	Canonical pathways	***P***	**Canonical pathways**	***P***
All cases ($n = 21$)				
1	PPAR signaling	0.0029	ATM signaling	0.0001
2	PPARα/RXRα activation	0.0032	Hypoxia signaling in the cardiovascular system	0.0005
3	Integrin signaling	0.0033	Cell cycle: G2/M DNA damage checkpoint regulation	0.0005
4	Nur77 signaling in T lymphocytes	0.0079	IL-1 signaling	0.0005
5	O-Glycan biosynthesis	0.0091	Sonic hedgehog signaling	0.0009
6	Circadian rhythm signaling	0.0120	D-glutamine and D-glutamate metabolism	0.0014
7	Estrogen receptor signaling	0.0141	PI3K/AKT signaling	0.0019
8	Wnt/β-catenin signaling	0.0155	IGF-1 signaling	0.0037
9	Aryl hydrocarbon receptor signaling	0.0174	Polyamine regulation in colon cancer	0.0044
10	Notch signaling	0.0195	Regulation of eIF4 and p70S6K signaling	0.0055
Basal-like cases ($n = 9$)				
1	CREB signaling in neurons	0.0032	Role of NFAT in regulation of the immune response	0.0007
2	TR/RXR activation	0.0035	Mismatch repair in eukaryotes	0.0046
3	Thrombin signaling	0.0044	Lipid antigen presentation by CD1	0.0065
4	Glutamate metabolism	0.0069	Systemic lupus erythematosus signaling	0.0095
5	Huntington's disease signaling	0.0072	Sonic hedgehog signaling	0.0148
6	Chondroitin sulfate biosynthesis	0.0117	Dendritic cell maturation	0.0166
7	CXCR4 signaling	0.0170	Role of CHK proteins in cell-cycle checkpoint control	0.0191
8	Propanoate metabolism	0.0170	Aminoacyl-tRNA biosynthesis	0.0209
9	Glioma invasiveness signaling	0.0178	Fcγ receptor-mediated phagocytosis in macrophages and monocytes	0.0214
10	GM-CSF signaling	0.0209	Natural killer cell signaling	0.0275
Non–basal-like cases ($n = 12$)				
1	Notch signaling	1.26E-05	IL-1 signaling	0.0001
2	Thrombopoietin signaling	0.0105	Androgen signaling	0.0002
3	ILK signaling	0.0162	BMP signaling pathway	0.0002
4	Hypoxia signaling in the cardiovascular system	0.0166	RAR activation	0.0002
5	Agrin interactions at neuromuscular junction	0.0182	Amyloid processing	0.0005
6	Caveolar-mediated endocytosis signaling	0.0219	Molecular mechanisms of cancer	0.0008
7	BMP signaling pathway	0.0219	Corticotropin releasing hormone signaling	0.0011
8	O-Glycan biosynthesis	0.0224	Cardiac hypertrophy signaling	0.0011
9	TNFR2 signaling	0.0224	Sonic hedgehog signaling	0.0012
10	Aryl hydrocarbon receptor signaling	0.0245	Melatonin signaling	0.0014

various immune-related pathways was particularly prominent in basal-like cancers (lipid antigen presentation by CD1, dendritic cell maturation, FC receptor–mediated phagocytosis and natural killer cell signaling).

To gain further insight into the molecular changes that occur in residual cancer specimens, we also carried out gene set analysis for 234 canonical biologic pathways and 2 gene signatures that were previously reported to be associated with residual disease after chemotherapy in breast cancer (18, 19). When all cases were analyzed together 7 gene sets were overexpressed and 12 were underexpressed at *P* value 0.05 or less, in post-NST samples (Table 3). When basal-like

Table 3. Gene set analysis of 236 gene sets between post- and pre-NST specimens ($P \leq 0.05$)[a]

	Gene sets	Number of genes	*P*	Elevated components
All samples ($n = 21$)				
1	Notch signaling	44	0.008	Post
2	Circadian rhythm signaling	37	0.021	Post
3	Fatty acid biosynthesis	17	0.032	Post
4	PPAR__RXR_ activation	261	0.036	Post
5	CXCR4 signaling	220	0.038	Post
6	Neurotrophin_TRK signaling	101	0.040	Post
7	O-Glycan biosynthesis	26	0.048	Post
1	Sonic hedgehog signaling	36	0.002	Pre
2	p38 MAPK signaling	135	0.009	Pre
3	Parkinson's signaling	19	0.019	Pre
4	Lysine biosynthesis	8	0.020	Pre
5	Nicotinate and nicotinamide metabolism	120	0.025	Pre
6	Role of PKR in IFN induction and antiviral response	69	0.025	Pre
7	Methionine metabolism	31	0.033	Pre
8	IL-10 signaling	87	0.039	Pre
9	Natural killer cell signaling	144	0.046	Pre
10	Role of RIG1-like receptors in antiviral innate immunity	54	0.049	Pre
11	B Cell receptor signaling	228	0.050	Pre
12	IL-15 signaling	85	0.050	Pre
Basal samples ($n = 9$)				
1	CXCR4 signaling	220	0.005	Post
2	Thrombin signaling	268	0.010	Post
3	Cardiac hypertrophy signaling	302	0.014	Post
4	Fatty acid biosynthesis	17	0.024	Post
5	Ascorbate and aldarate metabolism	20	0.027	Post
6	Propanoate metabolism	82	0.031	Post
7	Neurotrophin_TRK signaling	101	0.039	Post
8	TR_RXR activation	119	0.039	Post
9	IGF-1 signaling	138	0.045	Post
10	Alanine metabolism	55	0.046	Post
1	Lysine biosynthesis	8	0.001	Pre
2	Natural killer cell signaling	144	0.001	Pre
3	Fc Epsilon RI signaling	134	0.002	Pre
4	TREM1 signaling	71	0.002	Pre
5	B Cell Receptor Signaling	228	0.005	Pre
6	Role of NFAT in regulation of the immune response	272	0.006	Pre
7	Fc Receptor-mediated phagocytosis in macrophages and monocytes	153	0.007	Pre
8	IL-10 signaling	87	0.012	Pre
9	Dendritic cell maturation	188	0.018	Pre
10	p38 MAPK signaling	135	0.018	Pre
11	CTLA4 signaling in cytotoxic T lymphocytes	135	0.027	Pre
12	Sonic hedgehog signaling	36	0.027	Pre
13	IL-15 signaling	85	0.028	Pre
14	Role of PKR in IFN induction and antiviral response	69	0.030	Pre

(Continued on the following page)

Table 3. Gene set analysis of 236 gene sets between post- and pre-NST specimens ($P \leq 0.05$)[a] (Cont'd)

	Gene sets	Number of genes	*P*	Elevated components
15	IL-8 signaling	217	0.033	Pre
16	Fc_RIIB signaling in B lymphocytes	54	0.037	Pre
17	CD28 signaling in T helper cells	199	0.041	Pre
18	T Helper cell differentiation	57	0.044	Pre
19	Cytotoxic T lymphocyte-mediated apoptosis of target cells	51	0.048	Pre
20	NF-_B Signaling	183	0.048	Pre
Nonbasal samples ($n = 12$)				
1	Circadian rhythm signaling	37	0.005	Post
2	Notch signaling	44	0.006	Post
3	O-Glycan biosynthesis	26	0.011	Post
4	Chondroitin sulfate biosynthesis	38	0.017	Post
5	Caveolar-mediated endocytosis	119	0.027	Post
6	Thrombopoietin signaling	84	0.029	Post
7	Integrin signaling	280	0.034	Post
8	Actin cytoskeleton signaling	267	0.035	Post
9	Fatty acid biosynthesis	17	0.041	Post
10	PPAR signaling	124	0.043	Post
11	Semaphorin signaling in neurons	73	0.044	Post
12	Sonic hedgehog signaling	36	0.003	Pre
13	Role of RIG1-like receptors in antiviral innate immunity	54	0.020	Pre
14	Phototransduction pathway	49	0.026	Pre
15	BMP signaling pathway	116	0.031	Pre
16	Airway inflammation in asthma	3	0.044	Pre

[a]*P* value was calculated by Efron–Tibshirani's GSA test (under 1,000 times permutations); overlapping gene sets between basal and nonbasal are highlighted.

cancers were examined separately, again numerous metabolic pathways emerged as upregulated in residual cancers and various immune pathways dominated the downregulated gene sets. In non–basal-like cancers, 11 gene sets were overexpressed and 5 gene sets were underexpressed in post-NST ($P \leq 0.05$). The results were broadly consistent with the pathways that were derived from differentially expressed genes. Fatty acid biosynthesis was the only common upregulated pathway in both basal- and non–basal-like residual cancers, and the sonic hedgehog signaling was the only common downregulated pathway (Table 3, Fig. 2). None of the 2 previously reported residual cancer–associated gene sets showed any significant enrichment before or after therapy in any disease subset (Supplementary Table S6).

It has been suggested that the expression of genes involved with epithelial–mesenchymal transition (EMT) may be altered in cancers that survived chemotherapy (19). To test this hypothesis, we compared the expression of 8 key EMT genes (*CD44*, *PCNA*, *CDH1*, *SNAIL*, *SLUG*, *TWIST*, *SOX9*, and *TGFβ*) in pre- and post-NST samples. When all cancers ($n = 21$) were examined together (Table 4) only *SLUG*, *SOX9*, and *TWIST* were significantly overexpressed ($P = 0.029$, 0.001, and 0.004, respectively, unadjusted for multiple comparisons), and *SNAI1* was significantly underexpressed ($P = 0.001$) in post-NST samples. When basal-like and non–basal-like cancers were examined separately, *CD44*, *SLUG SOX9*, and *TWIST* were significantly overexpressed ($P = 0.012$, 0.050, 0.010, and 0.028, respectively), and *SNAI1* was significantly underexpressed ($P = 0.003$) in residual post-NST non–basal-like cancers. In basal-like cancers no gene was significantly altered (Table 4).

Discussion

Our aim was to determine gene expression differences between primary tumor and corresponding residual cancer after NST. We conducted gene expression profiling on baseline FNAs and surgical resections from residual cancer after 4 to 6 months of cytotoxic chemotherapy. Any such analysis strategy is fraught not only by the limited number of tumors analyzed but also with technical challenges that increase the probability for false discovery. In our case, baseline biopsies were obtained with an FNA and residual cancers represented surgically resected tissues. Tissue sampling methods can have a profound effect on gene

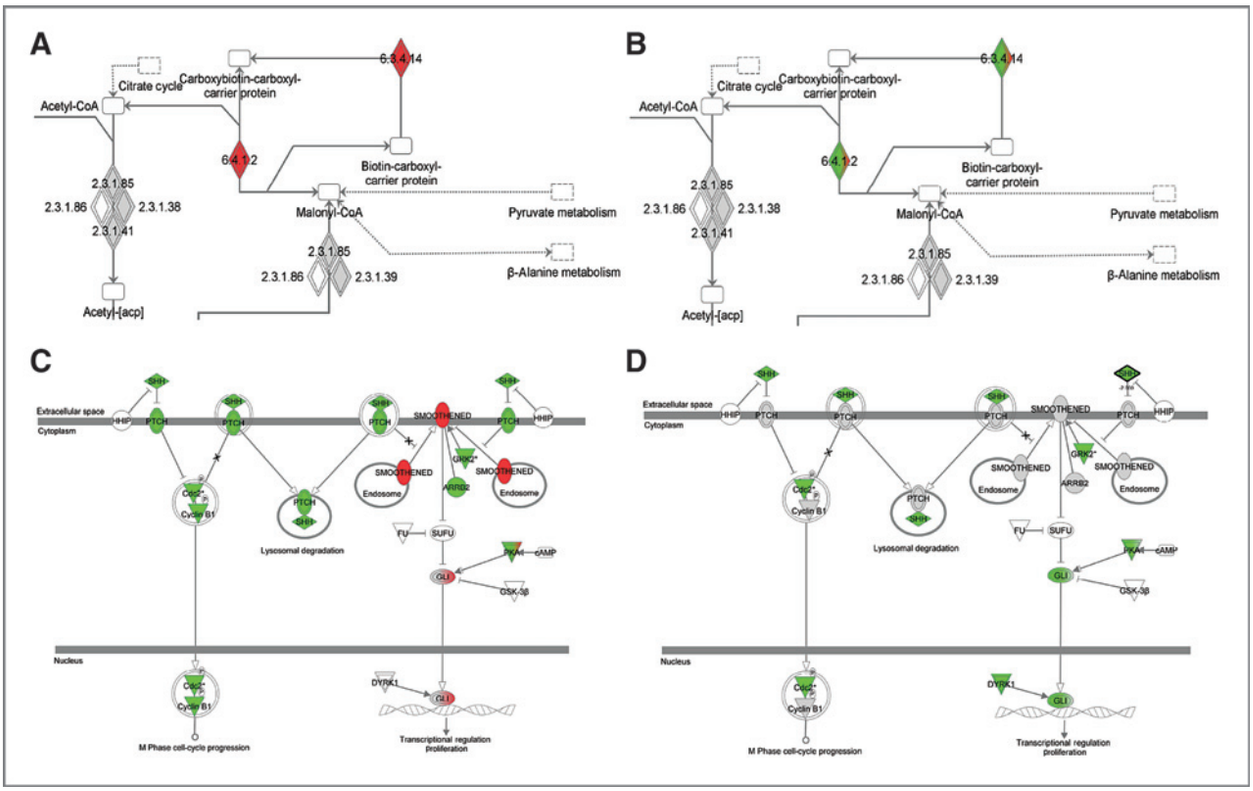

Figure 2. Ingenuity Pathway Maps of the fatty acid biosynthesis gene set that was enriched in both the basal-like (A) and non–basal-like residual cancers (B) and the Sonic Hedgehog signaling gene set that was depleted in basal-like (C) as well as non–basal-like (D) residual cancers. Red, genes overexpressed in residual samples; green, genes underexpressed in residual samples; grey, genes with no statistically significant differential expression ($P < 0.01$); white, pathway member not available in data. Mixed colors represent variable association for the same gene depending on probe set (6.3.4.14, biotin carboxylase; 6.4.1.2, acetyl-CoA carboxylase).

expression results from the same cancer (13). Also, chemotherapy alters tumor cellularity in most cancers shifting tissue composition toward stromal and fibrotic components in post-NST samples (5). These effects plus the inevitable technical noise in repeat gene expression profiling can mask real biologic changes in mRNA expression levels caused by NST in small studies. We tried to minimize tissue composition-related confounders by removing 1,618 probe sets that we previously found to be highly associated with tissue sampling and stromal "contamination" (10). We also filtered the lowest expressed 25% of probe sets because these may be the most susceptible to technical noise in repeat measurements (21).

After controlling for these confounders, ER and HER2 status based on mRNA expression remained stable in all but 2 cases (10%). There were also no significant changes in 2 single gene proliferation metrics, the expression of Ki67 and PCNA. Conversely, molecular class changed in 8 cases (38%) mostly to normal-like subtype. This shift to normal-like in post-NST specimens was observed in cancers with the lowest cancer cell cellularity in residual disease and probably reflects a major shift in tumor composition. Because pretreatment FNAs contain mostly neoplastic cells and are devoid of stroma, our analysis is biased to detect increase in stromal components in the surgical specimens even after our best attempt to control for this variable (none of the PAM50 classifier genes were included in our stromal associated and therefore filtered gene list; ref. 22). Nevertheless, these results indicate that molecular classification is susceptible to tissue sampling effects, and this should be considered when interpreting results from any repeat measurement of a tumor.

Using a stringent FDR threshold of 0.005, we found 200 to 600 probe sets differentially expressed between baseline and post-NST samples. When corresponding genes were mapped to biologic pathways, some interesting observations emerged. In basal-like cancers, pathways driven by PI3K, small G proteins, CAMK2, and several energy metabolism and biosynthetic pathways were enriched in post-NST samples, whereas immune cell–derived pathways and the sonic hedgehog pathway were depleted. The same broad picture emerged when gene set analysis was applied to the data. In non–basal-like breast cancers, notch signaling and also energy metabolism through fatty acid synthesis were upregulated and sonic hedgehog signaling as well as several immune-related pathways were depleted in residual cancers. However, due to the small sample size of this study, the variable cytotoxic therapies that these patients received and

Table 4. Paired sample *t* test for EMT-related genes

			All ($n = 21$)			Nonbasal ($n = 12$)			Basal ($n = 9$)		
Symbol	Name	ProbeSet	*P*	FDR	Ratio (Post/pre)	*P*	FDR	Ratio (Post/pre)	*P*	FDR	Ratio (Post/pre)
CD44	CD44 molecule (Indian blood group)	212063_at	0.154	0.283	1.03	0.012	0.043	1.07	0.507	0.797	1.07
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	201131_s_at	0.581	0.711	1.02	0.491	0.675	1.03	0.993	0.993	1.03
SLUG	Snail homolog 2 (*Drosophila*)	213139_at	0.029	0.065	1.12	0.050	0.110	1.17	0.371	0.681	1.17
SNAI1	Snail homolog 1 (*Drosophila*)	219480_at	0.001	0.010	0.80	0.003	0.037	0.80	0.076	0.247	0.80
SOX9	SRY (sex determining region Y)-box 9	202935_s_at	0.002	0.011	1.08	0.010	0.043	1.09	0.107	0.247	1.09
TGFB1	Transforming growth factor, beta 1	203085_s_at	0.848	0.853	1.00	0.654	0.797	1.01	0.912	0.993	1.01
TWIST	Twist homolog 1 (*Drosophila*)	213943_at	0.004	0.015	1.11	0.028	0.077	1.12	0.090	0.247	1.12

the possible tissue sampling related confounders, these observations remain hypothesis generating until confirmed in repeated observations and in larger studies. We also note that in our analysis, we could not observe an increase in EMT-related features or cancer stem cell signatures in residual cancers, as suggested previously (18, 19).

Two strong themes emerged from this analysis with potential therapeutic implications, the enrichment of energy- and metabolism-related processes in residual cancers that were most prominent in basal-like cancers and the depletion of immune-related signals. There is increasing interest in deciphering the deregulation of energy metabolism in cancer and attempts are being made to develop metabolism-targeted drugs (23). Our data suggests that in response to cytotoxic therapy several metabolic pathways become upregulated in residual cancer tissue. We hypothesize that the increased metabolic activity may enable cells to survive treatment and targeting these processes may enhance the efficacy of standard chemotherapy regimens. It is also increasingly evident that the presence of immune cells in the tumor microenvironment carries prognostic value, particularly between ER-negative and highly proliferative ER-positive cancers, and is also associated with higher probability of pCR to NST (10, 24–26). Because residual cancer tissue becomes depleted of immune-related transcriptional signals, probably due to the systematic immunosuppressive nature of chemotherapy, stimulating post-NST immune response may become a fruitful adjuvant therapeutic strategy for patients with residual cancer and high risk of recurrence.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

This work was supported in part by 1K23CA121994-01 (A.M. Gonzalez-Angulo), ASCO Career Development Award (A.M. Gonzalez-Angulo), Komen for the Cure Catalystic Award KG090341 (A.M. Gonzalez-Angulo), KG081099 (A.M. Gonzalez-Angulo, G.B. Mills), American Cancer Society Research Scholar Grant (A.M. Gonzalez-Angulo), The Breast Cancer Research Foundation (L. Pusztai and W.F. Symmans), and National Cancer Institute through The University of Texas MD Anderson's Cancer Center Support grant (P30 CA016672).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 27, 2011; revised December 12, 2011; accepted January 2, 2012; published OnlineFirst January 10, 2012.

References

1. Wolff AC, Berry D, Carey LA, Colleoni M, Dowsett M, Ellis M, et al. Research issues affecting preoperative systemic therapy for operable breast cancer. J Clin Oncol 2008;26:806–13.
2. Kuerer HM, Newman LA, Smith TL, Ames FC, Hunt KK, Dhingra K, et al. Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. J Clin Oncol 1999;17:460–9.
3. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 2008;26:1275–81.
4. Guarneri V, Broglio K, Kau SW, Cristofanilli M, Buzdar AU, Valero V, et al. Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors. J Clin Oncol 2006;24:1037–44.

5. Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 2007;25:4414–22.
6. O'Regan RM, Von Roenn JH, Carlson RW, Malik U, Sparano JA, Staradub V, et al. Final results of a phase II trial of preoperative TAC (Docetaxel/doxorubicin/cyclophosphamide) in stage III breast cancer. Clin Breast Cancer 2005;6:163–8.
7. Stearns V, Singh B, Tsangaris T, Crawford JG, Novielli A, Ellis MJ, et al. A prospective randomized pilot study to evaluate predictors of response in serial core biopsies to single agent neoadjuvant doxorubicin or paclitaxel for patients with locally advanced breast cancer. Clin Cancer Res 2003;9:124–33.
8. Andre F, Mazouni C, Liedtke C, Kau SW, Frye D, Green M, et al. HER2 expression and efficacy of preoperative paclitaxel/FAC chemotherapy in breast cancer. Breast Cancer Res Treat 2009;108:183–90.
9. Buzdar AU, Ibrahim NK, Francis D, Booser DJ, Thomas ES, Theriault RL, et al. Significantly higher pathologic complete remission rate after neoadjuvant therapy with trastuzumab, paclitaxel, and epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2-positive operable breast cancer. J Clin Oncol 2005;23:3676–85.
10. Bianchini G, Qi Y, Alvarez RH, Iwamoto T, Coutant C, Ibrahim NK, et al. Molecular anatomy of breast cancer stroma and its prognostic value in estrogen receptor-positive and -negative cancers. J Clin Oncol 2010; 28:4316–23.
11. Symmans WF, Ayers M, Clark EA, Stec J, Hess KR, Sneige N, et al. Total RNA yield and microarray gene expression profiles from fine needle aspiration and core needle biopsy samples of breast cancer. Cancer 2003;97:2960–71.
12. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Different molecular subtypes of breast cancer respond differently to preoperative chemotherapy. Clin Cancer Res 2005;11:5678–85.
13. Stec J, Wang J, Coombes K, Ayers M, Hoersch S, Gold DL, et al. Comparison of the predictive accuracy of DNA array based multigene classifiers across cDNA arrays and Affymetrix GeneChips. J Mol Diagn 2005;7:357–67.
14. Johnson WE, Rabinovic A, Li C. Adjusting batch effects in microarray expression data using Empirical Bayes methods. Biostatistics 2007;8:118–27.
15. Gong Y, Yan K, Lin F, Anderson K, Sotiriou C, Andre F, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. Lancet Oncol 2007;8:203–11.
16. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2010;27:1160–7.
17. Efron B, Tishirani R. On testing the significance of sets of genes. Ann Appl Stat 2007;1:107–29.
18. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proc Natl Acad Sci U S A 2009;106:13820–5.
19. Mego M, Mani SA, Lee BN, Li C, Evans KW, Cohen EN, et al. Expression of epithelial-mesenchymal transition-inducing transcription factors in primary breast cancer: The effect of neoadjuvant therapy. Int J Cancer 2012;130:808–16.
20. Jiang Z, Gentleman R. Extensions to gene set enrichment. Bioinformatics 2007;23:306–13.
21. Anderson K, Hess KR, Kapoor M, Tirrell S, Courtemanche J, Wang B, et al. Reproducibility of gene expression signature-based predictions in replicate experiments. Clin Cancer Res 2006;12:1721–7.
22. Symmans WF, Ayers M, Clark EA, Stec J, Hess KR, Sneige N, et al. Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma. Cancer 2003;97:2960–71.
23. Denkert C, Loibl S, Noske A, Roller M, Muller BM, Komor M, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol 2010;28:105–13.
24. Rody A, Holtrich U, Pusztai L, Liedtke C, Gaetje R, Ruckhaeberle E, et al. T-cell metagene predicts a favorable prognosis in estrogen receptor negative and HER2 positive breast cancers. Breast Cancer Res 2009;11:R15.
25. Karn T, Pusztai L, Ruckhaeberle E, Liedtke C, Muller V, Schmidt M, et al. Melanoma antigen family A identified by the bimodality index defines a subset of triple negative breast cancers as candidates for immune response augmentation. Eur J Cancer 2012; 48:12–23.
26. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer 2011;11:85–95.