Genome and Transcriptome Sequencing in Prospective Metastatic Triple-Negative Breast Cancer Uncovers Therapeutic Vulnerabilities

David W. Craig1, Joyce A. O'Shaughnessy2, Jeffrey A. Kiefer1, Jessica Aldrich1, Sheripad Sinari1, Tracy M. Moses1, Shukmei Wong1, Jennifer Dinh1, Alexis Christoforides1, Joanne L. Blum3, Cristi L. Altelli3, Cynthia R. Osborne4, Tyler Izzat1, Ahmet Kurdoglu1, Angela Baker1, Julie Koeman5, Catalin Barbacioru6, Anthony W. Tolcher4, Jeffrey M. Trent1, Spyro Mousses1, Daniel Von Hoff1, and John D. Carpten1

Abstract

Triple-negative breast cancer (TNBC) is characterized by the absence of expression of estrogen receptor, progesterone receptor, and HER-2. Thirty percent of patients recur after first-line treatment, and metastatic TNBC (mTNBC) has a poor prognosis with median survival of one year. Here, we present initial analyses of whole genome and transcriptome sequencing data from 14 prospective mTNBC. We have cataloged the collection of somatic genomic alterations in these advanced tumors, particularly those that may inform targeted therapies. Genes mutated in multiple tumors included TP53, LRP1B, HERC1, CDH5, RB1, and NF1. Notable genes involved in focal structural events were CTNNAI, PTEN, FBXW7, BRCA2, WT1, FGFR1, KRAS, HRAS, ARAF, Braf, and PGCP. Homozygous deletion of CTNNAI was detected in 2 of 6 African Americans. RNA sequencing revealed consistent overexpression of the FOXM1 gene when tumor gene expression was compared with nonmalignant breast samples. Using an outlier analysis of gene expression comparing one cancer with all the others, we detected expression patterns unique to each patient’s tumor. Integrative DNA/RNA analysis provided evidence for deregulation of mutated genes, including the monoallelic expression of TP53 mutations. Finally, molecular alterations in several cancers supported targeted therapeutic intervention on clinical trials with known inhibitors, particularly for alterations in the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways. In conclusion, whole genome and transcriptome profiling of mTNBC have provided insights into somatic events occurring in this difficult to treat cancer. These genomic data have guided patients to investigational treatment trials and provide hypotheses for future trials in this irremediable cancer. Mol Cancer Ther; 12(1); 104–16. ©2012 AACR.

Introduction

Triple-negative breast cancer (TNBC) is generally defined by histopathologies lacking expression of estrogen receptor (ER), progesterone receptor (PR), and HER-2 (1, 2). TNBC are believed to account for approximately 15% of breast cancer diagnoses, but approximately 25% of breast cancer-related deaths due to a more aggressive biology (1, 2). There is evidence that TNBC is more prevalent among young premenopausal African American women, and women of recent West-African descent (3–6). Furthermore, gene expression profiling has classified breast cancer into intrinsic subtypes (7, 8), among which is the basal-like subtype, representing ER/PR-negative with low HER2-neu–expressing tumors that are characterized as TNBC. Further investigations using gene expression data have fortified the heterogeneous nature of TNBC (9–12).

Although significant progress has been made toward treatment of Luminal A&B (hormone receptor–positive), and HER2-neu–enriched subtypes of breast tumors due to receptor target expression, women diagnosed with TNBC have cancers that lack receptor targets, and therefore therapeutic options are limited (13). Median survival in the approximate 30% of patients with TNBC who develop metastatic disease is 1 year.

Besides array-based gene expression analysis, a number of studies have reported genomic alterations that...
occur in TNBC clinical specimens and cell lines including comparative genomic hybridization and deep genomic profiling using next generation sequencing (NGS) technologies (14–20). Whole-genome sequencing of a single metastatic TNBC (mTNBC) patient’s germline, primary tumor, metastatic tumor, and xenograft has also been reported, which showed the complexity of the somatic events that arise within a given TNBC (19). More recently, genome-sequencing studies of large subsets of retrospectively collected TNBC have been reported, which implicate TP53, PIK3CA, NRAS, EGFR, RB1, and PTEN (20, 21). These deep exome-sequencing studies have also helped to uncover tumor heterogeneity within primary TNBC.

Here, we have applied deep whole genome and transcriptome sequencing to uncover somatic mutations occurring in a cohort of prospectively collected relapse mTNBC clinical specimens from 14 patients during the course of their clinical management. Our dataset includes 6 women who self-identified as African American and 8 who self-identified as European American. We have confirmed genomic alterations previously described in TNBC and have discovered novel, previously unreported mutations. Importantly, we have identified specific events and integrated concepts, which strongly infer both obvious and nonobvious targeted therapeutics that might be effective for treating subsets of mTNBC tumors, particularly agents targeting the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR axes.

Materials and Methods

Samples and DNA/RNA isolation

Samples were obtained following written informed content with US Oncology Central Institutional Review Board approval. High molecular weight DNA and total RNA were extracted from fresh frozen tumor tissue at Caris Diagnostics using the Qiagen All Prep system (Qiagen). No cell lines were used in this study.

Next generation sequencing

All NGS was carried out using the SOLiD version 4.0 system, long (1.5 kb) 50 bp × 50 bp mate-pair chemistry (Applied Biosystems by Life Technologies). SOLiD sequencing was conducted using the manufacturer’s recommendations.

Genome data analysis

Raw SOLiD NGS data (csfasta and qual files) were aligned against the reference human genome (NCBI Build 36, hg18) using Life Technologies BioScope version 1.3 software suite based on a seed-and-extend algorithm (22).

Compressed binary sequence alignment/map (BAM) formatted output files for germline and tumor genome alignments were generated and PCR duplicates subsequently removed using the Picard Tools (23). We used the algorithms (SolSNP; ref. 24) and MutationWalker to detect single-nucleotide variants (SNV). MutationWalker calculates a test of proportions for the tumor/normal set to construct a test-statistic for reads in the forward direction and the reverse direction separately. The minimum of these 2 comparisons is used as the reported test-statistic, insuring evidence is found in both the forward and reverse detection. Sites with evidence in the normal are filtered from the final report to reduce false-positives arising from undersampled polymorphic germline events. Calls common to both algorithms were considered for further examination. To reduce false-negative rate 2 sets of common calls were made. One was made with a strict and the other with lenient set of parameters for both the algorithms. Both the sets were visually examined for false-positives, which were then filtered to get a final list of true SNVs.

For detecting somatic indels, we used a 2-step strategy. In the first step, we removed from the tumor sample bam, reads whose insert size lay outside the interval 500 to 5,000 bp. Genome Analysis Toolkit (GATK; ref. 25) was then used to generate a list of potential small indels from this bam. A customized perl script, which used the Bio-SamTools library from BioPerl (26), then took these indel positions and for each of the indels assessed the region in the germline data consisting of 5 bases upstream from the start and 5 bases downstream from the end of the indel.

For copy number analysis, gains and losses were determined by calculating the log2 difference in normalized coverage between tumor and germline. We investigated regions in 100 bp windows in which the coverage in the germline was between 0.1 and 10 of the mode coverage to remove regions with high degree of repeat sequence. Normalized coverage was determined by the log2 coverage within a 100 bp bin over the overall modal coverage. We then reported the difference between the germline- and tumor-normalized coverage by a sliding window of size 2 kb.

A series of customized perl scripts were used in the detection of translocation. The genome was analyzed by a walker with step size equivalent to the insert size in which the number of anomalous reads were counted, reads whose mates align on a different chromosome. Outlier detection was done under the assumption that the normal distribution, of the proportion of hit discordant reads in 2 kb windows aggregated across the chromosome, will follow a normal distribution. We then computed the mean of the distributions and chose a cutoff of 3 SDs.

Transcriptome data analysis

SOLiD BioScope Whole Transcriptome Analysis (WTA) version 1.2.1 pipeline for single reads was used to align the reads, count aligned reads per exon, and calculate per base coverage. WTA pipeline CountTags module provides normalize RPKM (reads per kilobase of exon sequence, per million reads) values along with read counts per exon.

We used edgeR, a bioconductor package specifically for the differential expression analysis on the counts per gene from the tumor and normal samples (27). Gene counts were normalized using edgeR’s TMM [trimmed mean M (= log fold-change gene expression)] algorithm.
(28). For this study, the normal samples’ technical replicates were merged and the analysis was done on the biologic replicates. Gene expression was corrected using a moderated binomial dispersion correction then an exact test (similar to Fisher exact test) was used to assess differential expression. edgeR’s exact test produced a table of log concentrations, log fold-changes, P values, and false discovery rate (FDR). We also used RNA-seq data to independently conduct an outlier analysis to assess differential expression patterns for each tumor when compared against the other tumors within the cohort.

Knowledge mining
Therapeutic contexts were identified using a knowledge mining workflow on the combined whole genome and RNA-seq data on each individual patient. Prior knowledge curated from literature sources coupled with the use of online commercial databases and pathway tools were used. The latter resources included NextBio, Ingenuity IPA (Ingenuity Systems), and Metacore. The initial step consisted of cross-referencing–known therapeutic contexts and biomarkers identified previously. Drug candidates were also identified by target annotation of NGS data through IPA and Metacore databases. The main portal used to identify biologic information on individual genes of interest was NextBio’s genomic apps and literature search tool. This allowed for the rapid curation of gene-specific information represented in literature and placement of particular genes among various public data resources. In addition to binary matching of targets/biomarkers to drugs, we overlaid data onto signaling maps in IPA and Metacore to visualize the data within the context of known biologic signaling networks and concepts. This facilitated the biologic and therapeutic interpretation of the combined data.

PCR and Sanger sequencing
Methods used for PCR and Sanger sequencing are previously described (29). Primer sequences for PCR of DNA and cDNA for RB1 are available upon request.

Results
Integrated molecular profiling of mTNBC using NGS
The cohort consisted of 14 women clinically diagnosed with mTNBC (Supplementary Table S1). All tumor samples were sent to a laboratory that was compliant with the Clinical Laboratory Improvements Amendment (CLIA) regulations for sample quality control and analyte extraction, and for later validation of therapeutically actionable molecular alterations when applicable. All tumor specimens used for analyte extraction contained more than 50% tumor cellularity, with an average of 71% (range 50%–95%; Supplementary Table S1).

RNA-seq statistics can be found in Supplementary Table S2. We used edgeR (27) for differential gene expression analyses, in which RNA-seq data were compared for each tumor against RNA-seq data from ethnicity-matched nonmalignant samples (Supplementary Table S3). Individual tumor versus reference differential expression data were used to conduct pathway and gene ontology analysis on genes with log fold-change >2 (P < 0.001) using IPA (Ingenuity Systems). The most significant IPA canonical pathways for each tumor are provided in Supplementary Table S4. These analyses revealed that 8 of 14 tumors show expression profiles enriched with genes involved in cell-cycle control, G2–M checkpoint regulation, and mitosis, most indicative of the basal-like subtype of TNBC. Of those genes consistently overexpressed among these tumors is the FOXM1 gene. Moreover, 5 of 14 patients contained expression profiles with ontologies enriched for immune-related pathway genes (i.e., TREM1 signaling, primary immunodeficiency signaling, altered T cell, and B cell signaling in rheumatoid arthritis). Finally, a single patient (mTNBC8) showed a primary ontology enriched with genes involved in androgen and ER metabolism, and pentose and glucuronate metabolism.

We also independently conducted an outlier analysis to assess differential expression patterns for each tumor when compared against the other tumors within the cohort. Specific cancer genes (defined by Catalog of Somatic Mutations in Cancer frequently mutated genes or within Kyoto Encyclopedia of Genes and Genomes cancer pathways) differentially overexpressed among tumors in our cohort included ALK, AR, ARAF, BRAF, FGFR2, GLI1, GLI2, HRAS, HSP90AA1, KRAS, MET, NOTCH2, NOTCH3, and SHH, whereas significantly underexpressed cancer genes included BRCA1, BRCA2, CDK4R2, CTNNB1, DKK1, FBXW7, NF1, PTEN, and SFN (Supplementary Table S5). Importantly, this outlier analysis provided insights into the potential unique therapeutic vulnerabilities of each cancer.

Genome sequence coverage typically was more than 30× for both tumor and germline genomes (Supplementary Table S6). Somatic mutations including SNVs, indels, translocations, intrachromosomal rearrangements (inversions, etc.), and copy number alterations were determined from sequencing of tumor and germline pairs. Total numbers of somatic SNVs and structural variants are shown graphically by sample in Fig. 1. Leveraging RNA-seq data, we also assessed the expression of somatic SNVs at the transcript level using RNA-seq data, by calculating the number of mutations that were present in both DNA and RNA versus those only present in DNA. In this analysis, we took all data points including where a gene harbored a somatic SNV, but the gene was not expressed on the RNA level. The mean percentage of somatic SNVs that were expressed across our cohort of 14 mTNBC was 34% (range 22%–60%).

The list of somatic point mutations and small indels discovered is provided in Supplementary Table S5. TP53 mutations were most frequent in our mTNBC dataset, initially called in 7 of 14 mTNBC tumors (Table 1). Integrated analysis of TP53 along side RNA identified an additional 3 samples that showed strong evidence of a mutation in mRNA (more than 5 reads). The TP53 R273H SNV was the only recurring mutation in our
cohort. In 6 of 10 TP53 mutated tumors, LOH was detected without copy number loss at the TP53 locus, suggesting uniparental disomy as a common mechanism for TP53 LOH in mTNBC. Extending to a genome-wide analysis of allele-specific expression between the tumor RNA and DNA, TP53 was the only gene with more than 2 somatic mutations exhibiting significant transcriptional allelic imbalance. Integrating RNA-seq data revealed that the frequency of the mutated allele exhibited transcriptional allelic imbalance, in which the mutated allele was expressed at approximately 90% in 7 of 9 tumors with exon-coding mutations (Table 1). Importantly, this was observed in the face of upward of more than 30% normal contaminating stroma in some samples used for analyte extraction (Supplementary Table S1). Fifteen additional genes harbored somatic nonsynonymous SNVs in more than 1 tumor including CDH5 (n = 3), HERC1 (n = 2), LRP1B (n = 2), and TOP2A (n = 2; Supplementary Table S7).

Among genes containing somatic coding deletions in more than 1 tumor were RB1 and PTEN, both of which are known tumor suppressor genes. Importantly, combined genome and transcriptome data allow for immediate interpretation of the consequence of deletions. One example was an RB1 39 bp deletion in mTNBC1 that includes 37 bases of exon 12, and extends to also delete the 2

Table 1. TP53 somatic mutations and allelic expression

<table>
<thead>
<tr>
<th>Sample</th>
<th>TP53 tumor SNVs</th>
<th>SNV freq. tumor DNA</th>
<th>SNV freq. RNA</th>
<th>DNA LOH</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTNBC1</td>
<td>None detected</td>
<td>NA</td>
<td>NA</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC2</td>
<td>K320X</td>
<td>40%</td>
<td>100%</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC3</td>
<td>R273H*</td>
<td>50%</td>
<td>92%</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC4</td>
<td>Exon2 splice</td>
<td>50%</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC5</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>mTNBC6</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC7</td>
<td>E339X</td>
<td>47%</td>
<td>57%</td>
<td>No</td>
</tr>
<tr>
<td>mTNBC8</td>
<td>L111R*</td>
<td>23%</td>
<td>85%</td>
<td>No</td>
</tr>
<tr>
<td>mTNBC9</td>
<td>H193R</td>
<td>53%</td>
<td>83%</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC10</td>
<td>R81X</td>
<td>13%</td>
<td>20%</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC11</td>
<td>None detected</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>mTNBC12</td>
<td>R273H*</td>
<td>21%</td>
<td>92%</td>
<td>No</td>
</tr>
<tr>
<td>mTNBC13</td>
<td>Y220C</td>
<td>80%</td>
<td>85%</td>
<td>Yes</td>
</tr>
<tr>
<td>mTNBC14</td>
<td>R273H</td>
<td>21%</td>
<td>92%</td>
<td>No</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not available.

*Not detected by mutation caller although DNA evidence below the calling threshold was found in mTNBC8, mTNBC3, and mTNBC12.
nucleotides (GT) that define the conserved splice acceptor site (Supplementary Fig. S1). PCR and Sanger sequencing were conducted to validate the somatic DNA mutation and the resulting mutated cDNA confirming an in-frame splicing event of RB1 exon 11 to 13 (Supplementary Fig. S1). Importantly, this exon skipping results in an in-frame RB1 transcript but with 44 amino acids deleted within the conserved RB domain. However, this deleted RB1 transcript is expressed at normal levels by differential expression analysis of RNA-seq data.

Large structural changes across the genomes of mTNBC tumors were also assessed using genome-sequencing data (shown schematically in Fig. 2). Among the focal (15 Mb) homozygous deletions that occurred in more than 1 tumor were unique homozygous deletions in 2 tumors (mTNBC1, mTNBC6) that involved the adjacent CTNNA1 and SIL1 loci at 5q31.2 (Fig. 3). Integration of RNA-seq differential expression analyses using both comparison against nonmalignant samples and using our outlier analysis showed significant downregulation of CTNNA1 but not SIL1 in both tumors (Fig. 3). Interestingly, these deletions occurred in 2 of 6 tumors among African American patients.

Focal amplifications were also detected that encompass important oncogenes including WT1/WIT1, IRS2, MYC, WHSC1L1/FGFR1, MYB, PIK3CA, I,QGAP3, KRAS, HRAS, and ARAF (Supplementary Table S8). Of particular interest was a 1.5 Mb region detected as a copy number amplification (log2FC = 2.2) at 7q34 in the tumor from mTNBC2 that encompassed only 4 genes, including the BRAF locus. RNA-seq data showed increased BRAF expression when comparing against nonmalignant controls (log2FC = 1.7) and when comparing across tumors using the outlier analysis (log2FC = 2.2) in which this was the only tumor showing overexpression of BRAF. Deconvolution of mate-pair data revealed that this amplicon was part of a more complex rearrangement, likely a circular extrachromosomal double minute that includes chromosome 7 regions encompassing the BRAF oncogene along with genomic material from chromosomes 1 and 12 (Fig. 4). We conducted interphase FISH on paraffin sections from this and other tumors using a bacterial artificial chromosome (BAC) clone containing the BRAF locus and were able to validate the presence of amplified double minutes containing BRAF (Fig. 4). To our knowledge, this is the first report of BRAF amplification and double minutes in a TNBC tumor, the findings of which, may have therapeutic implications.

In addition, we detected distinct somatic alterations at the ERBB4 locus in 3 of 14 mTNBC. These include a 4 kb intronic homozygous deletion (mTNBC1), a somatic point mutation (mTNBC2), and a breakpoint defining a larger 21 Mb rearrangement at 2q34-q37.1 (mTNBC6). We also

Figure 2. Integrated genome-wide comparative genomic hybridization and structural rearrangements in mTNBC. Chromosome and genome positions are plotted across the X-axis, and log2 FC is plotted on the Y-axis. Copy number plots are colored red for amplified regions (log2 fold-change > 1.5) or green (log2 fold-change < 1). Blue lines show connection of trans regions based on discordant sequence read pairing.
observed significant downregulation of \( \text{ERBB4} \) in all mTNBC in our study, when comparing mTNBC tumors against nonmalignant controls. These data support frequent \( \text{ERBB4} \) deregulation in mTNBC both at the DNA and RNA levels. Furthermore, the \( \text{PGCP} \) gene was altered by multiple mechanisms in 3 different tumors, including 2 independent translocations, and represents the only gene involved in somatic translocation events in more than 1 tumor (Supplementary Table S9).

### Potential therapeutic targets in mTNBC

To facilitate genomic interpretation to uncover potential therapeutically relevant events, we developed and used a semiautomated knowledge mining workflow. For this study, the system was based on online commercial databases and pathway tools including NextBio, Ingenuity IPA (Ingenuity Systems), and Metacore, leveraging prior knowledge curated from the literature. Table 2 provides information on potentially informative genomic alterations within our cohort of patients. Clinical treatment histories are provided for each patient in Table 3. Of particular interest were alterations that converged on the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways in upward of 9 of 14 patients (Fig. 5).

Among the events detected in mTNBC1 were a single exon frameshift deletion within \( \text{PTEN} \), and a region of focal copy number gain (log\( _2 \)FC = 1.4) encompassing \( \text{PIK3CA} \). After CLIA validation of \( \text{PTEN} \) protein loss in this tumor, the patient received the dual phosphoinositide 3-kinase (PI3K) and mTOR inhibitor BEZ235 on a phase I study. Interestingly, this patient’s cancer also harbored a homozygous frameshift mutation in \( \text{NF1} \), an inhibitor of RAS and mTOR (30, 31).

Sequence analysis of mTNBC2 revealed a high-level \( \text{BRAF} \) amplification, which was validated as being part of complex extrachromosomal double minute by FISH (Fig. 4). This patient’s cancer also showed broad amplification and overexpression of \( \text{HRAS} \) and underexpression of \( \text{INPP4B} \), when compared with nonmalignant breast tissue control samples (Table 2). These findings suggested activation of both the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways in this cancer (Fig. 5). CLIA validation of \( \text{BRAF} \) overexpression and \( \text{INPP4B} \) underexpression by microarray was completed, and the patient was then treated on a phase I study with the oral mitogen-activated protein/extracellular signal–regulated kinase (MEK) inhibitor, trametinib (GSK1120212), in combination with the oral AKT inhibitor, GSK 2141795 (32). With 2 months of treatment, her breast mass nearly completely regressed. She developed toxicities including skin ulceration, diarrhea, anorexia, and significant fatigue. Shortly thereafter, the patient had a seizure and was found to have a hemorrhagic brain metastasis and she discontinued the investigational therapies.

Finally, in addition to cancers with alterations in the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways, NGS also identified mTNBC with homologous recombination defects. Patient mTNBC9 had a complete response with preoperative chemotherapy (doxorubicin, cyclophosphamide then paclitaxel), and at bilateral mastectomy. Approximately 2.5 years later, she developed metastatic disease and had a complete response documented on positron emission tomography computed tomography scan with 4 months of treatment with chemotherapy (gemcitabine, carboplatin, and iniparib) on an expanded access protocol. Of interest, sequencing analysis of patient mTNBC9 revealed somatic events in

---

**Figure 3.** Integrated view of DNA and RNA fold-changes within tumor for chromosome 5q31.2-31.3 encompassing \( \text{CTNNA1} \). For each individual, log\( _2 \) fold-change of RNA and DNA are shown as compared with a population-based reference sample for RNA or the germline sample for DNA. Within each, RNA is shown as red for significantly overexpressed genes (P < 0.001 and fold-change >2) and DNA is shown on a blue-yellow scale in which deletions are blue lines. Inaccessible regions are clear and regions not exhibiting a significant fold-change are gray. Two individuals, mTNBC1 and mTNBC6, both show an approximate 150 kb homozygous deletion encompassing part of \( \text{CTNNA1} \) leading to a significant downregulation of expression.
multiple genes encoding proteins involved in DNA repair, double-strand break repair, and homologous recombination repair (Fig. 6). This included homozygous deletion and underexpression of BRCA2, significant underexpression of BRCA1 by outlier analysis, an SNV in DCLRE1C (Artemis), and a structural event encompassing BRIP1 (Table 2; Fig. 6).

Discussion

This study represents a comprehensive sequence-based genomic survey of somatic events occurring in an ethnically diverse cohort of 14 mTNBC, and describes potential genotype–phenotype correlations through the identification of targetable mutations and clinical outcomes. Recently, Shah and colleagues provided insights into the genomic makeup of primary TNBC through genome, exome, and transcriptome sequencing of a cohort of approximately 100 tumors (20). Other recent studies have also applied whole genome or exome sequencing to a large series of primary breast cancers, including TNBCs (21, 33, 34). These studies have provided an understanding of tumor cell clonality and heterogeneity and have suggested new biologic subsets of TNBC.

Transcriptome sequencing and differential expression analysis against nonmalignant hyperplastic breast tissues revealed significant ontologic profiles associated with G2–M checkpoint and proliferation indicative of basal-like breast cancers. The FOXM1 gene was overexpressed in 12 of 14 tumors. FOXM1 is a transcriptional activator involved in proliferation, cell-cycle control, and mitosis, through the regulation of many genes involved in G2–M and mitotic checkpoint, such as AURKA, AURKB, PLK1, and CENPF among others (35). The FOXM1 gene was also recently highlighted as significantly deregulated in serous ovarian tumors and breast cancer (36–38).

The present study integrated comprehensive DNA and RNA sequencing data in mTNBCs and confirmed the presence of frequent known mutations including TP53 mutation (21). This study also revealed allele-specific expression of TP53 mutations. This observation is consistent with the absence of expression of wild-type TP53 within neighboring normal cells adjacent to TP53-mutated cancer cells as has been previously reported, for which the actual mechanism remains unknown (39). Recent studies of malignant brain tumors suggest that similar monoallelic expression of TP53 due to LOH have prognostic value toward outcome (40). Integrated analysis of DNA and RNA data also helped to reconcile transcriptional consequences of mutations encompassing consensus splice sites. This was evidenced by a 39 bp deletion in RB1 encompassing the consensus splice acceptor site that results in exon skipping and an in-frame, but likely non-functional form of RB1 that was expressed at normal levels.
levels. In this case, the use of a 3’ probe microarray would show normal RB1 expression, however, this form of RB1 is likely nonfunctional due to loss of a large number of amino acids within a highly conserved domain.

This study also highlights additional genes that may play a role in mTNBC pathogenesis including CDH5, HERC1, LRP1B, and TOP2A. CDH5 is a member of the cadherin family of cell adhesion molecules, and has been previously identified as a strong candidate tumor suppressor gene (41). HERC1 is an E3 ubiquitin ligase, and is known to interact with and destabilize the tumor suppressor TSC2, a negative regulator of mTOR (42). We also detected mutations in multiple tumors within the LRP1B gene, which is frequently mutated in non–small cell lung carcinoma and recently associated with acquired chemotherapeutic resistance in ovarian cancer (43, 44). TOP2A

### Table 2. Highlighted cancer gene alterations in mTNBC

<table>
<thead>
<tr>
<th>Patient</th>
<th>Amp</th>
<th>Del</th>
<th>Mutated Structural variant</th>
<th>Overexpressed genes</th>
<th>Underexpressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTNBC1</td>
<td>LRP6</td>
<td>CTNNA1</td>
<td>NF1</td>
<td>ERBB4</td>
<td>EGF, HSP90AA1, RASGRF1, TFDI, MDM4, CCNE2, AR</td>
</tr>
<tr>
<td></td>
<td>PK3CA</td>
<td>PTEN</td>
<td>RB1</td>
<td></td>
<td>CTNNA1, NF1</td>
</tr>
<tr>
<td></td>
<td>PTGS2</td>
<td>GSK3B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mTNBC2</td>
<td>BRAF</td>
<td>MYC</td>
<td>TP53</td>
<td>ERBB4</td>
<td>BRAF, NRG3, NFKB2, PAPR1, NFKBIA, HRAS</td>
</tr>
<tr>
<td></td>
<td>NOTCH2</td>
<td>TERT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mTNBC3</td>
<td>SMAD3</td>
<td>SMAD6</td>
<td></td>
<td></td>
<td>PDGFRB, VEGFC, PDGFR, GLI2, VEGFA</td>
</tr>
<tr>
<td>mTNBC4</td>
<td>WT1</td>
<td>KLF8</td>
<td>TP53</td>
<td></td>
<td>WT1, PAK7</td>
</tr>
<tr>
<td>mTNBC5</td>
<td>IQGAP3</td>
<td>ERCC2</td>
<td>LRP1B</td>
<td></td>
<td>IL2, ALK, AKT2, CBLC, CCNE1, MYCN</td>
</tr>
<tr>
<td>mTNBC6</td>
<td>WHSC1L1</td>
<td>FBXW7</td>
<td>TOP2A</td>
<td>ERBB4</td>
<td>ID2, RPRM, EREP</td>
</tr>
<tr>
<td></td>
<td>FGR1</td>
<td>CTNNA1</td>
<td></td>
<td></td>
<td>FBXW7, CTNNA1, DKK1</td>
</tr>
<tr>
<td>mTNBC7</td>
<td>NOTCH2</td>
<td></td>
<td>DNER</td>
<td></td>
<td>CDKN2A, SFN[14-3-3 sigma]</td>
</tr>
<tr>
<td>mTNBC8</td>
<td>ATG5</td>
<td></td>
<td>TP53</td>
<td>MDM2</td>
<td>CAMK4D, ATG5, PTGS2, MET, TGFB2, AR, PDGFR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TOP2A</td>
<td>BRD4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CDH5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mTNBC9</td>
<td>BRCA2</td>
<td></td>
<td>TP53</td>
<td>DCLRE1C</td>
<td>ERCC4, CCNE2, CDK1, MET, SMC1B, ALK, MYC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DDB1</td>
<td>PTEN</td>
<td>BRCA1, BRCA2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rif1</td>
<td>BRI1p1</td>
<td></td>
</tr>
<tr>
<td>mTNBC10</td>
<td>KRAS</td>
<td></td>
<td>TP53</td>
<td></td>
<td>KRAS, DNTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CDH5</td>
<td></td>
<td>ID4, SFRP1</td>
</tr>
<tr>
<td>mTNBC11</td>
<td>SMAD3</td>
<td>PTEN</td>
<td>BRCAl</td>
<td>RB1</td>
<td>NOTCH3, TERT, CCNE2, NFKB1, MET, RET</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PTCH2</td>
<td>RHEB</td>
<td>NFATC2, TRAF3, SFRP1, DTX1, PTEN, PIK3CG</td>
</tr>
<tr>
<td>mTNBC12</td>
<td>MYB</td>
<td>ARAF</td>
<td>TP53</td>
<td>CDH5</td>
<td>ELK1, ARAF, HMG2A, FGR2, GLI1, SHH, TGFB2, PDGFRB, PDGFR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDKN2A</td>
</tr>
<tr>
<td>mTNBC13</td>
<td>CADM2,</td>
<td>TNNI2</td>
<td>TP53</td>
<td>DLR1</td>
<td>SFRP2, PTCH2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERCC2</td>
<td>CDH5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAM2</td>
<td>HMMR</td>
<td></td>
</tr>
<tr>
<td>mTNBC14</td>
<td>FOXM1</td>
<td>PTPRM</td>
<td>TP53</td>
<td>STAT1</td>
<td>IGFBP3, PRKCG, NRG4, NRG1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RUNX1</td>
<td></td>
<td></td>
<td>CDKN2A</td>
</tr>
</tbody>
</table>

---

aFocal amplification from outlier analysis with log fold-change ≤ or ≥ 2.0 and P value P ≤ 0.05.
bFocal homozygous deletion from outlier analysis with log fold-change ≤ or ≥ 2.0 and P value P ≤ 0.05.
cGermline mutation detected from outlier analysis with log fold-change ≤ or ≥ 2.0 and P value P ≤ 0.05.
controls the topologic states of DNA (45), and alterations in TOP2A have been associated with drug resistance in breast cancer (46).

Copy number analysis using whole-genome sequencing data suggest that homozygous deletion of CTNNA1 may be enriched among women of recent West-African descent, who typically have more aggressive and treatment-resistant disease. CTNNA1 encodes catenin (cadherin-associated protein), α-1, a protein that forms a complex to anchor E-cadherin to the cell membrane to maintain normal cell adhesion properties (47). It is well known that aberrations deregulating this complex result in dissociation of cancer cells from tumor foci and represent a key primer for invasion and metastasis (48). Furthermore, a neuronal knockout model of α-catenin showed abnormal CNS cell growth with spreading of ventricular zone cells throughout the brain, which formed invasive tumor-like masses similar to those seen in human

### Table 3. Clinical history and treatment outcome of mTNBC

<table>
<thead>
<tr>
<th>Participant</th>
<th>Treatment while on sequencing study</th>
<th>Best-response disease site</th>
<th>Additional history and best response</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTNBC1</td>
<td>BEZ235 (PI3K/mTOR inhibitor)</td>
<td>SD</td>
<td>pCR: preoperative AC/T; PR: iniparib/gemcitabine/carboplatin; PD: bortezomib/cyclophosphamide/pegylated liposomal doxorubicin</td>
</tr>
<tr>
<td>mTNBC2</td>
<td>Trametinib (MEK inhibitor) + GSK2141795 (AKT inhibitor)</td>
<td>PR</td>
<td>PR: gemcitabine/carboplatin; docetaxel/cyclophosphamide; paclitaxel/bevacizumab</td>
</tr>
<tr>
<td>mTNBC3</td>
<td>BEZ235 (PI3K/mTOR inhibitor)</td>
<td>Breast</td>
<td>PD: doxorubicin/cyclophosphamide</td>
</tr>
<tr>
<td>mTNBC4</td>
<td>Eribulin</td>
<td>Lung</td>
<td>SD: iniparib/gemcitabine/carboplatin</td>
</tr>
<tr>
<td>mTNBC5</td>
<td>EZN 2208 (TOPO1 inhibitor)</td>
<td>PR</td>
<td>PR: preoperative AC/T; PR: iniparib/gemcitabine/carboplatin; PD: bortezomib/cyclophosphamide/pegylated liposomal doxorubicin</td>
</tr>
<tr>
<td>mTNBC6</td>
<td>Paclitaxel + bevacizumab + everolimus (mTOR inhibitor)</td>
<td>CR</td>
<td>No response to preoperative AC/T</td>
</tr>
<tr>
<td>mTNBC7</td>
<td>Trametinib (MEK inhibitor) + GSK2141795 (AKT inhibitor)</td>
<td>PD</td>
<td>Primary refractory to all cytotoxic agents</td>
</tr>
<tr>
<td>mTNBC8</td>
<td>Iniparib + gemcitabine + carboplatin</td>
<td>Response Bone</td>
<td>Bone-only metastatic disease at presentation</td>
</tr>
<tr>
<td>mTNBC9</td>
<td>Iniparib + gemcitabine + carboplatin</td>
<td>CR</td>
<td>PR: preoperative AC/T; PD: iniparib/gemcitabine/carboplatin</td>
</tr>
<tr>
<td>mTNBC10</td>
<td>EZN 2208 (TOPO1 inhibitor)</td>
<td>PR</td>
<td>PR: preoperative AC/T; PD: iniparib/gemcitabine/carboplatin</td>
</tr>
<tr>
<td>mTNBC11</td>
<td>Iniparib + gemcitabine + carboplatin</td>
<td>PD</td>
<td>Primary refractory to preoperative AC/T; PD: iniparib/gemcitabine/carboplatin</td>
</tr>
<tr>
<td>mTNBC12</td>
<td>Nab paclitaxel + capecitabine</td>
<td>PR</td>
<td>Bone and diffuse adenopathy at presentation</td>
</tr>
<tr>
<td>mTNBC13</td>
<td>Iniparib + gemcitabine + carboplatin</td>
<td>PR</td>
<td>BRCA2 mutation; ovarian cancer (paclitaxel/carboplatin) and bilateral breast cancer; PD: ixabepilone/capecitabine</td>
</tr>
<tr>
<td>mTNBC14</td>
<td>BEZ235 (PI3K/mTOR inhibitor)</td>
<td>SD</td>
<td>Regional lymphadenopathy primary refractory to all cytotoxic agents/radiation</td>
</tr>
</tbody>
</table>

Abbreviations: AC/T, doxorubicin, cyclophosphamide, paclitaxel; CR, complete response; pCR, pathologic complete response; PD, progressive disease; PR, partial response; SD, stable disease.
ERK and PI3K/AKT/mTOR signaling pathways in mTNBC. Up to 9 of 14 patients with mTNBC in this study showed potentially biologically significant somatic alterations in genes involved in PI3K/AKT/mTOR and/or RAS/RAF/MEK/ERK signaling pathways (Fig. 5). Three cancers contained single events including focal KRAS amplification, focal ARAF amplification, and focal FBXW7 homozygous deletion. However, 2 cancers (mTNBC1 and mTNBC2) harbored multiple events supporting concomitant activation of both signaling pathways. Observations supporting activation of the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways in TNBC have been previously suggested both in vitro and in vivo (53). A study based on a systems approach observed that basal breast cancer cell lines were susceptible to MEK inhibitors in vitro (54). Interestingly, both tumors showing dual PI3K/AKT/mTOR and RAS/RAF/MEK/ERK pathway alterations in our study have gene expression ontologies indicative of basal tumors with enrichment of genes involved in cell-cycle control, G2–M checkpoint regulation, and mitosis (Supplementary Table S4; refs. 9–12). Furthermore, MEK inhibition was shown to illicit a feedback loop through the PI3K/AKT pathway, leading to limited anti-tumor efficacy with MEK inhibition alone; however, synergistic antitumor effects occurred with dual MEK and PI3K inhibition (54). This hypothesis was further supported using an independent set of MEK and PI3K inhibitors both in vitro and in vivo (55). It was shown that PTEN loss reduces the effectiveness of single agent MEK inhibition, and that combined inhibition of MEK and PI3K was required to effectively reduce cell-cycle progression and basal breast cancer cell growth (55).

This study also highlights the potential of using whole genome and transcriptome analysis for detecting targetable events mTNBC. For example, the observation of high-level BRAF amplification and overexpression in mTNBC2 raises the question of how best to therapeutically target this alteration. In the context of wild-type BRAF, RAF inhibitors can enhance tumor growth in a RAS-dependent manner, whereas this is not observed with MEK inhibitors (55). As mTNBC2 had wild-type BRAF amplification, a PI3K/AKT inhibitor in combination with a MEK inhibitor might be more a rational therapeutic strategy than a combination including a RAF inhibitor. Furthermore, mTNBC2 also had significant underexpression of INPP4B, a negative regulator of PI3K (56).

Here, we applied whole genome and transcriptome sequencing to the problem of mTNBC. One limitation of our study is that of sensitivity of 30× genome coverage in identifying low-level mutations in tumors. Deep exome or panel-based sequencing approaches have higher sensitivity, thus are likely to have lower false-negative rates for coding mutation detection. These approaches can also identify low-level mutations associated with minor tumor cell clones (34). However, these approaches have the limitation of missing important structural events, as coding mutations and small indels are not the only mechanisms of biologically and clinically relevant somatic

CNS tumors, such as medulloblastoma, neuroblastoma, and retinoblastoma (49). Interestingly, the 2 tumors in this study that exhibit CTNN1A1 loss were both from African American patients. In a previous study describing somatic alterations by whole-genome sequencing of both a primary and metastatic lesion from a single 44-year-old African American patient with mTNBC, a homozygous deletion was also detected at the CTNN1A1/SIL1 locus (19). Therefore, taken together, these data suggest an important role for loss of CTNN1A1 in providing some TNBC tumors with invasive metastatic potential, and suggest enrichment of this aberration in African American TNBC.

Furthermore, we uncovered a number of somatic events involving the ERBB4 locus. The role of ERBB4 in mammary physiology, including maturation of mammary glands during pregnancy and lactation through Stat5 activation is well known (50). ERBB4 mutations have been reported in multiple tumor types including lung carcinoma and melanoma, in which mutations are believed to be oncogenic (43, 51). However, ERBB4 has also been implicated as a tumor suppressor with growth inhibitory functions, and reactivation of epigenetically silenced ERBB4 using 5-aza-2′-deoxycytidine resulted in increased apoptosis in BT20 breast cancer cells (52). These are among the first data supporting somatic events occurring at the ERBB4 locus in TNBC.

These analyses have further revealed genomic contexts that may have significant downstream therapeutic implications including dual activation of the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR signaling pathways in mTNBC.
change. The use of long insert whole-genome sequencing offers a greater power for detecting breakpoints and structural events that might encompass important cancer-related genes. For instance, some events such as inversions and moderately sized copy number changes in important cancer-related genes might be missed if using an exome or panel approach. Of note, we detected 2 copy neutral tandem overlapping inversion events within the INPP5F locus in mTNBC3. As both inversions span multiple exons of the INPP5F locus, it is likely that this structural alteration would lead to significant gene disruption in this tumor. INPP5F encodes inositol polyphosphate 5-phosphatase F, which modulates AKT/GSK3β signaling by decreasing AKT and GSK3β phosphorylation (57). A number of genes were involved in translocation events that might also be relevant in cancer (Supplementary Table S9). These types of events are likely to be missed by exome or panel-sequencing approaches. Thus, future approaches might take advantage of deep (>100×) exome sequencing for high sensitivity detection of low-level point mutations, in conjunction with low coverage (>8×) long insert genome sequencing to capture relevant structural changes in tumors. Through the analysis of mTNBC, these results add to our nascent understanding of these generally irremediable tumors. This study is integrative and comprehensive in nature through the generation and analysis of multiple dimensions of genomic data in conjunction with prospective clinical outcomes. These observations must be further explored in independent cohorts of molecularly and clinically characterized mTNBCs to evaluate the translational impact of these findings.

Disclosure of Potential Conflicts of Interest
D.W. Craig has honoraria from Speakers Bureau of Life Technology. O. Sakarya is employed as Bioinformatics Scientist at Life Technologies and as Computational Biologist at Genomic Health Inc. F.M. De La Vega is a Distinguished Fellow of Life Technologies. A. Siddiqui has ownership interest (including patents) in Life Technologies. A.W. Tolcher is consultant/advisory board member of Abgenomics, Abraxis, Bayer, Bind Biosciences, Celator, Celgene, Curis, Complete Genomics, Cytox, Daiichi Sankyo, Dendreon, Dicerma, Actavis, Eli Lilly, EMD Serono, Endo, Enzon, Evergent, Exelixis, Five Prime, Galapagos, Genentech, Geron, Adnexus, GSK, HUYA, Icon Clinical Research, Insert Therapeutics, Intellikine, Invirviv, Janssen, Johnson & Johnson, Merck, Micromet, Ambit, Nanoworks, Nektar, Neumedicines, Novartis, OncoMed, Onyx, Otsuka, Pfizer, PPD Development, Precision Health Holdings, Amgen, ProNai, Regeneron, Sanofi, Spectrum, Symphogen, Triphase Accelerator Corp., Vaccines, Veeva Oncology, Zynex, Ariad, Arreto, Astellas, and Astex. J.D. Carpten has commercial research grant from Life Technologies. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Genome Sequencing in Recurrent Triple-Negative Breast Cancer

Development of methodology: D.W. Craig, J.A. O'Shaughnessy, A. Christoforides, F.M. De La Vega, A. Siddiqui, L. Hoang, S. Mousses, J.D. Carpten

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.A. O'Shaughnessy, S. Wong, J. Dinh, J.L. Blum, C.L. Attelli, C.R. Osborne, J. Koeman, F.M. De La Vega, B. Salbia, A. Tolcher, J.M. Trent, J.D. Carpten


Writing, review, and/or revision of the manuscript: D.W. Craig, J.A. O'Shaughnessy, J.A. Kiefer, J.L. Blum, P.R. Billings, A.W. Tolcher, J.M. Trent, S. Mousses, D. Von Hoff, J.D. Carpten

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D.W. Craig, J. Aldrich, A. Baker, L. Hoang, D. Von Hoff, J.D. Carpten

Study supervision: D.W. Craig, J.A. O'Shaughnessy, P.R. Billings, A.W. Tolcher, J.D. Carpten

Samples processing and their preparation for analysis: T.M. Moses

Acknowledgments

The authors thank Dr. Shannon Morris of GlaxoSmithKline, Inc.

Grant Support

Funding for this work was provided by the Life Technologies Foundation (J.D. Carpten).

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 August 2, 2012; revised October 12, 2012; accepted October 30, 2012; published OnlineFirst November 19, 2012.

References


www.aacrjournals.org

Mol Cancer Ther; 12(1) January 2013 115

Published OnlineFirst November 19, 2012; DOI: 10.1158/1535-7163.MCT-12-0781

Downloaded from mct.aacrjournals.org on June 22, 2017. © 2013 American Association for Cancer Research.
Molecular Cancer Therapeutics

Genome and Transcriptome Sequencing in Prospective Metastatic Triple-Negative Breast Cancer Uncovers Therapeutic Vulnerabilities

David W. Craig, Joyce A. O'Shaughnessy, Jeffrey A. Kiefer, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1535-7163.MCT-12-0781

Supplementary Material
Access the most recent supplemental material at:
http://mct.aacrjournals.org/content/suppl/2012/11/19/1535-7163.MCT-12-0781.DC1

Cited articles
This article cites 56 articles, 21 of which you can access for free at:
http://mct.aacrjournals.org/content/12/1/104.full.html#ref-list-1

Citing articles
This article has been cited by 13 HighWire-hosted articles. Access the articles at:
/content/12/1/104.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.