Systematic Review and Meta-Analysis

Medicine®

OPEN

Integrative exploration of genomic profiles for triple negative breast cancer identifies potential drug targets

Xiaosheng Wang, MD, PhD[a,*], Chittibabu Guda, PhD[b,c,d,e]

Abstract

Background: Triple negative breast cancer (TNBC) is high-risk due to its rapid drug resistance and recurrence, metastasis, and lack of targeted therapy. So far, no molecularly targeted therapeutic agents have been clinically approved for TNBC. It is imperative that we discover new targets for TNBC therapy.

Objectives: A large volume of cancer genomics data are emerging and advancing breast cancer research. We may integrate different types of TNBC genomic data to discover molecular targets for TNBC therapy.

Data sources: We used publicly available TNBC tumor tissue genomic data in the Cancer Genome Atlas database in this study.

Methods: We integratively explored genomic profiles (gene expression, copy number, methylation, microRNA [miRNA], and gene mutation) in TNBC and identified hyperactivated genes that have higher expression, more copy numbers, lower methylation level, or are targets of miRNAs with lower expression in TNBC than in normal samples. We ranked the hyperactivated genes into different levels based on all the genomic evidence and performed functional analyses of the sets of genes identified. More importantly, we proposed potential molecular targets for TNBC therapy based on the hyperactivated genes.

Results: Some of the genes we identified such as *FGFR2*, *MAPK13*, *TP53*, SRC family, MUC family, and BCL2 family have been suggested to be potential targets for TNBC treatment. Others such as CSF1R, EPHB3, TRIB1, and LAD1 could be promising new targets for TNBC treatment. By utilizing this integrative analysis of genomic profiles for TNBC, we hypothesized that some of the targeted treatment strategies for TNBC currently in development are more likely to be promising, such as poly (ADP-ribose) polymerase inhibitors, while the others are more likely to be discouraging, such as angiogenesis inhibitors.

Limitations: The findings in this study need to be experimentally validated in the future.

Conclusion: This is a systematic study that combined 5 different types of genomic data to molecularly characterize TNBC and identify potential targets for TNBC therapy. The integrative analysis of genomic profiles for TNBC could assist in identifying potential new therapeutic targets and predicting the effectiveness of a targeted treatment strategy for TNBC therapy.

Abbreviations: AML = acute myeloid leukemia, BLBC = basal-like breast cancer, CNA = copy number alteration, EGFR = epidermal growth factor receptor, ER = estrogen receptor, FDR = false discovery rate, GO = gene ontology, GSEA = gene set enrichment analysis, HER2 = human epidermal growth factor receptor 2, MAF = mutation annotation format, miRNA = microRNA, mTOR = mammalian target of rapamycin, PARP1 = poly (ADP-ribose) polymerase 1, PR = progesterone receptor, TCGA = the Cancer Genome Atlas, TNBC = triple negative breast cancer, VEGF = targeting vascular endothelial growth factor.

Keywords: copy number variation, gene expression profiling, methylation, microRNA, somatic mutation, targeted therapy, triple negative breast cancer

1. Introduction

Approximately 10% to 20% of breast cancers are triple negative breast cancer (TNBC), a breast tumor subtype that is clinically negative for expression of the estrogen receptor (ER) and progesterone receptor (PR) and lacks overexpression of the human epidermal growth factor receptor 2 (HER2).[1] TNBC often has a poor prognosis due to its aggressive clinical behavior and lack of response to hormonal therapy or therapies that target HER2 receptors. So far, chemotherapy remains the only possible therapeutic strategy in the adjuvant or metastatic setting in TNBC.[2] For example, a latest neoadjuvant trial has shown that the addition of carboplatin to a standard neoadjuvant chemotherapy regimen significantly increased the pathologic complete response in TNBC patients.[3]

Some potential targeted-therapy-based approaches to TNBC treatment have been investigated such as targeting vascular endothelial growth factor (VEGF), epidermal growth factor receptor (EGFR), mammalian target of rapamycin (mTOR), and poly (ADP-ribose) polymerase 1 (PARP1).[4] One encouraging result from clinical trials has shown that the PARP inhibitor,

Editor: Konstantinos Dimas.

Authors' contributions: XW conceived of the research, performed the research, and wrote the manuscript. CG proposed suggestions for improvement and edited the manuscript.

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

[a] School of Basic Medicine and Clinic Pharmacy, China Pharmaceutical University, Nanjing, China, [b] Department of Genetics, Cell Biology and Anatomy, [c] Bioinformatics and Systems Biology Core, [d] Department of Biochemistry and Molecular Biology, [e] Fred and Pamela Buffet Cancer Center, University of Nebraska Medical Center, Omaha, Nebraska, USA.

[*] Correspondence: Xiaosheng Wang, China Pharmaceutical University, Nanjing, Jiangsu, China (e-mail: xiaosheng.wang@cpu.edu.cn).

Medicine (2016) 95:30(e4321)

Received: 10 February 2016 / Received in final form: 3 June 2016 / Accepted: 29 June 2016

http://dx.doi.org/10.1097/MD.0000000000004321

Veliparib, can improve pathologic complete response for TNBC patients by combined addition of carboplatin to standard presurgery chemotherapy.[5] However, clinical efficacies for most of targeted therapy remain unclear so far. Thus, discovery of new treatment targets and strategies for TNBC therapy is pressing and of significant interest.

A large volume of cancer genomics data are emerging and advancing breast cancer research.[6,7] The Cancer Genome Atlas (TCGA) Network gave comprehensive molecular portraits of human breast cancer by integrating various types of "omics" data including genomic DNA copy number arrays,[8] DNA methylation, exome sequencing, messenger RNA arrays, microRNA (miRNA) sequencing, and reverse-phase protein arrays. The related investigations have greatly advanced our understanding of breast cancer in molecular profiles, although translation of genomic findings into clinical applications remains challenging. The high-quality TCGA primary breast tumor samples and their comprehensive molecular profiles could be an invaluable source of information for molecular exploration of TNBC and discovery of new treatment targets.

In cancer research, gene expression measure is of great importance as it reflects gene activity directly and has successfully been used to stratify cancer into different subtypes.[9] Lehmann et al[10] identified 6 TNBC subtypes based on gene expression profiling and revealed that each subtype was related to unique gene ontologies and pathways. For example, the immunomodulatory subtype was enriched in immune cell processes and signal transduction pathways, while the luminal androgen receptor (LAR) subtype was enriched in androgen receptor signaling pathways. Further, they found that the different subtypes were uniquely sensitive to different agents. For example, the LAR cell lines were uniquely sensitive to bicalutamide (an androgen receptor [AR] antagonist), and the basal-like cell lines preferentially responded to cisplatin.

One major limitation of gene expression analysis is its variability and unsteadiness, as a single measure often leads to misinterpretation. To overcome the limitation, it is crucial to collect other genomic evidence such as DNA copy number alteration (CNA), DNA methylation, miRNA gene expression, and gene mutation data that also reflect gene activity and could cause gene expression change. Although previous studies have associated cancer with genomic changes in copy number, methylation, miRNAs, and gene mutations,[8,11] integration of the different types of genomic data into cancer research remains challenging, but promising. Some previous studies have used integrative approaches to analysis of TNBC. The following study of Lehmann et al[10] revealed that PIK3CA kinase domain mutations were frequent in the LAR subtype, and the combination of AR antagonism and PI3K inhibition could synergistically inhibit LAR TNBC cell growth.[12] This study exemplifies the importance of integrating different types of genomic data into exploration of discovery of cancer treatment targets. Shah et al[13] revealed that TNBCs exhibit a wide and continuous spectrum of genomic evolution by analyzing somatic mutation, CNA, gene fusions, and gene expression patterns of 104 primary TNBCs. Craig et al[14] integrated gene expression and somatic mutation profiling of 14 metastatic TNBCs using next-generation sequencing data and proposed potential therapeutic targets in metastatic TNBC.

Although these integrative analyses have provided important insights into TNBC,[12–14] a broader exploration of genomic profiles for TNBC could improve our understanding of this disease and detect potential targets for TNBC treatment. In this study, we carried out an integrative exploration of wide genomic profiles (gene expression, copy number, DNA methylation, miRNA gene expression, and gene mutation) for TNBC using the TCGA breast cancer data. In addition to dissecting the biology of TNBC, we attempt to find genes or pathways that could be targets for treatment of TNBC by identification of abnormally hyperactivated genes and pathways in TNBC. Here, we defined the abnormally hyperactivated genes as those genes that have higher expression, more copy numbers, lower methylation level, or are targets of miRNAs with lower expression in TNBC than in normal samples. Based on the different genomic evidences, we categorized the abnormally hyperactivated genes into different levels. The greater the indication that a gene is hyperactivated, the higher the level the gene belongs to. The genes in high levels are more likely to be associated with the pathogenesis of TNBC and therefore could be preferential targets for TNBC treatment.

2. Methods

2.1. Datasets

We downloaded the breast carcinoma gene expression (microarray), copy number, methylation, miRNA (Level 3), and gene somatic mutation data (Level 2) from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm). In the gene expression data, we found a total of 55 TNBC samples. Considering that the gene expression activity is our primary concern, and for statistical consistency, we analyzed the same 55 TNBC samples in the other 4 data types. There are 2, 0, 2, and 1 samples missing in copy number, methylation, miRNA, and gene somatic mutation data, respectively. Thus, we analyzed 55 TNBC samples in the gene expression and methylation data, 53 TNBC samples in the copy number and miRNA data, and 54 TNBC samples in the gene mutation data. Ethical approval was waived since we used only publicly available data and materials in this study.

2.2. Identification of genes with differential expression, copy number, or methylation level between TNBC and normal samples

Based on the microarray gene expression data, we identified significantly upregulated genes in the TNBC samples, relative to the paired normal samples with at least two-fold mean expression difference (Wilcox signed-rank test, FDR [false discovery rate] ≤ 0.05).

For the CNA, we used the "*.nocnv_hg19.seg.txt" data (SNP array 6.0). We annotated the overlapping genes with the genomic regions in the data using the tool PennCNV[15] and obtained the gene copy number by averaging the segment values of the same gene. We identified the genes having significant copy number gain in the TNBC samples relative to the paired normal samples with at least 1.2-fold mean copy number difference (Wilcox signed-rank test, FDR ≤ 0.05). Because the copy number difference was generally low with the maximum being 1.5, we used the 1.2-fold threshold to define the genes with significant copy number difference between the TNBC and normal samples.

In the methylation analysis, we used the data produced by 2 different platforms: HumanMethylation27 (HM27) BeadChip and HumanMethylation450 (HM450) BeadChip. The HM27 data include 32 TNBC samples versus 27 normal samples, and the HM450 data includes 23 TNBC samples versus 47 normal samples. We identified the hypomethylated genes in the TNBC

samples relative to normal samples with mean methylation level (β value) difference no less than 5% (Wilcox sum-rank test, FDR $\leq$ 0.05) in both datasets and selected the genes overlapping between both analyses as the hypomethylated genes in TNBC.

The FDR was estimated using the method of Benjami and Hochberg.[16]

2.3. Identification of genes that are targets of miRNAs with differential expression between TNBC and normal samples

We identified significantly downregulated *miRNA* genes in the 53 TNBC samples relative to 103 normal samples with at least two-fold mean expression difference (*t* test, FDR $\leq$ 0.05). Using the tool TargetScanHuman for predicting miRNA targets,[17] we identified the genes that are targets of the downregulated *miRNA* genes. As previously, the FDR was estimated using the method of Benjami and Hochberg.[16]

2.4. Identification of genes frequently mutated in TNBC

In the gene somatic mutation analysis, we used the MAF (mutation annotation format) data by exome-sequencing data analysis. We first constructed an $m \times n$ mutation matrix, where m is the number of genes and n is the number of breast cancer samples in the MAF data. The entry (i, j) in the matrix is 1 if at least 1 mutation in gene i was detected in sample j, otherwise 0. Based on the mutation matrix, we identified some frequently mutated genes in TNBC (Fisher exact test, P value <0.05) and compared their mutation rates in TBNC with those in general breast cancer (992 samples). For convenience, in some cases hereafter, we also call the frequently mutated genes abnormally hyperactivated in TNBC, although a gene mutation does not necessarily result to the hyperactivation of the gene.

2.5. Evaluation of significance of hyperactivated genes in TNBC

We categorized the identified genes into different levels based on all the genomic evidence. Level 1 includes those genes with significantly higher expression in TNBC samples than in normal samples; Level 2 includes those genes that belong to Level 1 and were identified as abnormally hyperactivated in at least one of the other genomic analyses (copy number, methylation, miRNA, and gene mutation); Level 3 includes those genes that belong to Level 1 and were identified as abnormally hyperactivated in at least two of the other genomic analyses; Level 4 includes those genes that belong to Level 1 and were identified as abnormally hyperactivated in at least three of the other genomic analyses; Level 5 includes those genes that belong to Level 1 and were identified as abnormally hyperactivated in all the other genomic analyses. The higher the level a gene belongs to, the more likely the gene is to be hyperactivated in TNBC.

2.6. Functional analysis of the gene sets identified

Using the gene set enrichment analysis (GSEA) software, we classified the hyperactivated genes into different gene families and identified the gene sets that are significantly overlapping with them. We inferred significant networks associated with gene sets using Ingenuity Pathway Analysis tool (IPA, Ingenuity® Systems, www.ingenuity.com). IPA is a system that yields a set of networks relevant to a list of genes based on the preserved records contained in the Ingenuity Pathways Knowledge Base. We identified significant gene ontology (GO) biological processes that are associated with gene sets using the PANTHER classification system.[18]

3. Results and discussion

3.1. Identification of the abnormally hyperactivated genes in TNBC

We identified 1800 upregulated genes in the TNBC samples with at least 2-fold higher mean expression compared to the normal samples (Wilcox signed-rank test, FDR $\leq$ 0.05). We identified 1655 genes that have at least 1.2-fold mean copy number gain in the TNBC samples compared to the normal samples (Wilcox signed-rank test, FDR $\leq$ 0.05). We identified 731 genes that have lower methylation level (β value) in TNBC samples compared to normal samples in both the HM27 and HM450 data analysis with mean β value difference no less than 5% (Wilcox sum-rank test, FDR $\leq$ 0.05). We identified 2020 *mRNA* genes that are targets of the 52 downregulated miRNAs in the TNBC samples compared to normal samples with at least 2-fold mean expression difference (*t* test, FDR $\leq$ 0.05). We also identified 18 genes that are frequently mutated in the TNBC samples (Fisher exact test, P value <0.05). Here, we refer to the groups of genes identified by gene expression, copy number, methylation, miRNA, and gene mutation analyses as GE, CN, ME, MR, and GM, respectively. These gene lists are shown in the Additional File 1, Table S1, http://links.lww.com/MD/B149.

Figure 1 illustrates overlaps between the gene sets identified in the different genomic analyses (also see the Additional File 2, Table S2, http://links.lww.com/MD/B150). For example, there are 209, 154, 167, and 2 genes overlapping between GE and CN, ME, MR, and GM, respectively; there are 32 genes overlapping among GE, CN, and ME. We categorized the identified genes into different levels based on all the genomic evidence. Level 1 includes the 1800 genes that were highly expressed in TNBC samples compared to normal samples; Level 2 includes 474 genes that lie in Level 1 and were hyperactivated in TNBC by at least one of the other genomic analyses; Level 3 includes 59 genes that lie in Level 1 and were hyperactivated in TNBC by at least two of the other genomic analyses. Both Levels 4 and 5 contain 0 genes. We only explored the genes in Levels 1, 2, and 3 (see the Additional File 1, Table S1, http://links.lww.com/MD/B149). Figure 2 is a heatmap that presents the Level 3 genes and their hyperactivated status in the different genomic analyses.

3.2. Functional analysis of the abnormally hyperactivated genes in TNBC

We are more interested in those genes in Levels 2 and 3 because they were not only highly expressed in TNBC but also identified abnormally hyperactivated by other genomic evidence. We classified the Level 2 genes into different gene families using the GSEA software as shown in Table 1.[19]

We used the "Compute Overlaps" tool in GSEA to get the gene sets (positional, curated, or oncogenic) that were significantly overlapping with the Level 2 gene list (FDR $< 10^{-10}$). Among them, a number of gene sets (Table 2) are related to breast cancer, other cancer types, and stem cells. Table 2 shows that the hyperactivated genes we identified in TNBC tend to be upregulated in the basal subtype of breast cancer, breast cancer resistant to chemotherapy, ER breast cancer, and aggressive

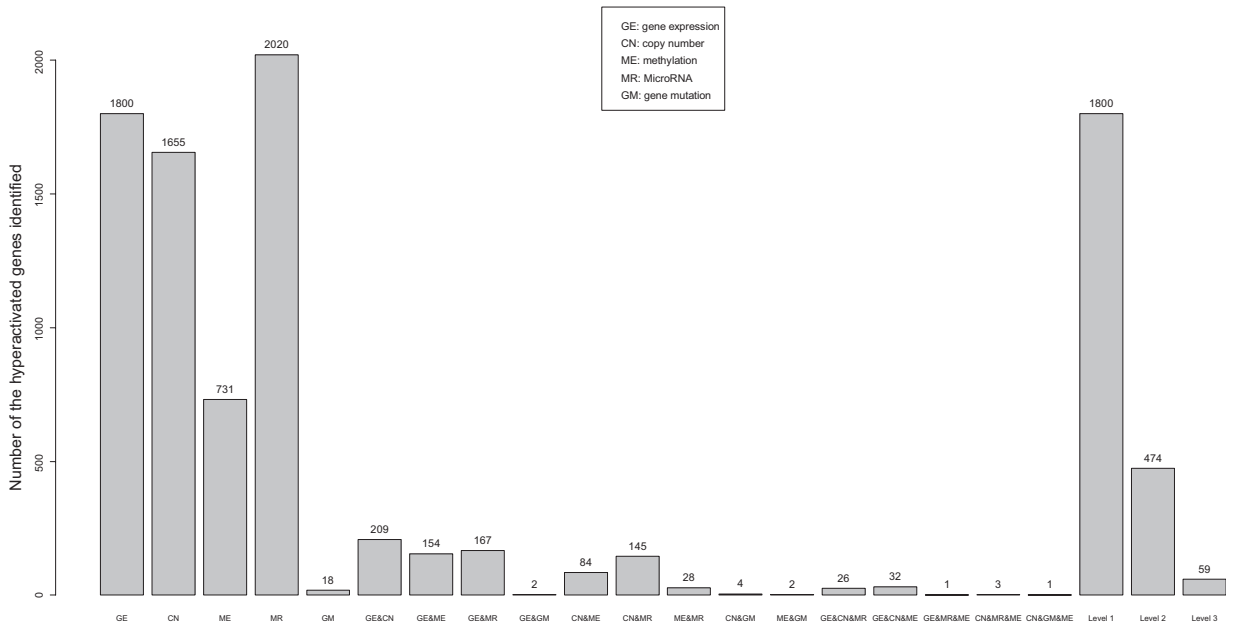

Figure 1. Numbers of the genes identified in all the genomic analyses.

prostate cancer, lymphoma, acute myeloid leukemia (AML), and hepatocellular carcinoma, underlying the molecular commonalities between TNBC and aggressive cancer types or subtypes. Many of the hyperactivated genes are also highly expressed in stem cells, indicating that the TNBC cells may harbor a substantial number of cancer stem cells that result in invasive activities of TNBC. In addition, Table 2 shows that many of the hyperactivated genes in TNBC are associated with dysregulation of TP53, aberrant activation of the Wnt signaling pathway, and immune system processes. These features have been correlated with aggressiveness and poor prognosis of breast cancers.[20–22]

We performed a network analysis of the Level 2 gene set with the addition of the tumor suppressor gene *TP53*, since dysregulation of TP53 has been found in the vast majority of TNBC cases.[23] In our analyses, TP53 mutation was highly frequent (78% mutation rate) in TNBC, and its expression was

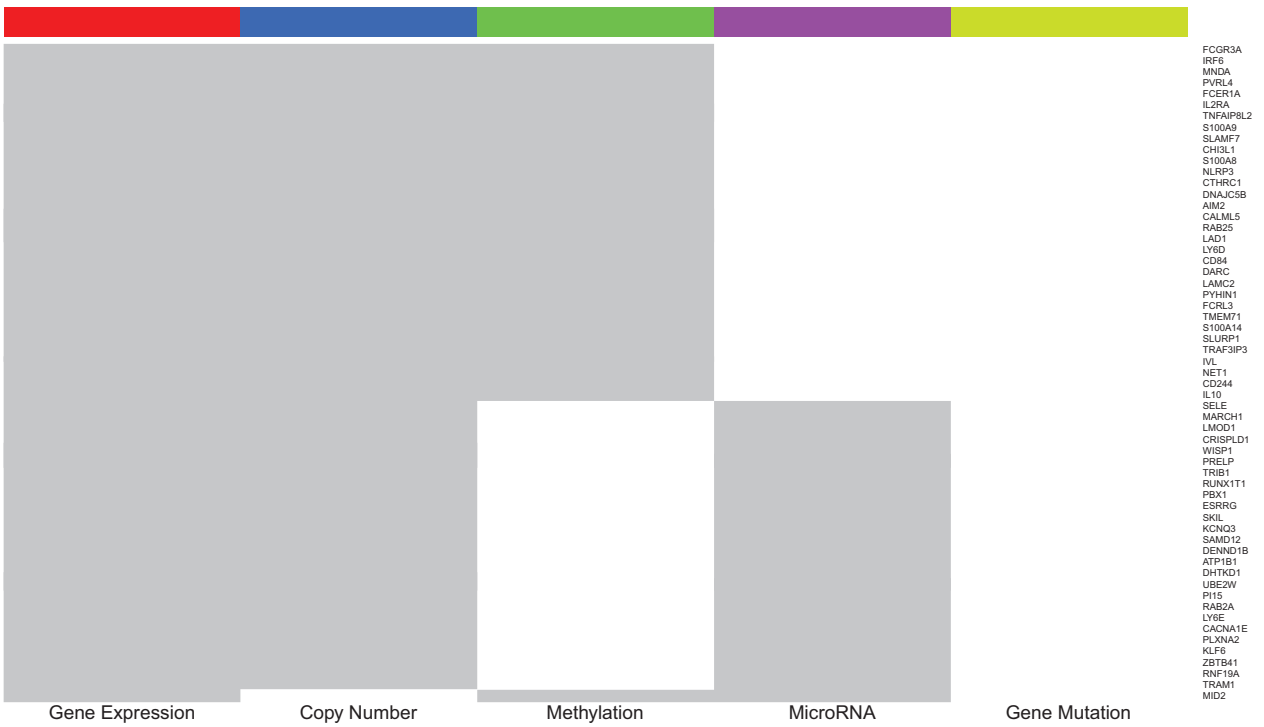

Figure 2. The Level 3 genes and their hyperactivated status in the different genomic analyses. The gray color indicates that the gene is hyperactivated in the analysis, and the white indicates that the gene isn't hyperactivated in the analysis.

Table 1

Category of the Level 2 genes.

Gene family*	Gene†
Tumor suppressors	*CYLD*, *KLF6*, *PIK3R1*, *PRF1*
Oncogenes	*BCL11A*, *BCL6*, *BTG1*, *CCDC6*, *CD79A*, *CIITA*, *FCGR2B*, *FGFR2*, *FLT3*, *KDR*, *LPP*, *MAFB*, *MYH9*, *NFIB*, *PBX1*, *RUNX1T1*, *ZNF521*
Translocated cancer genes	*BCL11A*, *BCL6*, *BTG1*, *CCDC6*, *CIITA*, *FCGR2B*, *LPP*, *MAFB*, *MYH9*, *NFIB*, *PBX1*, *RUNX1T1*, *ZNF521*
Protein kinases	*BMPR2*, *CASK*, *FGFR2*, *FLT3*, *INSR*, *KDR*, *MAPK13*, *MYLK*, *RIPK2*, *SCYL3*, *STK3*, *STK38L*, *TRIB1*
Cell differentiation markers	*BTLA*, *C5AR1*, *CCR7*, *CD163*, *CD244*, *CD48*, *CD79A*, *CD84*, *CD86*, *CD93*, *DARC*, *F11R*, *FASLG*, *FCGR2A*, *FCGR2B*, *FCGR3A*, *FCRL5*, *FGFR2*, *FLT3*, *FUT3*, *IL10RA*, *IL1R2*, *IL2RA*, *IL2RB*, *INSR*, *ITGA6*, *KDR*, *LAIR1*, *LILRA3*, *LILRA4*, *LILRB2*, *LILRB4*, *LILRB5*, *PROM1*, *PTPRC*, *SELE*, *SELL*, *SELP*, *SIGLEC9*, *SIRPG*, *SLAMF1*, *SLAMF7*, *TNFRSF9*, *TNFSF10*, *TNFSF13B*
Homeodomain proteins	*EN1*, *PBX1*, *TSHZ2*, *TSHZ3*, *ZHX2*
Transcription factors	*BCL6*, *BCL11A*, *CASK*, *CEBPE*, *CIITA*, *DACH1*, *EHF*, *ELF3*, *ELF5*, *EN1*, *ESRRG*, *FOXI1*, *GRHL1*, *HIVEP2*, *ID4*, *IRF6*, *KIAA0040*, *KLF5*, *KLF6*, *LITAF*, *LMO4*, *MAFB*, *MNDA*, *MTF1*, *NCALD*, *NFIA*, *NFIB*, *NFKBIA*, *OPTN*, *PBX1*, *RARB*, *RUNX1T1*, *SOX11*, *SOX4*, *SOX9*, *TBX19*, *TFCP2L1*, *TFEC*, *TRIM22*, *TRPS1*, *TSHZ2*, *TSHZ3*, *VGLL1*, *ZHX2*, *ZMYND11*, *ZNF238*, *ZNF532*
Cytokines and growth factors	*CALCB*, *CAMP*, *CCL11*, *CCL18*, *CCL5*, *CCL7*, *CCL8*, *CX3CL1*, *CXCL10*, *FASLG*, *FGF1*, *FGF7*, *IGF1*, *IL10*, *IL1B*, *KL*, *MIA*, *OSM*, *SEMA4A*, *SLURP1*, *TG*, *TGFB3*, *TNF*, *TNFSF10*, *TNFSF13B*, *XCL1*, *XCL2*

* The definition of gene families refers to the website: http://www.broadinstitute.org/gsea/msigdb/gene_families.jsp.

Tumor suppressors: both alleles of these genes need to be mutated for oncogenesis.

Oncogenes: a single mutated allele is sufficient to contribute to oncogenesis.

Translocated cancer genes: genes mutated by translocation.

Protein kinase: the protein kinase complement of the human genome.

Cell differentiation markers: human leukocyte and stromal cell molecules: the CD markers.

Homeodomain proteins: human homeodomain proteins.

Transcription factors: a compilation of human transcription factors.

Cytokines and growth factors: human cytokine and growth factor genes.

† Some Level 2 genes are not present in any gene family above; the same gene may belong to different gene families; the genes also belonging to Level 3 are underlined.

Table 2

Gene sets that are significantly overlapping with the Level 2 gene list.

Description of gene sets	No. of overlapping genes	FDR
Genes downregulated in the luminal B subtype of breast cancer	76	4.96E – 57
Genes upregulated in basal subtype of breast cancer samples	62	4.96E – 37
Genes upregulated in atypical ductal hyperplastic tissues from patients with breast cancer vs those without the cancer	42	4.67E – 32
Genes upregulated in invasive ductal carcinoma relative to ductal carcinoma in situ	41	1.2E – 27
Genes upregulated in lobular carcinoma vs normal ductal breast cells	15	3.71E – 14
Genes upregulated in the normal-like subtype of breast cancer	45	1.29E – 26
Genes upregulated in breast cancer tumors (formed by MCF-7 xenografts) resistant to tamoxifen	40	8.37E – 19
Genes downregulated in MCF7 cells (breast cancer) at 24 h of estradiol treatment	39	5.89E – 20
Genes upregulated in luminal-like breast cancer cell lines compared to the mesenchymal-like ones	39	1.23E – 21
Genes upregulated in basal-like breast cancer cell lines as compared to the mesenchymal-like ones	21	1.25E – 17
Downregulated genes from the optimal set of 550 markers discriminating breast cancer samples by ESR1 expression: ER+ vs ER– tumors	21	6.66E – 12
Genes downregulated in early primary breast tumors expressing ESR1 vs the ESR1 negative ones	10	2.25E – 12
Genes within amplicon 8q12-q22 identified in a copy number alterations study of 191 breast tumor samples	22	4.68E – 18
Genes within amplicon 8q23-q24 identified in a copy number alterations study of 191 breast tumor samples	21	2.42E – 15
Genes upregulated in prostate cancer samples from African-American patients compared with those from the European-American patients	37	8.08E – 26
Genes upregulated in PC3 cells (prostate cancer) after knockdown of EZH2 by RNAi	47	3.48E – 15
Upregulated genes that best discriminate plasmablastic plasmacytoma from plasmacytic plasmacytoma tumors	32	3.59E – 22
Genes upregulated in papillary thyroid carcinoma compared to normal tissue	37	5.17E – 20
Upregulated genes in angioimmunoblastic lymphoma compared to normal T lymphocytes	23	3.48E – 15
Genes upregulated in AML patients with mutated NPM1	25	1.96E – 18
Genes from “subtype S1" signature of hepatocellular carcinoma: aberrant activation of the Wnt signaling pathway	21	5.27E – 12
The “adult tissue stem" module: genes coordinately upregulated in a compendium of adult tissue stem cells	48	1.96E – 24
Genes upregulated in cultured stromal stem cells from adipose tissue, compared with the freshly isolated cells	35	9.4E – 19
Set “Suz12 targets": genes identified as targets of the Polycomb protein SUZ12 in human embryonic stem cells	43	1.26E – 12
Genes downregulated in ES (embryonic stem cells) with deficient SUZ12	24	2.78E – 12
Genes consistently upregulated in mammary stem cells both in mouse and human species	30	7.2E – 13
Genes upregulated in the HMEC cells (primary mammary epithelium) upon expression of TP53 off adenoviral vector	44	1.44E – 11
Genes involved in immune system	43	3.71E – 14

AML = acute myeloid leukemia, FDR = false discovery rate.

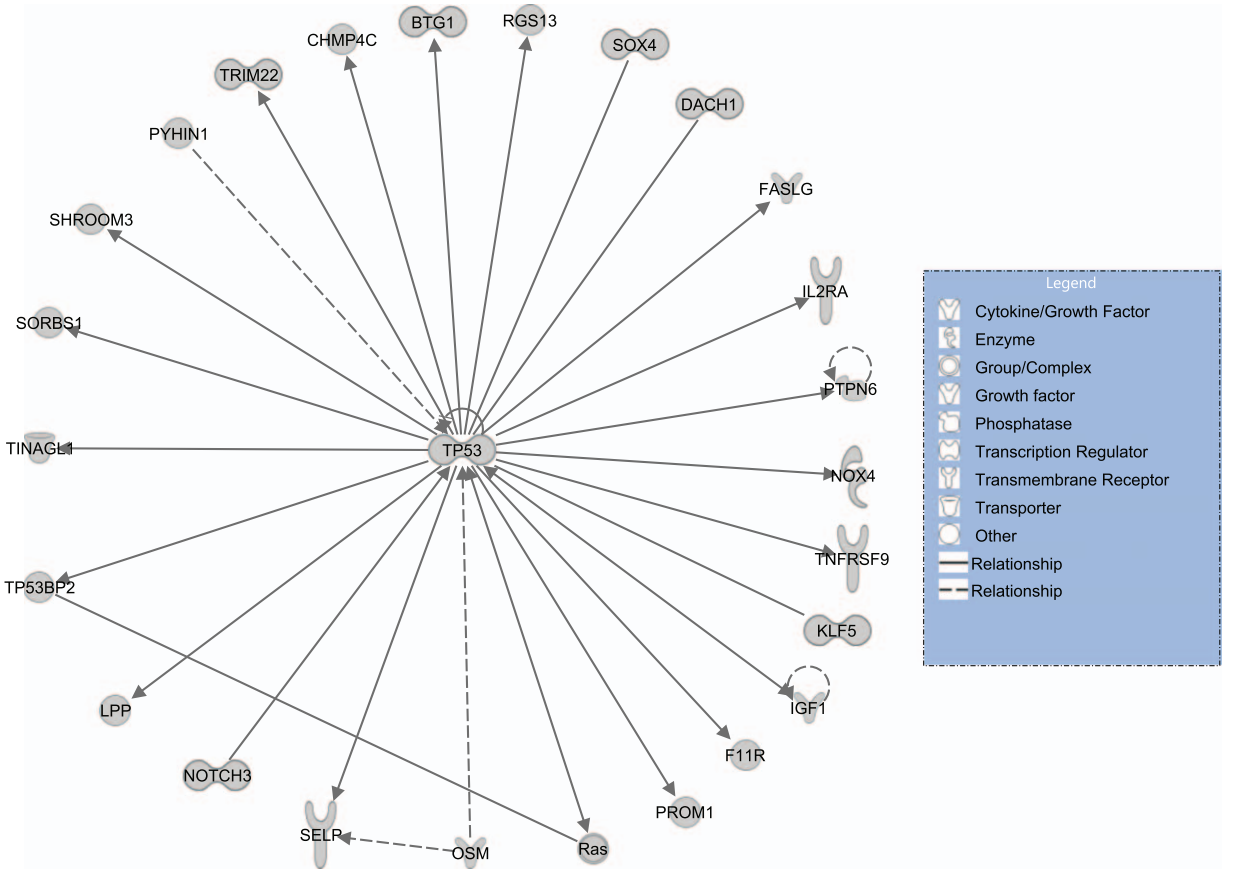

Figure 3. TP53-centered protein–protein interaction network identified based on the Level 2 gene set using Ingenuity Pathway Analysis.

significantly lower in TNBC than in normal samples (1.6-fold mean expression difference, Wilcox signed-rank test, FDR = 0.002). We generated a TP53-centered network (Fig. 3), in which TP53 connects to all the other nodes. Figure 3 shows that TP53 regulates many hyperactivated genes such as *RGS13*, *SOX4*, *NOTCH3*, *TRIM22*, and *IGF1*, and genes associated with RAS signaling. Dysregulation of TP53 may be associated with abnormal hyperactivation of these regulated genes and pathways that significantly contribute to pathogenesis and progression of TNBC.

3.3. Identification of genes that are frequently mutated in TNBC

In the 54 TNBC samples with exome-sequencing data, we found 18 genes that were frequently mutated (Fisher exact test, P value <0.05) as shown in Table 3. Notably, TP53 has the highest mutation rate (78%) that is much higher than its 31% mutation rate across all the TCGA breast cancers (odds ratio: 7.7, Fisher exact test P value $= 10^{-11}$), suggesting that TP53 mutations might significantly contribute to aggressiveness of TNBC. TTN has the second highest mutation rate (22%) in TNBC, slightly higher than its 19% mutation rate across all the breast cancers. Table 3 and Fig. 4 show that a majority of the frequently mutated genes in TNBC have significantly higher mutation frequency compared to breast cancer in general, suggesting that mutations in these genes may contribute to higher aggressiveness of TNBC compared to non-TNBC breast cancers.

Using the PANTHER classification system,[18] we identified significant GO biological processes associated with the 18 frequently mutated genes as shown in Table 4. Table 4 shows that these genes are mostly involved in important biological processes that underlie the pathogenesis of cancer.

In Table 3, 2 members of the MUC gene family, *MUC4* and *MUC16*, show high frequency of mutation in TNBC. It has been shown that MUC4 could promote invasive activities of TNBC and be associated with metastasis of breast cancer[24,25] and MUC16 could increase proliferation and antiapoptosis in breast cancer cells,[26] consistent with their high mutation rate in the aggressive TNBC. Interestingly, both MUC4 and MUC16 had decreased expression in TNBC compared to normal samples (Wilcox signed-rank test, P value $= 2 \times 10^{-5}$ and 0.035 for MUC4 and MUC16, respectively), but highly expressed in TNBC compared to non-TNBC tumor samples (t test, P value $= 2.2 \times 10^{-6}$ and $<10^{-7}$ for MUC4 and MUC16, respectively). This is similar to a previous finding that MUC4 expression was depressed in primary breast tumors relative to normal tissue, but was elevated in metastatic lesions compared to primary breast tumors,[24] suggesting that MUC4 may play an important role in promoting TNBC metastasis. Except for MUC4 and MUC6, other *MUC* genes also have mutations in TNBC (Table 5). In fact, *MUC* genes have been identified as attractive therapeutic targets since their deregulation has been associated with unfavorable prognosis of cancers.[27]

3.4. Identification of potential targets for TNBC therapy

3.4.1. The hyperactivated kinase-encoding genes could be promising targets for TNBC therapy. It has been recognized that many kinase-encoding genes are upregulated in cancer and the development of anticancer drugs that inhibit overexpression

Table 3

Genes frequently mutated in TNBC.

Symbol	Name	Mutation rate in TNBC (%)	Mutation rate in breast cancer (%)	Difference in mutation rate (odds ratio)*
TP53	Tumor protein p53	78	31	7.7 (10^{-11})
TTN	Titin	22	19	1.2 (0.6)
FAT3	FAT atypical cadherin 3	15	4	4.2 (0.002)
MUC4	Mucin 4, cell surface associated	15	7	2.2 (0.06)
USH2A	Usher syndrome 2A (autosomal recessive, mild)	15	5	3.1 (0.01)
F5	Coagulation factor V (proaccelerin, labile factor)	11	2	6.1 (0.001)
HYDIN	HYDIN, axonemal central pair apparatus protein	11	4	3.3 (0.02)
MUC16	Mucin 16, cell surface associated	11	10	1.1 (0.82)
OBSCN	Obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	11	5	2.5 (0.05)
APOB	Apolipoprotein B	9	4	2.6 (0.07)
CACNA1B	Calcium channel, voltage-dependent, N type, alpha 1B subunit	9	2	5.5 (0.005)
CSMD2	CUB and Sushi multiple domains 2	9	3	3 (0.04)
FLG	Filaggrin	9	6	1.7 (0.24)
FRAS1	Fraser extracellular matrix complex subunit 1	9	2	4.3 (0.01)
LAMA3	Laminin, alpha 3	9	2	5.5 (0.005)
MXRA5	Matrix-remodeling associated 5	9	3	3 (0.04)
RYR1	Ryanodine receptor 1 (skeletal)	9	4	2.4 (0.08)
ANKRD30A	Ankyrin repeat domain 30A	9	1	9.1 (0.0009)

TNBC = triple negative breast cancer.

* Comparison of mutation rate in TNBC versus all breast cancers (the Fisher exact test *P* values presented in parenthesis).

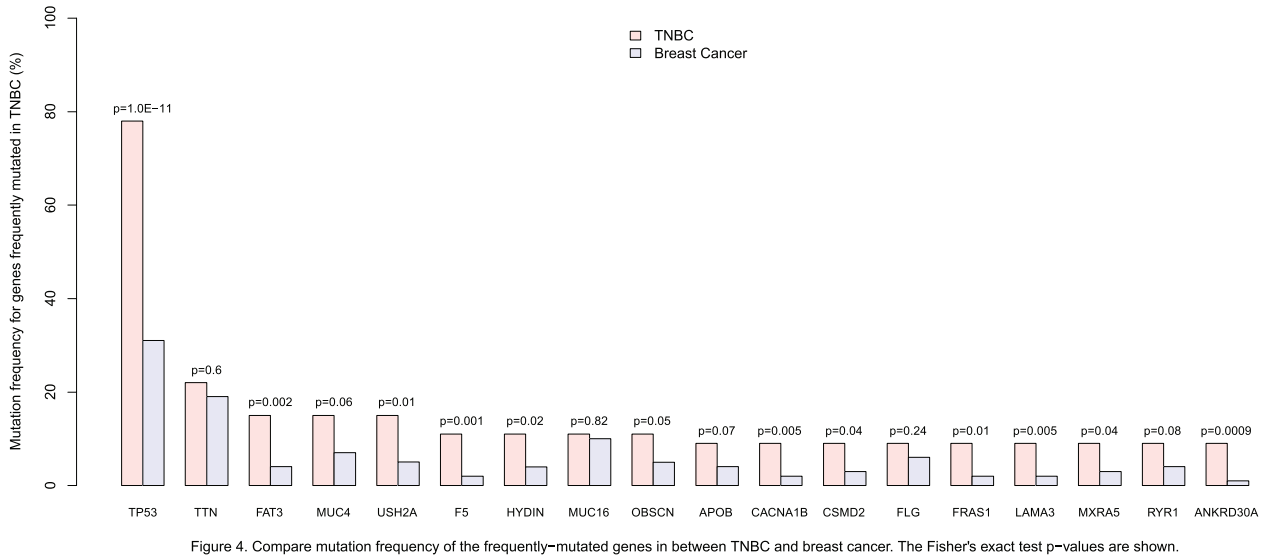

Figure 4. Compare mutation frequency of the frequently−mutated genes in between TNBC and breast cancer. The Fisher's exact test p−values are shown.

Figure 4. Compare mutation frequency of the frequently-mutated genes in between TNBC and breast cancer in general. The Fisher exact test *P* values are presented.

Table 4

Gene ontology related to the highly mutated genes in TNBC.

GO term	Associated genes
Apoptotic process (GO:0006915)	*MXRA5*, *TP53*
Biological adhesion (GO:0022610)	*OBSCN*, *MXRA5*, *FAT3*, *USH2A*, *LAMA3*, *F5*, *CSMD2*, *TTN*
Biological regulation (GO:0065007)	*ANKRD30A*, *F5*, *TP53*
Cellular component organization or biogenesis (GO:0071840)	*FAT3*, *TP53*
Cellular process (GO:0009987)	*OBSCN*, *MXRA5*, *FAT3*, *USH2A*, *LAMA3*, *F5*, *CSMD2*, *TTN*, *FRAS1*, *CACNA1B*, *RYR1*, *TP53*
Developmental process (GO:0032502)	*OBSCN*, *MXRA5*, *FAT3*, *USH2A*, *LAMA3*, *F5*, *TP53*
Immune system process (GO:0002376)	*MXRA5*, *RYR1*, *F5*, *CSMD2*
Localization (GO:0051179)	*CACNA1B*, *RYR1*, *F5*, *CSMD2*
Metabolic process (GO:0008152)	*ANKRD30A*, *OBSCN*, *MXRA5*, *F5*, *CSMD2*, *TP53*
Multicellular organismal process (GO:0032501)	*OBSCN*, *MXRA5*, *USH2A*, *LAMA3*, *F5*, *CACNA1B*, *RYR1*
Response to stimulus (GO:0050896)	*F5*, *TP53*, *RYR1*, *CSMD2*

Table 5

***MUC* genes mutated in TNBC.**

Gene*	No. of mutated TNBC samples	Mutation rate (%)	Expression change (TNBC vs normal)	Expression change (TNBC vs non-TNBC)	Identified in other analyses†
MUC4	8	15	Down	Up	No
MUC16	6	11	Down	Up	No
MUC5B	4	7	Up	Up	No
MUC12	3	6	Not significant	Not significant	No
MUC6	3	6	Up	Not significant	No
MUC2	2	4	Not significant	Not significant	No
MUC1	1	2	Up	Down	Yes (CN)
MUC13	1	2	Down	Not significant	Yes (ME)
MUC17	1	2	Up	Not significant	Yes (ME)
MUC20	1	2	Up	Not significant	No
MUC21	1	2	Not significant	Not significant	No
MUC7	1	2	Up	Not significant	No

TNBC = triple negative breast cancer.
* MUC15 is not mutated in any TNBC sample, but more highly expressed in TNBC compared to normal samples and non-TNBC tumors; MUCL1 is downregulated in TNBC compared to non-TNBC tumors.
† CN = copy number, ME = methylation.

of protein kinases is promising in cancer treatment.[28,29] Therefore, of the hyperactivated genes identified in TNBC, the druggable kinase genes could be good candidates for development of molecularly targeted therapy for TNBC. Table 6 presents the highly expressed kinase genes (Levels 1 and 2) in TNBC compared to normal samples (at least 2-fold expression elevation, Wilcox signed-rank test, FDR < 0.05).

Of the kinase genes in Table 6, CSF1R has the highest expression elevation in TNBC (24.68-fold expression elevation, FDR $= 1.30 \times 10^{-8}$). Previous studies have revealed that overexpression of CSF1R was associated with ipsilateral breast cancer recurrence and poor prognosis of breast cancer.[30] This is in line with our result that CSF1R is highly expressed in TNBC, which often has unfavorable clinical outcome. Therefore, CSF1R could be an important target for TNBC therapy. In fact, it has been shown that CSF1R activity could be inhibited by some small molecule inhibitors such as imatinib, dasatinib, sunitinib, CEP-701, and PKC-412.[31] These compounds may be worth clinical trial for TNBC therapy.

Table 6 presents many kinase genes that belong to the same gene families (including *SRC*, *EPH*, *FLT*, *MAP*, *NTRK*, *PAK*, *PRK*, *RIPK*, and *STK*) that are worth investigation. For example, HCK has the second highest expression elevation in TNBC (12.57-fold expression elevation, FDR $= 1.75 \times 10^{-9}$). The gene encodes a member of the SRC family of tyrosine kinases, which are potential therapeutic targets for TNBC.[32,33] In Table 6, there is another SRC family kinase gene *FGR* that are overexpressed (3.64-fold expression elevation, FDR $= 2.05 \times 10^{-9}$) and amplified in TNBC (1.5-fold copy number gain, FDR $= 1.35 \times 10^{-8}$). Some small molecule inhibitors such as dasatinib have been shown to be effective in TNBC therapy, possibly because they can inhibit the activity of the SRC family kinases.[34] EPHB3, a member of the EPH receptor gene family, has the third highest expression elevation in TNBC (9.43-fold expression elevation, FDR $= 1.91 \times 10^{-8}$). Another EPH receptor family gene, *EPHB1*, is also highly expressed in TNBC (5.23-fold expression elevation, FDR $= 1.29 \times 10^{-9}$). It has been reported that increased expression of the EPH receptor was correlated with more malignant and metastatic tumors,[35] which is consistent with our results.

The kinase genes in Level 2 are especially worthy of note since their hyperactivation in TNBC was confirmed or demonstrated by multiple genomic evidences. For example, FGFR2 (fibroblast growth factor receptor 2) has more than 2-fold higher expression in TNBC (FDR $= 2.94 \times 10^{-6}$) and is targeted by miRNA-410 and miRNA-381, both of which were significantly downregulated in TNBC compared to the normal samples (t test, FDR $< 10^{-6}$). This gene has been found to be hyperactive in breast cancer and is associated with increased breast cancer risk.[36] Another study has shown that FGFR2 was amplified in TNBC cell lines that were highly sensitive to FGFR2 inhibitors.[37] MAPK13 (mitogen-activated protein kinase 13) has more than 4-fold higher expression (FDR $< 1.38 \times 10^{-9}$) and much lower methylation level in TNBC than in the normal samples (FDR $< 3 \times 10^{-7}$). The gene is involved in MAPK pathways that have been suggested to be potential targets for TNBC treatment.[38] TRIB1 has more than 2-fold higher expression in TNBC (FDR $= 1.32 \times 10^{-9}$), 1.4-fold copy number gain (FDR $= 10^{-8}$), and is targeted by miRNA-144, which was significantly downregulated in TNBC compared to the normal samples (t test, FDR $= 5.85 \times 10^{-11}$). This gene plays a role in mediating proliferation, apoptosis, and differentiation in cells through binding to MAPKK signaling proteins of MAPK pathways, and has been suggested as a therapeutic target for prostate cancer.[39] Our results suggest that this gene could be a promising target for TNBC therapy.

In summary, the kinase genes hyperactivated in TNBC provide potential targets for development of molecularly targeted therapy for TNBC.

3.4.2. Identification of the hyperactivated genes that are targets of TNBC-sensitive agents. TNBC is highly concordant with basal-like breast cancer (BLBC), defined by gene expression profiling, in that both share many clinical features such as lack of expression of ER, PR, and HER2, high p53 mutation rate, unfavorable clinical outcome, and so on.[2,8] In addition, the majority of claudin-low tumors are triple negative and with poor prognosis.[40] In a previous study, Heiser et al[41] revealed that different breast cancer subtypes (luminal, basal, HER2-enriched, and claudin-low) exhibited differential sensitivities to most therapeutic compounds by performing a systematic drug screening of breast cancer cell lines. They identified a list of compounds that showed significant subtype specificity (Table 1 of Ref. [41]), in which we found that three of the seven basal-like and claudin-low subtype sensitive compounds target genes in the list of hyperactivated genes we identified. The three compounds include docetaxel, PD173074, and CGC-11047, which could be

Table 6

Kinase-encoding genes highly expressed in TNBC.

Gene*	Name	Fold change	FDR
ACVRL1	Activin A receptor type II-like 1	3.75	1.91E – 09
ALPK1	Alpha-kinase 1	3.37	2.01E – 09
BMPR2	Bone morphogenetic protein receptor, type II (serine/threonine kinase)	2	1.41E – 09
CASK	Calcium/calmodulin-dependent serine protein kinase (MAGUK family)	2.12	1.42E – 09
CLUL1	Clusterin-like 1 (retinal)	4.33	2.76E – 09
CSF1R	Colony stimulating factor 1 receptor	24.68	1.30E – 08
DAPK2	Death-associated protein kinase 2	4.81	2.47E – 09
DDR1	Discoidin domain receptor tyrosine kinase 1	2.69	1.35E – 09
EIF2AK2	Eukaryotic translation initiation factor 2-alpha kinase 2	2.16	5.30E – 09
EPHB1	EPH receptor B1	5.53	1.29E – 09
EPHB3	EPH receptor B3	9.43	1.91E – 08
FGFR2	Fibroblast growth factor receptor 2	2.6	2.94E – 06
FGR	Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene homolog	3.64	2.05E – 09
FLT1	Fms-related tyrosine kinase 1	2.01	1.95E – 08
FLT3	Fms-related tyrosine kinase 3	3.22	1.80E – 09
FRK	Fyn-related kinase	3.4	2.63E – 09
HCK	Hemopoietic cell kinase	12.57	1.75E – 09
INSR	Insulin receptor	2.51	1.81E – 09
IRAK3	Interleukin-1 receptor-associated kinase 3	3.54	1.29E – 09
KDR	Kinase insert domain receptor (a type III receptor tyrosine kinase)	6.16	2.67E – 08
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	5.03	1.31E – 08
LIMK2	LIM domain kinase 2	3.29	2.13E – 09
MAP3K5	Mitogen-activated protein kinase kinase kinase 5	2.07	1.49E – 08
MAPK13	Mitogen-activated protein kinase 13	4.37	1.38E – 09
MYLK	Myosin light chain kinase	3.08	1.30E – 09
NEK11	NIMA (never in mitosis gene a)- related kinase 11	2.25	1.28E – 09
NTRK2	Neurotrophic tyrosine kinase, receptor, type 2	5.24	1.01E – 08
NTRK3	Neurotrophic tyrosine kinase, receptor, type 3	2.27	4.08E – 09
PAK3	p21 protein (Cdc42/Rac)-activated kinase 3	3.09	2.43E – 08
PAK6	p21 protein (Cdc42/Rac)-activated kinase 6	2.63	5.48E – 09
PDGFRB	Platelet-derived growth factor receptor, beta polypeptide	2.94	2.82E – 09
PRKCH	Protein kinase C, eta	2.63	1.29E – 09
PRKX	Protein kinase, X-linked	2.64	1.28E – 09
PTK7	PTK7 protein tyrosine kinase 7	2.33	1.42E – 09
RIPK2	Receptor-interacting serine-threonine kinase 2	2.09	1.31E – 09
RIPK3	Receptor-interacting serine-threonine kinase 3	6.36	9.94E – 09
SCYL3	SCY1-like 3 (*Saccharomyces cerevisiae*)	2.36	1.34E – 09
STK3	Serine/threonine kinase 3 (STE20 homolog, yeast)	2.13	3.38E – 09
STK31	Serine/threonine kinase 31	3.09	1.52E – 07
STK38L	Serine/threonine kinase 38 like	3.13	1.29E – 09
STYK1	Serine/threonine/tyrosine kinase 1	4.71	1.31E – 09
SYK	Spleen tyrosine kinase	3.41	1.82E – 09
TESK1	Testis-specific kinase 1	2.13	1.41E – 08
TRIB1	Tribbles homolog 1 (Drosophila)	3.36	1.32E – 09
ULK2	Unc-51-like kinase 2 (*Caenorhabditis elegans*)	2.53	1.61E – 09

FDR = false discovery rate, TNBC = triple negative breast cancer.

* The genes also belonging to Level 2 are underlined (TRIB1 also belongs to Level 3).

promising in molecularly targeted TNBC therapy (Table 7). In fact, docetaxel has been reported to be effective in TNBC treatment[42]; PD173074 has been shown to be able to impair breast cancer metastasis by inhibiting FGFR signaling[43]; CGC-11047 has been suggested to be preferentially effective against aggressive breast cancer subtypes.[44]

In Table 7, BCL2 is the target of docetaxel that has been used in the neoadjuvant treatment for TNBC.[45] We identified several other *BCL2* family genes that were hyperactivated in TNBC including *BCL2L2* (MR), *BCL2L10* (GE), *BCL2L11* (MR), *BCL2L14* (GE and ME), and *MCL1* (CN). In addition, BCL2L11 and BCL2L12 were found to be mutated in 1 sample, and BCL2A1 has higher expression in TNBC compared to normal samples (mean expression 1.4-fold change, FDR = 0.048). Our results are consistent with previous findings that alterations in *BCL2* family genes were associated with pathogenesis and progression of human cancers.[46–48] Thus, *BCL2* family genes could provide targets for cancer therapy including TNBC.

In another study, Shiang et al[49] identified 224 genes that critically sustain the viability of TBNC cell lines by siRNA screening (Appendix Table A2 in Ref. [49]). Of them, 1 (LAD1), 20, and 58 genes were presented in our Levels 3, 2, and 1 gene list, respectively (Additional File 3, Table S3, http://links.lww.com/MD/B151). The Level 3 gene *LAD1* encodes a protein that may contribute to the stability of the association of the epithelial layers with the underlying mesenchyme. Its role in TNBC is unappreciated, but worth further investigation, since the gene was highly expressed (4-fold expression elevation, FDR = 5.93×10^{-9}), amplified (1.34-fold copy number gain, FDR = $1.03 \times$

Table 7

Compounds that are potentially effective in TNBC therapy.

Compound	Target	Subtype specificity	Target hyperactivated in TNBC*
Etoposide	TOP2A	Claudin-low	No
Cisplatin	DNA cross-linker	Basal/Claudin-low	No
Docetaxel	TUBB1, BCL2	Basal/Claudin-low	BCL2 (MR)
GSK1070916	AURK B/C	Claudin-low	No
PD173074	FGFR3	Claudin-low	FGFR3 (MR)
CGC-11047	Polyamine analog	Basal	†LAMA3 (GE, GM), CYLD (GE, MR), PRPF18 (GE, CN), AMFR (GE), PPP1R2 (GE, MR)
Erlotinib	EGFR‡	Basal	No

EGFR = epidermal growth factor receptor, TNBC = triple negative breast cancer.

* The genomic evidences are shown in parenthesis (GE = gene expression, CN = copy number, MR = miRNA, GM = gene mutation).

† High expression levels of these genes were associated with increased sensitivity of breast cancer cells to CGC-11047[44].

‡ EGFR has higher expression in TNBC than in normal samples (mean expression fold change: 1.4, FDR = 0.023).

10^{-7}), and lower-methylated (β value depression >5%, FDR < 2×10^{-5}) in TNBC compared to normal samples.

3.4.3. Genomic profiles for targets of the agents currently being explored in clinical trials. Currently, there are several targeted agents in development for the treatment of metastatic TNBC.[4] The targets of these agents include VEGF, EGFR, PARP, mTOR, FGFR, JAK2, AR, NOTCH, HDAC, and MET (Table 3 of Ref. [4]). We examined the genomic profiles for these genes in TNBC as shown in Table 8. It can be seen from Table 8 that some of the genes (families) are generally upregulated in TNBC such as EGFR, PARP family, and NOTCH family, while some others are downregulated in TNBC such as VEGF family. It could partially explain that in experimental and clinical trials to test new treatment for TNBC, the agents targeting EGFR and PARP family have shown encouraging results,[50,51] while the agents targeting VEGF showed conflicting results.[45]Table 8 indicates that the FGFR family member FGFR2 could be a good therapeutic target for TNBC relative to the other FGFR family members. The *NOTCH* family genes are consistently upregulated in TNBC, indicating that NOTCH inhibition could be effective in TNBC therapy. In the HDAC family, some genes are hyperactivated in TNBC such as *HDAC2*, *HDAC5*, *HDAC6*, *HDAC9*, and *HDAC11*. Inhibition of them could be promising in TNBC therapy. For the targeted treatments against mTOR, JAK2, AR, or MET, Table 8 shows no strong evidence supporting that they could be effective in TNBC therapy. Certainly, the association between genomic profiles and efficacy of the targeted therapy needs to be confirmed by more clinical experiments with genomic data available.

4. Conclusion

TNBC is high-risk due to its rapid drug resistance and recurrence, metastasis, and lack of targeted therapy. So far, no molecularly targeted therapeutic agents have been clinically approved for

Table 8

Genomic profiles for targets of the agents currently explored in clinical trials.

Gene (family)	Expression	Copy number	Methylation	miRNA	Mutation
VEGF family	VEGFB: down (1.3); VEGFC: down (3.8)	VEGFA: up (1.2); VEGFC: down (1.2)	–	VEGFA: down	0
EGFR	Up (1.4)	Up (1.2)	Hyper	–	0
PARP family	PARP2: down (1.3); PARP3: down (1.5); PARP7: down (1.8); PARP8: up (1.4); PARP9: up (4); PARP10: up (1.6); PARP12: up (2); PARP14: up (2); PARP15: up (1.7); PARP16: down (1.4)	PARP1: up (1.3); PARP3: down (1.2); PARP7: up (1.2); PARP10: up (1.3); PARP11: up (1.2)	PARP6: hyper; PARP8: hypo	PARP7: up; PARP8: up	PARP1: 1; PARP3: 1; PARP4: 1; PARP6: 1; PARP8: 1; PARP9: 1; PARP11: 1; PARP12: 1; PARP15: 1
MTOR	Down (1.2)	–	–	Down	1
FGFR family	FGFR1: down (1.7); FGFR2: up (2.6); FGFR3: down (3); FGFR4: down (3.2); FGFRL1: down (1.7)	FGFR3: down (1.2); FGFR4: down (1.2)	FGFR1: hyper	FGFR1: down; FGFR2: down FGFR3: down	FGFR1: 0; FGFR2: 1; FGFR3: 0; FGFR4: 1; FGFRL1: 0
JAK2	–	–	–	Up	0
AR	–	–	–	Up	0
NOTCH family	NOTCH1: up (1.5); NOTCH2: up (1.5); NOTCH3: up (3.7); NOTCH4: up (3.6)	–	–	NOTCH2: down; NOTCH3: down	NOTCH1: 2; NOTCH2: 1; NOTCH3: 3; NOTCH4: 1
HDAC family	HDAC1: down (1.2); HDAC2: up (1.7); HDAC3: down (1.2); HDAC4: down (1.7); HDAC5: up (1.4); HDAC6: down (1.4); HDAC7: down (1.2); HDAC8: down (1.2); HDAC9: up (1.3); HDAC10: down (2.2); HDAC11: up (1.2)	HDAC3: down (1.2)	HDAC6: hypo; HDAC7: hyper; HDAC8: hypo; HDAC11: hypo	HDAC4: down; HDAC9: up	HDAC2: 1; HDAC5: 1; HDAC6: 2 HDAC9: 1; HDAC10: 1
MET	Down (2)	–	–	up	1

miRNA = microRNA.

TNBC. Treatments that target molecules such as EGFR, VEGF, PARP, and mTOR are still at an early stage of research. It is essential for us to discover new treatment targets for TNBC. The cancer genomics data are becoming an invaluable source for development of molecular targets for TNBC therapy.[8] In the present study, we integrally explore genomic profiles (gene expression, copy number, methylation, miRNA, and gene mutation) in TNBC. To our knowledge, this is the first study that combined the 5 different types of genomic data to molecularly characterize TNBC and identify potential targets for TNBC therapy. We identified hyperactivated genes in TNBC based on multiple genomic evidences, which could significantly contribute to pathogenesis and progression of TNBC. Our results confirm previous findings that TNBC has common molecular profiles with BLBC subtype. Moreover, we revealed that many of the hyperactivated genes in TNBC were also highly active in invasive cancer types or subtypes such as lymphoma, AML, hepatocellular carcinoma and invasive prostate cancer, and stem cells, suggesting that their high activities may contribute to the aggressiveness of cancer.

In the present study, we identified potential molecular targets for TNBC therapy. Some of them such as FGFR2, MAPK13, TP53, SRC family, MUC family, and BCL2 family have been suggested to be potential targets for TNBC treatment by previous studies.[23,27,33,36,38] The others such as CSF1R, EPHB3, TRIB1, and LAD1 could be promising new targets for TNBC treatment for which further investigation is worth doing, whereas their importance in TNBC has not been recognized.

Targeted treatment strategies for TNBC have been developed, some of which were encouraging while others were discouraging.[45] Integrative genomic profiles for TNBC could assist in predicting the effectiveness of a targeted treatment strategy and identifying potential new targets.

In the present study, we treated all the TNBC samples as a single homogeneous group instead of dividing them into several heterogeneous subgroups as shown in Ref. [10]. As a result, the hyperactivated genes we identified could show varied "hyperactivity" across the different subgroups. Dissection of TNBC into different subtypes and discovery of subtype-specific molecular targets for TNBC therapy could be a promising direction for us to make efforts in the future. In addition, based on the same method, using the TCGA and other comprehensive cancer genomic data, we can explore other cancer types to find potential molecular targets for their treatment.

Acknowledgments

We thank Kristin E. Wipfler (PhD candidate in CG's lab) for editing and providing valuable comments and the Bioinformatics and Systems Biology Core at the University of Nebraska Medical Center for providing computational infrastructure. This work was supported by the startup funds to XW from the China Pharmaceutical University, and the startup funds to CG from the University of Nebraska Medical Center.

References

[1] Anders C, Carey LA. Understanding and treating triple-negative breast cancer. Oncology (Williston Park) 2008;22:1233–9. discussion 1239–1240, 1243.
[2] Gluz O, Liedtke C, Gottschalk N, et al. Triple-negative breast cancer—current status and future directions. Ann Oncol 2009;20:1913–27.
[3] Sikov W, Berry D, Perou C, et al. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant weekly paclitaxel followed by dose-dense AC on pathologic complete response rates in triple-negative breast cancer. San Antonio:Breast Cancer Symposium; 2013.
[4] Bayraktar S, Gluck S. Molecularly targeted therapies for metastatic triple-negative breast cancer. Breast Cancer Res Treat 2013;138:21–35.
[5] Rugo H, Olopade O, DeMichele A, et al. I-SPY 2 Site PI'sVeliparib/carboplatin plus standard neoadjuvant therapy for high-risk breast cancer: First efficacy results from the I-SPY 2 TRIAL. San Antonio:Breast Cancer Symposium; 2013.
[6] Patel L, Parker B, Yang D, et al. Translational genomics in cancer research: converting profiles into personalized cancer medicine. Can Biol Med 2013;10:214–20.
[7] Marcotte R, Sayad A, Brown KR, et al. Functional genomic landscape of human breast cancer drivers, vulnerabilities, and resistance. Cell 2016;164:293–309.
[8] Cancer Genome Atlas NetworkNetwork TCGA: Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70.
[9] van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.
[10] Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011;121:2750–67.
[11] Cancer Genome Atlas Research NetworkComprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 2008;455:1061–8.
[12] Lehmann BD, Bauer JA, Schafer JM, et al. PIK3CA mutations in androgen receptor-positive triple negative breast cancer confer sensitivity to the combination of PI3K and androgen receptor inhibitors. Breast Cancer Res 2014;16:406.
[13] Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012; 486:395–9.
[14] Craig DW, O'Shaughnessy JA, Kiefer JA, et al. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. Mol Cancer Ther 2013; 12:104–16.
[15] Wang K, Li M, Hadley D, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 2007;17:1665–74.
[16] Benjami Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995; 57:289–300.
[17] Friedman RC, Farh KK, Burge CB, et al. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009;19:92–105.
[18] Mi H, Muruganujan A, Casagrande JT, et al. Large-scale gene function analysis with the PANTHER classification system. Nat Protoc 2013; 8:1551–66.
[19] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
[20] Baker L, Quinlan PR, Patten N, et al. p53 mutation, deprivation and poor prognosis in primary breast cancer. Br J Cancer 2010;102:719–26.
[21] Lamb R, Ablett MP, Spence K, et al. Wnt pathway activity in breast cancer sub-types and stem-like cells. PLoS One 2013;8:e67811.
[22] Alistar A, Chou JW, Nagalla S, et al. Dual roles for immune metagenes in breast cancer prognosis and therapy prediction. Genome medicine 2014;6:80.
[23] Turner N, Moretti E, Siclari O, et al. Targeting triple negative breast cancer: is p53 the answer? Cancer Treat Rev 2013;39:541–50.
[24] Workman HC, Miller JK, Ingalla EQ, et al. The membrane mucin MUC4 is elevated in breast tumor lymph node metastases relative to matched primary tumors and confers aggressive properties to breast cancer cells. Breast Cancer Res 2009;11:R70.
[25] Mukhopadhyay P, Lakshmanan I, Ponnusamy MP, et al. MUC4 overexpression augments cell migration and metastasis through EGFR family proteins in triple negative breast cancer cells. PLoS One 2013;8: e54455.
[26] Lakshmanan I, Ponnusamy MP, Das S, et al. MUC16 induced rapid G2/M transition via interactions with JAK2 for increased proliferation and anti-apoptosis in breast cancer cells. Oncogene 2012;31:805–17.
[27] Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer 2009;9:874–85.
[28] Wang X, Simon R. Identification of potential synthetic lethal genes to p53 using a computational biology approach. BMC Med Genomics 2013;6:30.
[29] Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nat Rev Cancer 2009;9:28–39.

[30] Kluger HM, Dolled-Filhart M, Rodov S, et al. Macrophage colony-stimulating factor-1 receptor expression is associated with poor outcome in breast cancer by large cohort tissue microarray analysis. Clin Cancer Res 2004;10(1 pt 1):173–7.
[31] Mashkani B, Griffith R, Ashman LK. Colony stimulating factor-1 receptor as a target for small molecule inhibitors. Bioorg Med Chem 2010;18:1789–97.
[32] Peddi PF, Ellis MJ, Ma C. Molecular basis of triple negative breast cancer and implications for therapy. Int J Breast Cancer 2012;2012: 217185.
[33] Mayer EL, Krop IE. Advances in targeting SRC in the treatment of breast cancer and other solid malignancies. Clin Cancer Res 2010;16: 3526–32.
[34] Zhang S, Yu D. Targeting Src family kinases in anti-cancer therapies: turning promise into triumph. Trends Pharmacol Sci 2012;33: 122–8.
[35] Surawska H, Ma PC, Salgia R. The role of ephrins and Eph receptors in cancer. Cytokine Growth Factor Rev 2004;15:419–33.
[36] Fletcher MN, Castro MA, Wang X, et al. Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun 2013;4:2464.
[37] Turner N, Lambros MB, Horlings HM, et al. Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. Oncogene 2010;29:2013–23.
[38] Giltnane JM, Balko JM. Rationale for targeting the Ras/MAPK pathway in triple-negative breast cancer. Discov Med 2014;17:275–83.
[39] Mashima T, Soma-Nagae T, Migita T, et al. TRIB1 supports prostate tumorigenesis and tumor-propagating cell survival by regulation of endoplasmic reticulum chaperone expression. Cancer Res 2014;74: 4888–97.
[40] Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010;12:R68.
[41] Heiser LM, Sadanandam A, Kuo WL, et al. Subtype and pathway specific responses to anticancer compounds in breast cancer. Proc Natl Acad Sci U S A 2012;109:2724–9.
[42] Huober J, von Minckwitz G, Denkert C, et al. Effect of neoadjuvant anthracycline-taxane-based chemotherapy in different biological breast cancer phenotypes: overall results from the GeparTrio study. Breast Cancer Res Treat 2010;124:133–40.
[43] Ye T, Wei X, Yin T, et al. Inhibition of FGFR signaling by PD173074 improves antitumor immunity and impairs breast cancer metastasis. Breast Cancer Res Treat 2014;143:435–46.
[44] Kuo WL, Das D, Ziyad S, et al. A systems analysis of the chemosensitivity of breast cancer cells to the polyamine analogue PG-11047. BMC Med 2009;7:77.
[45] von Minckwitz G, Martin M. Neoadjuvant treatments for triple-negative breast cancer (TNBC). Ann Oncol 2012;23(suppl 6):vi35–9.
[46] Yip KW, Reed JC. Bcl-2 family proteins and cancer. Oncogene 2008;27:6398–406.
[47] Xie M, Doetsch PW, Deng X. Bcl2 inhibition of mitochondrial DNA repair. BMC Cancer 2015;15:586.
[48] Xie M, Park D, You S, et al. Bcl2 inhibits recruitment of Mre11 complex to DNA double-strand breaks in response to high-linear energy transfer radiation. Nucleic Acids Res 2015;43:960–72.
[49] Shiang C. Novel therapeutic targets identified by high-throughput technologies for triple-negative breast cancer. Houston, Texas:Texas Medical Center Library: The University of Texas Health Science Center at Houston; 2013.
[50] Hudis CA, Gianni L. Triple-negative breast cancer: an unmet medical need. Oncologist 2011;16(suppl 1):1–1.
[51] Dent RA, Lindeman GJ, Clemons M, et al. Phase I trial of the oral PARP inhibitor olaparib in combination with paclitaxel for first- or second-line treatment of patients with metastatic triple-negative breast cancer. Breast Cancer Res 2013;15:R88.