Published in final edited form as:
Breast Cancer Res Treat. 2014 January ; 143(1): 57–68. doi:10.1007/s10549-013-2780-y.

Characterizing the heterogeneity of triple-negative breast cancers using microdissected normal ductal epithelium and RNA-sequencing

Milan Radovich[1], **Susan E. Clare**[1], **Rutuja Atale**[1], **Ivanesa Pardo**[1], **Bradley A. Hancock**[1], **Jeffrey P. Solzak**[1], **Nawal Kassem**[2], **Theresa Mathieson**[3], **Anna Maria V. Storniolo**[2,3], **Connie Rufenbarger**[3], **Heather A. Lillemoe**[1], **Rachel J. Blosser**[1], **Mi Ran Choi**[1], **Candice A. Sauder**[1], **Diane Doxey**[1], **Jill E. Henry**[3], **Eric E. Hilligoss**[6], **Onur Sakarya**[6], **Fiona C. Hyland**[6], **Matthew Hickenbotham**[6], **Jin Zhu**[7], **Jarret Glasscock**[7], **Sunil Badve**[4], **Mircea Ivan**[2], **Yunlong Liu**[5], **George W. Sledge**[8], and **Bryan P. Schneider**[2,5]

[1]Department of Surgery, Indiana University Simon Cancer Center, Indiana University School of Medicine, Indianapolis, Indiana

[2]Department of Medicine-Division of Hematology/Oncology, Indiana University Simon Cancer Center, Indiana University School of Medicine, Indianapolis, Indiana

[3]Susan G. Komen for the Cure Tissue Bank, Indiana University Simon Cancer Center, Indiana University School of Medicine, Indianapolis, Indiana

[4]Pathology and Laboratory Medicine, Indiana University Simon Cancer Center, Indiana University School of Medicine, Indianapolis, Indiana

[5]Medical & Molecular Genetics, Indiana University Simon Cancer Center, Indiana University School of Medicine, Indianapolis, Indiana

[6]Life Technologies Corporation, Foster City, CA 94404

[7]Cofactor Genomics, LLC, St. Louis, Missouri

[8]Division of Oncology, Stanford University School of Medicine, Stanford, California

Abstract

Triple-negative breast cancers (TNBCs) are a heterogeneous set of tumors defined by an absence of actionable therapeutic targets (ER−,PR−,HER2−). Microdissected normal ductal epithelium from healthy volunteers represents a novel comparator to reveal insights into TNBC heterogeneity and to inform drug development. Using RNA-sequencing data from our institution and The Cancer Genome Atlas (TCGA) we compared the transcriptomes of 94 TNBCs, 20 microdissected normal breast tissues from healthy volunteers from the Susan G. Komen for the Cure Tissue Bank, and 10 histologically normal tissues adjacent to tumor. Pathway analysis comparing TNBCs to optimized normal controls of microdissected normal epithelium versus classic controls composed of adjacent normal tissue revealed distinct molecular signatures. Differential gene expression of TNBC compared with normal comparators demonstrated important findings for TNBC-specific

Corresponding Author: Milan Radovich, PhD, Assistant Professor, Department of Surgery, Division of General Surgery, Indiana University School of Medicine, 980 W. Walnut St., Room C312, Indianapolis, IN 46202, Phone: 317-278-0189 Fax: 317-274-0396.

CONFLICTS OF INTEREST

M. Radovich (Speaker honoraria from Life Technologies Corp.), E. E. Hilligoss (employee and stock ownership Life Technologies Corp.), O. Sakarya (former employee Life Technologies Corp; current employee, stock ownership, and funding Genomic Health), F. C. Hyland (Employee and Stock Ownership, Life Technologies Corp), M. Hickenbotham (employee Life Technologies), J. Zhu (former employee Cofactor Genomics, LLC), and J. Glasscock (employee and stock ownership Cofactor Genomics, LLC).

clinical trials testing targeted agents; lack of over-expression for negative studies and over-expression in studies with drug activity. Next, by comparing each individual TNBC to the set of microdissected normals, we demonstrate that TNBC heterogeneity is attributable to transcriptional chaos, is associated with non-silent DNA mutational load, and explains transcriptional heterogeneity in addition to known molecular subtypes. Finally, chaos analysis identified 146 core genes dysregulated in >90% of TNBCs revealing an over-expressed central network. In conclusion, Use of microdissected normal ductal epithelium from healthy volunteers enables an optimized approach for studying TNBC and uncovers biological heterogeneity mediated by transcriptional chaos.

Keywords

triple-negative breast cancer; RNA-seq; TCGA; normal breast; adjacent normal; ductal epithelium

INTRODUCTION

TNBC preferentially affects pre-menopausal women and women of African descent and has been plagued by the absence of targeted therapies leading to poor survival[1–5]. Because these tumors do not over-express the estrogen, progesterone, or HER-2 receptors (triple-negative), these patients do not respond to targeted therapies that are successfully used in patients who over-express these proteins. A major impediment to therapeutic development in TNBC is an inadequate understanding of the transcriptional biology of the normal breast as a comparator. The use of microdissected ductal epithelium from healthy women as the optimal control is not commonly used secondary to sample availability from healthy volunteers and laborious sample preparation. Many prior gene expression studies have used undissected reduction mammoplasty or histologically “non-cancerous” tissue adjacent to the tumor. Both of these controls are fraught with problems. Specifically hyperplastic breasts that require surgical reduction may harbour neoplasms or pathological atypia[6–9]. In addition, these tissues are more likely to contain pertubations in global gene expression[10,11], changes in epigenetic markers[12], and loss of heterozygosity[13,14].

Recent studies have begun to shed light on the heterogeneity of TNBC using genome-wide technologies. Work by Lehmann et al. using TNBC gene expression data from publically available microarrays demonstrated that TNBC can be divided into 6 reproducible subtypes (plus an unclassified type), with potential therapeutic implications[15]. On the DNA level, recent reports from Shah et al.[16], and the TCGA[17] using exome sequencing have reported extensive mutational heterogeneity among TNBCs/Basal-like tumors with very low frequency mutations in a variety of genes, with common recurrent mutations restricted primarily to TP53 and the PI3K pathway. In addition, previous studies using copy number analysis have also demonstrated frequent RB1 loss-of-heterozygosity as well as Chromosome 5q loss and 8q, 10p,and 12p gains[18–20]. Building on this knowledge of mutational heterogeneity, we used RNA sequencing (RNA-seq) to analyze TNBCs, donated microdissected normal breast epithelium and adjacent normal tissues to better understand the transcriptional heterogeneity of this disease.

METHODS

RNA from 20 normal frozen breast tissues from healthy pre-menopausal volunteers with no history of disease were procured from the Susan G. Komen for the Cure® Tissue Bank (KTB) at the IU Simon Cancer Center (IUSCC). As ductal epithelium (the presumed origin of breast cancer) comprises a minority of cells in the normal breast, these tissues were laser capture microdissected in order to enrich for epithelial RNA. RNA from 10 frozen TNBCs

was extracted from tissues with high tumor content and did not necessitate microdissection. Normal and TNBC RNA was sequenced on a Life Technologies SOLiD sequencer with subsequent read mapping to the genome using LifeScope 2.5.1[21]. RNA-seq data from the normal tissues is available for download from dbGAP (http://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000644.v1.p1). Data for non-TCGA TNBCs are pending NCBI GEO submission. Normal and TNBC RNA-seq data from Indiana University were then merged with RNA-seq data from 84 TNBCs and 10 adjacent normals from the TCGA downloaded from the UCSC cgHUB database (dbGAP approval #3317: Transcriptional and Mutational Landscape of Triple-Negative Breast Cancer, P.I. Milan Radovich). Samples IDs of all samples used in this study are in Supplementary Table 4. Sequencing data was imported into Partek Genomics Suite for gene expression and statistical analysis, and IPA 9.0 for network and pathway analyses. Full methods of sample preparation, sequencing, bioinformatics analysis, qPCR, and IHC are in the Supplemental Methods. All studies on these samples were approved by the IU Institutional Review Board.

RESULTS

Microdissected normal epithelium is a distinct control compared to adjacent normal tissue

We performed next-generation RNA sequencing on 20 microdissected normal breast tissues from our Susan G. Komen for Cure® Tissue Bank at the Indiana University Simon Cancer Center as well as 10 TNBCs and merged the mapped sequencing data with 84 TNBCs and 10 adjacent normal tissues available from the TCGA (Supplementary Figures 1–2, Supplementary Methods). Unsupervised Principal Components Analysis (PCA) of 14,271 expressed genes demonstrated a significant separation of TNBCs from microdissected normal tissues and adjacent normal tissues illustrating the vast differences in their transcriptomic profiles (Figure 1). In order to better delineate the individual genes that differentiate the tissue types, we compared the expression values between TNBCs, microdissected normal breast tissues, and adjacent normal tissues. When considering a false discovery rate (FDR) < 0.05 with a fold change more than ±2, we report 3,197 differentially expressed genes for TNBC vs. microdissected normal tissue; 3,217 genes for TNBC vs. adjacent normal, and 933 genes for adjacent normal vs. microdissected normal tissue (Figure 2, Supplementary Table 1). To better understand the biological differences between these tissue types, we employed canonical pathway analysis to compare them. Observing the most statistically significant pathways, TNBC vs. microdissected normal tissue reveals key pathways known to be implicated in TNBC biology (BRCA1/DNA damage, immune system, chromosomal abnormalities)[22], whereas TNBC vs. adjacent normal reveals some of the same pathways but others that are not as intuitive to TNBC biology but instead stromal biology (atherosclerosis signalling, hepatic fibrosis). The difference in these tissue types becomes more evident when we examined pathways with genes specific to each comparison (that do not overlap on the Venn diagram). TNBC vs. microdissected normal tissue primarily reveals immune pathways (well known to be implicated in TNBC), with the vast majority of gene ontology biological functions associated with leukocyte and lymphocyte biology. Conversely, TNBC vs. adjacent normal tissue reveals a diverse set of pathways with considerably lower p-values with genes indicative of stroma and with gene ontology biological functions associated with death, edema, angiogenesis, microtubule dynamics, and neuronal and organ development. These observations can be attributed to the fact that the adjacent normals are not microdissected and represent a milleu of various stromal cells. In a further analysis, we also performed a pathway analysis of those genes that were differentially expressed between both TNBC vs. microdissected normal and adjacent normal (Figure 2 Venn diagram overlap of 1,267 genes). This analysis recapitulated pathways seen in both comparators, but was absent of immune pathways, thus excluding potentially important microenvironmental cues observed in the TNBC vs. microdissected

normal comparison. Overall, this data suggests that both types of normal controls discover key genes, but that using microdissected normal tissue provides increased accuracy of understanding gene dysregulation in TNBC, and thus is used as the standard control for the rest of our analyses.

To further support the use of microdissected normal epithelium, we performed an upstream transcriptional regulator analysis which predicts transcriptional regulators that are either "inhibited" or "activated" based on differentially expressed genes (Supplementary Table 2). Atop the list of inhibited transcription factors is TP53, which is known to be mutated in 80% of basal-like breast cancers and is the most common recurrently mutated in gene in TNBC[17]. In addition, we observe RB1 as significantly inhibited in TNBC. Previous data has demonstrated functional loss of RB1 with loss of heterozygosity observed in 72% of basal-like breast cancers[18]. Further, genomic sequencing has also demonstrated mutations within the RB1 gene and an enrichment of somatic mutations within RB-associated protein binding sites in TNBC[16]. As RB1 is a canonical suppressor of the E2F1 transcription factor, our analysis shows significant activation of E2F1 as would be expected. In addition our analysis demonstrates inhibition of the tumour suppressor CDKN2A (p16) most likely due to loss of function of RB1, activation of the MYC oncogene whose network is known to be activated in basal-like breast cancer[17], and activation of the FOXM1 transcription factor (discussed later in the results).

Congruency of gene expression with results of TNBC clinical trials

We next examined genes that had been targeted in clinical trials enriched for TNBC patients, specifically, EGFR, KIT, and PARP1. EGFR and KIT, which have previously been shown to be over-expressed in TNBC by microarray[23] and immunohistochemistry when compared to other breast cancer subtypes, were not differentially expressed and down-regulated, respectively, when compared to normal breast in our study (Table 1). Interestingly, the expression of PARP1, whose inhibitors have shown clinical activity in BRCA1 mutated and sporadic TNBCs, was significantly upregulated compared to normal. To further validate these findings, we assessed the gene expression of EGFR, KIT, and PARP1 in a separate cohort of 26 frozen TNBCs and 10 microdissected normal samples by qPCR (Supplementary Figure 3). The qPCR data from the validation cohort confirmed the findings from the next-generation sequencing of a lack of differential expression of EGFR, downregulation of KIT, and upregulation of PARP1. To further confirm at the protein level, we performed immunohistochemistry (IHC) for EGFR and KIT on 20 normal breast tissues and 11 TNBCs (Supplementary Figure 4). The IHC also demonstrates no difference in EGFR expression and downregulation of KIT in TNBC compared to normal. The lack of transcriptional upregulation (compared with normal breast), and the lack of recurrent activating mutations in these two genes might explain the disappointing outcomes to several clinical trials implementing agents designed to target these pathways (Table 1)[24–28]. To further validate the role of over-expression of drug targets with efficacy, we used data from the Cancer Cell Line Encyclopedia (CCLE) [29] which contains data of 13 TNBC cell lines treated with 24 cancer drugs (Supplementary Table 3). As seen in Supplementary Figure 5, we find that drugs that target genes that are over-expressed in our TNBC vs. Microdissected Normal dataset had significantly lower IC50s in treated TNBC cell lines than those that did not ($p<0.0001$).

Transcriptional chaos contributes to TNBC heterogeneity

In order to gain a better understanding of how individual TNBCs differ from microdissected normal breast tisssues, we compared each of our 94 TNBCs individually versus the set of 20 microdissected normal breast tissues. We then plotted the number of differentially expressed genes per TNBC on a waterfall plot (Figure 3A) to demonstrate that there is a significant

range in the number of dysregulated genes ("transcriptional chaos") between individual TNBCs (1328 – 3594 differentially expressed genes). To validate this transcriptional chaos, we correlated the number of differentially expressed genes for each TNBC with their number of non-silent somatic DNA mutations as reported by the TCGA Broad Firehose application[30]. 76 of the 94 TNBCs had DNA mutational data available. In Figure 3B, there was indeed a significant correlation between transcriptional chaos and non-silent somatic DNA mutations (p=0.0007) suggesting that DNA mutational events play a significant role in the transcriptional chaos that is being observed. To delve deeper into this association, we used analyzed TNBC TCGA data from the Memorial-Sloan Kettering Cancer Center cBioPortal for Cancer Genomics[31,32]. Using the same 76 TNBCs that had DNA mutational data available, we first checked to see if genes that are commonly mutated in TNBC (either by base mutations, amplification, or deletion) are associated with transcriptional chaos. Interestingly, we see no significant association between mutations in TP53 (p=0.57), MYC (p=0.86), PIK3CA (p=0.28), or RB1 (p=0.73) and transcriptional chaos. In addition, we further examined whether there was an association between transcriptional chaos and the fraction of somatic copy number altered genome as reported by cBioPortal and again saw no association (p=0.25). Taken together, we observe that transcriptional chaos is mutationally dictated primarily by the conglomerate of DNA mutations, both common and rare.

We then sought to determine whether this transcriptional chaos can add additional information to the currently known subtypes of TNBC. To better discriminate, we focused on TNBCs within the top and bottom quartiles of transcriptional chaos, and plotted them on an unsupervised PCA (using RPKM values of all expressed genes) (Figure 4A). In addition, to avoid any bias from data merging, we plotted only the TCGA samples using the raw log2 transformed RPKM values. We indeed observed a separation of TNBCs based on unsupervised PCA into low and high transcriptional chaos groups. To further determine whether this observation adds to what is currently known about TNBC heterogeneity, we subtyped our TNBCs using the Vanderbilt TNBCtype tool (https://cbc.mc.vanderbilt.edu/tnbc/) (Supplementary Table 4). We then compared the PCA of the TNBC samples colored by transcriptional chaos to the same PCA colored by Vanderbilt TNBCtype, and demonstrate that transcriptional chaos adds additional information along with the TNBC subtype in explaining the heterogeneity (Figure 4B). Samples were also subtyped by PAM50[33,34], but because only 7 of 94 TNBCs were non-basal, a proper comparison between transcriptional chaos and PAM50 could not be performed. To determine whether known clinical factors can explain transcriptional chaos, we found no significant association of transcriptional chaos with age, stage, or race (Supplementary Table 5). TCGA did not have grade available, though the vast majority of TNBCs are Grade 3 [22], and a survival analysis was not performed secondary to too few reported survival events (7 deaths out of 94 samples). To further understand the nature of the transcriptional chaos, we performed a correlation analysis of all expressed genes with transcriptional chaos (Supplementary Table 6). Analysis of the top positive-correlated genes ($r > 0.5$, $p < 2.57\times10^{-7}$) revealed strong involvement in proliferation, cell cycle, and DNA replication, including: TUBG1, AURKA, EBNA1BP2, FAM58A, CENPA, SUV39H1, CCNB1, RRM2, RNASEH1, DKC1, and NABP2.

Chaos analysis reveals a set of TNBC core genes regulated by a FOXM1 network

Finally, we sought to determine whether there were any transcriptional denominators that served as underlayment for TNBC biology. We filtered the results of our transcriptional chaos analysis to only those genes that were differentially expressed in >90% of our TNBC samples (85 or more of 94 TNBCs). This resulted in 146 genes referred to as the "TNBC core genes" (Supplementary Table 7). Network analysis of the 146 core genes identified a

major hub network regulated by the transcription factor FOXM1. As shown in Figure 5, when focusing only on the 146 core genes, FOXM1 directly regulates 13 of these core genes. When taken to second level interactions, those 13 genes directly regulate an additional 32 core genes. Altogether, the FOXM1 hub network directly regulates 47 of 146 TNBC core genes as known by the IPA database. Of interest, the FOXM1 gene itself is 17.2-fold over-expressed in TNBC compared to microdissected normal. Altogether, these data suggest that FOXM1 acts as a regulator of a substantial number of genes that define the core transcriptional dysregulation present in TNBC.

DISCUSSION

Using differential gene expression and pathway analysis we demonstrate that microdissected normal tissues are an optimal comparator to adjacent normal tissues for studying TNBC gene expression. Of interest, we were able to identify key pathways using both comparators, but adjacent normal tissues added pathways indicative of stroma. This is not surprising, as histologically “normal” tissue adjacent to tumor is comprised of a milieu of stromal cells that complicate a normal epithelium vs. tumor epithelium analysis. This observation is obviously due to the adjacent normal not being microdissected, though use of adjacent normal as a comparator for breast studies is commonly used. Further, factors secreted by tumors can have a substantial effect on the transcriptomes of normal epithelium near the tumor, referred to as “field effect”[35,36]. Indeed, it has been previously demonstrated that microdissected normal epithelium adjacent to tumor were subject to gene expression dysregulation, aberrant methylation, and loss of heterozygosity events[10–14]. In addition, previous data have demonstrated that adjacent normal tissue contains gene expression patterns indicative of wound healing[37] (congruent with our observations of genes involved in edema and angiogenesis in the adjacent normal), as well as can serve as a predictor of clinical outcome[38], further reinforcing abnormal dysregulation in the adjacent normal. A commonly used alternative is the use of reduction mammoplasty tissue, both dissected and undissected. While usually derived from healthy patients, the need for surgery secondary to hyperplasticity combined with the relative occurrence of pathologic atypia[7–9], as well as adipose contamination if not dissected, makes these samples less optimal as controls. Another possible alternative control is matched contralateral normal breast tissue from breast cancer patients. Data is limited on this type of control, but recently published data points to changes in lipid metabolism genes in contralateral normal as biomarkers of ER-specific breast cancer risk[39]. In practical terms, matched contralateral breast tissue is not normally collected at the time of surgery and not widely available for research.

We also observed that the canonical pathway analysis comparing TNBC vs. microdissected normal detected cues of non-epithelial cell-types, in particular, immune cells. While tumor cellularity in our TNBC sample set was high as was required of both TCGA breast cancer tissues[17] as well as the IU TNBC tissues, non-epithelial infiltrating cells are present. While having these non-epithelial cells in the analysis did not mask finding tumor epithelial specific pathways as evidenced by some of the top hits (Role of BRCA1 in DNA Damage Response, Cell Cycle: G2/M DNA Damage Checkpoint, and Cell Cycle Control of Chromosomal Replication), we find having the intrinsic microenvironment in our analysis an advantage. In particular, the intrinsic microenvironment can assist in identifying therapeutic targets that would not be identified if the tumor microenvironment was absent due to microdissection. A key example is PD-1 (Programmed Death 1) which is expressed on the surface of tumor infiltrating immune cells and is up-regulated in our dataset (Gene Symbol PDCD1, Fold-Change=10.23, p=0.0009), has been recently demonstrated to be associated with poor survival in all subtypes of breast cancer including basal-like/TNBC[40]. This is further evidenced by a recently initiated Phase Ib clinical trial testing the PD-1 inhibitor, Lambrolizumab, in solid tumors with a specific focus on TNBC (http://

clinicaltrials.gov/show/NCT01848834). Taken together, we find the role of these non-epithelial cells present in TNBCs (even though a minority population) an important aspect of the analysis. Further, to independently validate our findings, we compared our differentially expressed genes (Supplementary Table 1) to a recently published set of differentially expressed genes derived from 30 microdissected TNBCs and 13 normal ductal epithelium using microarrays [41]. We validated by determining the number of overlapping genes between the two datasets where the direction of the fold-change was the same and the p-value<0.05. We found 80% of the significant genes in the smaller microarray dataset had validated in our larger RNA-seq dataset, providing an independent validation.

To clinically associate our results, we compared the differential expression of previously tested targeted agents to clinical trial outcomes. This was most strikingly illustrated by the fact that some genes previously reported to be over-expressed in TNBC by microarray and immunohistochemistry (e.g. EGFR and KIT) were not up-regulated when compared to normal ductal epithelium in this study [42,43]. In contradistinction, one of the few targeted agents that has shown clinical activity in a randomized trial enriched for patients with TNBC was Iniparib (BSI-201), a PARP inhibitor[44]. The target for Iniparib (PARP1) is 3-fold over expressed in TNBC compared to normal in our study. While the subsequent randomized phase III trial did not support clear and uniform activity for all patients with sporadic TNBC, there were multiple confounding variables including the presumed mechanism of activity[45]. Other trials using agents with robust PARP inhibition in selected populations with TNBC have demonstrated exquisite sensitivity[46–48]. Taken together, our data is in congruence with current clinical trial outcomes of agents that target these proteins (Table 1), and suggests that comparing TNBC to microdissected ductal epithelium versus other comparators may yield better therapeutic targets.

We then sought to determine if we could use microdissected normal epithelium to better understand TNBC transcriptional heterogeneity. We performed a transcriptional chaos analysis by comparing each individual TNBC to the set of 20 microdissected normals, demonstrating a wide range in the number of genes that are dysregulated in each individual TNBC. “Chaos” is a proper term for this analysis as any number of 14,156 genes was dysregulated in at least 2 or more TNBCs. Further the term “chaos” is supported by the observation that transcriptional chaos is associated with the number of non-silent DNA mutations, of which the vast majority are not recurrent. This suggests an interesting link between the burden of non-recurrent, non-silent DNA mutations with observed transcriptional heterogeneity. Of interest, transcriptional chaos was not associated with stage, age, or race, but was correlated with the expression of genes involved in proliferation, cell cycle, and DNA damage repair. Recently, studies using various modalities have sought to subtype TNBCs into distinct molecular subtypes with varying degrees of overlap[15,19,17]. To understand how transcriptional chaos plays a role, we compared our chaos results with TNBCtype and PAM50, and demonstrate that transcriptional chaos adds additional information to molecular subtypes. This suggests that while the commonly used subtyping methods do separate samples into various groups, its the individual uniqueness of each TNBC and its difference compared to normal that also dictates heterogeneity.

Lastly, in the midst of the transcriptional chaos, we sought to determine whether any genes were present that served as transcriptional denominators for TNBC. We identified 146 genes that were dysregulated in >90% of TNBCs (“TNBC core genes”). Strikingly, out of 14,271 expressed RefSeq genes in this study, these core genes represent only 1%. Of this small fraction we demonstrate that over-expressed FOXM1 is a master regulator of a significant fraction of these core genes. FOXM1 is a transcription factor known for its role in mediating cell cycle progression and metastasis[49,50]. Indeed, several of the genes involved in the FOXM1 TNBC core gene network are involved in proliferation and cellular movement

(Figure 5). To support the importance of FOXM1, data from the TCGA has identified activation of FOXM1 as a basal-like specific network when compared to the other intrinsic subtypes[17]. Taken together, these data suggest that targeting FOXM1 or its network members may serve as potential therapeutic targets for TNBC.

In summation, we present a comprehensive and novel characterization of the differential expression of a lethal disease with no FDA-approved targeted therapies using RNA-seq technology. By using microdissected normal epithelium from healthy volunteers we demonstrate the utility of this tissue to uncover novel biological insights into TNBC biology and for informing future drug development. Further, we show that a significant portion of observed transcriptional heterogeneity can be explained by transcriptional chaos that was uncovered only through the use of a normal control.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Mark Mooney, James Elliott, Darryl Leon, and Ryan Richt for discussions of next-generation sequencing and analysis. We also like to thank Benjamin Haibe-Kains for assistance with PAM50 analysis and to thank Carla Bullitt, Stuart Tugendreich, Gordon Janaway, and Bryant Macy for assistance with Ingenuity Pathway Analysis. We also like to thank the IUSCC Tissue Procurement and Distribution Core for providing tissues for IHC and qPCR validations. Lastly, we would like to thank and acknowledge the TCGA for the sample procurement, production of the TNBC RNA-seq data, and the clinical annotation of TCGA samples used in this publication. This work was supported by the Susan G. Komen for the Cure® (S.E.C., G.W.S., A.V.S., S.B., B.P.S.), Breast Cancer Research Foundation (S.E.C., A.V.S.) and the Catherine Peachey Fund (M.R., S.E.C., G.W.S., A.V.S., C.R., B.P.S.).. M.R. was supported by pre-doctoral fellowships from the National Institutes of Health, NRSA 1T32CA111198 Cancer Biology Training Program and 5TL1RR025759 Indiana Clinical and Translational Sciences Institute Career Development Award.

REFERENCES

1. Schneider BP, Winer EP, Foulkes WD, Garber J, Perou CM, Richardson A, Sledge GW, Carey LA. Triple-negative breast cancer: risk factors to potential targets. Clin Cancer Res. 2008; 14(24):8010–8018. doi:14/24/8010 [pii] 10.1158/1078-0432.CCR-08-1208. [PubMed: 19088017]
2. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA. Triple-negative breast cancer: clinical features and patterns of recurrence. Clin Cancer Res. 2007; 13(15 Pt 1):4429–4434. doi:13/15/4429 [pii] 10.1158/1078-0432.CCR-06-3045. [PubMed: 17671126]
3. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MC, Nielsen TO, Moorman PG, Earp HS, Millikan RC. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006; 295(21):2492–2502. doi:295/21/2492 [pii] 10.1001/jama.295.21.2492. [PubMed: 16757721]
4. Stark A, Kleer CG, Martin I, Awuah B, Nsiah-Asare A, Takyi V, Braman M, Quayson SE, Zarbo R, Wicha M, Newman L. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. Cancer. 2010; 116(21):4926–4932. [PubMed: 20629078]
5. Huo D, Ikpatt F, Khramtsov A, Dangou JM, Nanda R, Dignam J, Zhang B, Grushko T, Zhang C, Oluwasola O, Malaka D, Malami S, Odetunde A, Adeoye AO, Iyare F, Falusi A, Perou CM, Olopade OI. Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer. J Clin Oncol. 2009; 27(27):4515–4521. [PubMed: 19704069]
6. Ambaye AB, MacLennan SE, Goodwin AJ, Suppan T, Naud S, Weaver DL. Carcinoma and atypical hyperplasia in reduction mammaplasty: increased sampling leads to increased detection. A

prospective study. Plast Reconstr Surg. 2009; 124(5):1386–1392. doi:10.1097/PRS.0b013e3181b988da 00006534-200911000-00002 [pii]. [PubMed: 20009822]

7. Ishag MT, Bashinsky DY, Beliaeva IV, Niemann TH, Marsh WL Jr. Pathologic findings in reduction mammaplasty specimens. Am J Clin Pathol. 2003; 120(3):377–380. [PubMed: 14502800]
8. Ramakrishnan R, Bhandare D, Fine N, Khan SA, Lal A, Nayar R. Pathologic findings in contralateral reduction mammaplasty specimens in patients with breast cancer. Breast J. 2005; 11(5):372–373. doi:TBJ00059 [pii] 10.1111/j.1075-122X.2005.00059.x. [PubMed: 16174168]
9. Degnim AC, Visscher DW, Hoskin TL, Frost MH, Vierkant RA, Vachon CM, Shane Pankratz V, Radisky DC, Hartmann LC. Histologic findings in normal breast tissues: comparison to reduction mammaplasty and benign breast disease tissues. Breast Cancer Res Treat. 2012; 133(1):169–177. [PubMed: 21881938]
10. Tripathi A, King C, de la Morenas A, Perry VK, Burke B, Antoine GA, Hirsch EF, Kavanah M, Mendez J, Stone M, Gerry NP, Lenburg ME, Rosenberg CL. Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. Int J Cancer. 2008; 122(7):1557–1566. [PubMed: 18058819]
11. Graham K, Ge X, de Las Morenas A, Tripathi A, Rosenberg CL. Gene expression profiles of estrogen receptor-positive and estrogen receptor-negative breast cancers are detectable in histologically normal breast epithelium. Clin Cancer Res. 2011; 17(2):236–246. doi:1078-0432.CCR-10-1369 [pii] 10.1158/1078-0432.CCR-10-1369. [PubMed: 21059815]
12. Yan PS, Venkataramu C, Ibrahim A, Liu JC, Shen RZ, Diaz NM, Centeno B, Weber F, Leu YW, Shapiro CL, Eng C, Yeatman TJ, Huang TH. Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. Clin Cancer Res. 2006; 12(22):6626–6636. doi:12/22/6626 [pii] 10.1158/1078-0432.CCR-06-0467. [PubMed: 17121881]
13. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. Science. 1996; 274(5295):2057–2059. [PubMed: 8953032]
14. Lakhani SR, Chaggar R, Davies S, Jones C, Collins N, Odel C, Stratton MR, O'Hare MJ. Genetic alterations in 'normal' luminal and myoepithelial cells of the breast. J Pathol. 1999; 189(4):496–503. doi:10.1002/(SICI)1096-9896(199912)189:4<496::AIDPATH485>3.0.CO;2-D [pii] 10.1002/(SICI)1096-9896(199912)189:4<496::AID-PATH485>3.0.CO;2-D. [PubMed: 10629549]
15. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest. 2011; 121(7):2750–2767. [PubMed: 21633166]
16. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, Turashvili G, Ding J, Tse K, Haffari G, Bashashati A, Prentice LM, Khattra J, Burleigh A, Yap D, Bernard V, McPherson A, Shumansky K, Crisan A, Giuliany R, Heravi-Moussavi A, Rosner J, Lai D, Birol I, Varhol R, Tam A, Dhalla N, Zeng T, Ma K, Chan SK, Griffith M, Moradian A, Cheng SW, Morin GB, Watson P, Gelmon K, Chia S, Chin SF, Curtis C, Rueda OM, Pharoah PD, Damaraju S, Mackey J, Hoon K, Harkins T, Tadigotla V, Sigaroudinia M, Gascard P, Tlsty T, Costello JF, Meyer IM, Eaves CJ, Wasserman WW, Jones S, Huntsman D, Hirst M, Caldas C, Marra MA, Aparicio S. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012; 486(7403):395–399. [PubMed: 22495314]
17. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490(7418):61–70. [PubMed: 23000897]
18. Herschkowitz JI, He X, Fan C, Perou CM. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. Breast Cancer Res. 2008; 10(5):R75. [PubMed: 18782450]
19. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C, Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012; 486(7403):346–352. [PubMed: 22522925]
20. Dawson SJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. EMBO J. 2013; 32(5):617–628. [PubMed: 23395906]

21. Sakarya O, Breu H, Radovich M, Chen Y, Wang YN, Barbacioru C, Utiramerur S, Whitley PP, Brockman JP, Vatta P, Zhang Z, Popescu L, Muller MW, Kudlingar V, Garg N, Li CY, Kong BS, Bodeau JP, Nutter RC, Gu J, Bramlett KS, Ichikawa JK, Hyland FC, Siddiqui AS. RNA-Seq mapping and detection of gene fusions with a suffix array algorithm. PLoS Comput Biol. 2012; 8(4):e1002464. [PubMed: 22496636]
22. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010; 363(20): 1938–1948. [PubMed: 21067385]
23. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L, Akslen LA, Ragaz J, Gown AM, Gilks CB, van de Rijn M, Perou CM. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res. 2004; 10(16):5367–5374. doi:10.1158/1078-0432.CCR-04-0220 10/16/5367 [pii]. [PubMed: 15328174]
24. Baselga J, Gomez P, Greil R, Braga S, Climent MA, Wardley AM, Kaufman B, Stemmer SM, Pego A, Chan A, Goeminne JC, Graas MP, Kennedy MJ, Ciruelos Gil EM, Schneeweiss A, Zubel A, Groos J, Melezinkova H, Awada A. Randomized Phase II Study of the Anti-Epidermal Growth Factor Receptor Monoclonal Antibody Cetuximab With Cisplatin Versus Cisplatin Alone in Patients With Metastatic Triple-Negative Breast Cancer. J Clin Oncol. 2013
25. Baselga J, Albanell J, Ruiz A, Lluch A, Gascon P, Guillem V, Gonzalez S, Sauleda S, Marimon I, Tabernero JM, Koehler MT, Rojo F. Phase II and tumor pharmacodynamic study of gefitinib in patients with advanced breast cancer. J Clin Oncol. 2005; 23(23):5323–5333. doi:JCO. 2005.08.326 [pii] 10.1200/JCO.2005.08.326. [PubMed: 15939921]
26. Modi S, Seidman AD, Dickler M, Moasser M, D'Andrea G, Moynahan ME, Menell J, Panageas KS, Tan LK, Norton L, Hudis CA. A phase II trial of imatinib mesylate monotherapy in patients with metastatic breast cancer. Breast Cancer Res Treat. 2005; 90(2):157–163. [PubMed: 15803362]
27. Finn RS, Bengala C, Ibrahim N, Roche H, Sparano J, Strauss LC, Fairchild J, Sy O, Goldstein LJ. Dasatinib as a single agent in triple-negative breast cancer: results of an open-label phase 2 study. Clinical cancer research : an official journal of the American Association for Cancer Research. 2011; 17(21):6905–6913. [PubMed: 22028489]
28. Carey LA, Rugo HS, Marcom PK, Mayer EL, Esteva FJ, Ma CX, Liu MC, Storniolo AM, Rimawi MF, Forero-Torres A, Wolff AC, Hobday TJ, Ivanova A, Chiu WK, Ferraro M, Burrows E, Bernard PS, Hoadley KA, Perou CM, Winer EP. TBCRC 001: randomized phase II study of cetuximab in combination with carboplatin in stage IV triple-negative breast cancer. J Clin Oncol. 2012; 30(21):2615–2623. [PubMed: 22665533]
29. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P Jr, de Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palescandolo E, Gupta S, Mahan S, Sougnez C, Onofrio RC, Liefeld T, MacConaill L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, Warmuth M, Finan P, Harris JL, Meyerson M, Golub TR, Morrissey MP, Sellers WR, Schlegel R, Garraway LA. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483(7391):603–607. [PubMed: 22460905]
30. Broad Institute TCGA Genome Data Analysis Center. Breast Invasive Carcinoma (Primary solid tumor cohort)--21 April 2013: Mutation Analysis (MutSig v2.0). Broad Institute of MIT and Harvard. 2013
31. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2(5):401–404. [PubMed: 22588877]
32. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling. 2013; 6(269):pl1. [PubMed: 23550210]

33. Haibe-Kains B, Schroeder M, Bontempi G, Sotiriou C, Quackenbush J. genefu: Relevant Functions for Gene Expression Analysis, Especially in Breast Cancer. R package version 191. 2012 http://compbiodfciharvardedu.
34. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009; 27(8):1160–1167. [PubMed: 19204204]
35. Heaphy CM, Griffith JK, Bisoffi M. Mammary field cancerization: molecular evidence and clinical importance. Breast Cancer Res Treat. 2009; 118(2):229–239. [PubMed: 19685287]
36. Trujillo KA, Heaphy CM, Mai M, Vargas KM, Jones AC, Vo P, Butler KS, Joste NE, Bisoffi M, Griffith JK. Markers of fibrosis and epithelial to mesenchymal transition demonstrate field cancerization in histologically normal tissue adjacent to breast tumors. Int J Cancer. 2011; 129(6): 1310–1321. [PubMed: 21105047]
37. Troester MA, Lee MH, Carter M, Fan C, Cowan DW, Perez ER, Pirone JR, Perou CM, Jerry DJ, Schneider SS. Activation of host wound responses in breast cancer microenvironment. Clin Cancer Res. 2009; 15(22):7020–7028. [PubMed: 19887484]
38. Roman-Perez E, Casbas-Hernandez P, Pirone JR, Rein J, Carey LA, Lubet RA, Mani SA, Amos KD, Troester MA. Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. Breast Cancer Res. 2012; 14(2):R51. [PubMed: 22429463]
39. Wang J, Scholtens D, Holko M, Ivancic D, Lee O, Hu H, Chatterton RT Jr, Sullivan ME, Hansen N, Bethke K, Zalles CM, Khan SA. Lipid metabolism genes in contralateral unaffected breast and estrogen receptor status of breast cancer. Cancer Prev Res (Phila). 2013; 6(4):321–330. [PubMed: 23512947]
40. Muenst S, Soysal SD, Gao F, Obermann EC, Oertli D, Gillanders WE. The presence of programmed death 1 (PD-1)-positive tumor-infiltrating lymphocytes is associated with poor prognosis in human breast cancer. Breast Cancer Res Treat. 2013; 139(3):667–676. [PubMed: 23756627]
41. Komatsu M, Yoshimaru T, Matsuo T, Kiyotani K, Miyoshi Y, Tanahashi T, Rokutan K, Yamaguchi R, Saito A, Imoto S, Miyano S, Nakamura Y, Sasa M, Shimada M, Katagiri T. Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. International journal of oncology. 2013; 42(2):478–506. [PubMed: 23254957]
42. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001; 98(19):10869–10874. 98/19/10869 [pii]. [PubMed: 11553815]
43. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. Nature. 2000; 406(6797):747–752. [PubMed: 10963602]
44. O'Shaughnessy J, Osborne C, Pippen JE, Yoffe M, Patt D, Rocha C, Koo IC, Sherman BM, Bradley C. Iniparib plus chemotherapy in metastatic triple-negative breast cancer. N Engl J Med. 2011; 364(3):205–214. [PubMed: 21208101]
45. O'Shaughnessy J, Schwartzberg LS, Danso MA, Rugo HS, Miller K, Yardley DA, Carlson RW, Finn RS, Charpentier E, Freese M, Gupta S, Blackwood-Chirchir A, Winer EP. A randomized phase III study of iniparib (BSI-201) in combination with gemcitabine/carboplatin (G/C) in metastatic triple-negative breast cancer (TNBC). J Clin Oncol (Meeting Abstracts). 2011; 29 suppl(15) abstr 1007.
46. Puhalla, SL.; Appleman, LJ.; Beumer, JH.; Tawbi, H.; Stoller, RG.; Owonikoko, TK.; Ramalingam, SS.; Belani, CP.; Brufsky, AM.; Abraham, J.; Shephard, SP.; Giranda, V.; Chen, AP.; Chu, E. Two Phase I Trials Exploring Different Dosing Schedules of Carboplatin (C), Paclitaxel (P), and the Poly-ADP-Ribose Polymerase (PARP) Inhibitor, Veliparib (ABT-888) (V) with Activity in Triple Negative Breast Cancer (TNBC); San Antonio Breast Cancer Symposium Poster Discussion; 2012.

47. Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergui-Roelvink M, Mortimer P, Swaisland H, Lau A, O'Connor MJ, Ashworth A, Carmichael J, Kaye SB, Schellens JH, de Bono JS. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med. 2009; 361(2):123–134. doi:NEJMoa0900212 [pii] 10.1056/NEJMoa0900212. [PubMed: 19553641]

48. Tutt A, Robson M, Garber JE, Domchek SM, Audeh MW, Weitzel JN, Friedlander M, Arun B, Loman N, Schmutzler RK, Wardley A, Mitchell G, Earl H, Wickens M, Carmichael J. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet. 2010; 376(9737):235–244. doi:S0140-6736(10)60892-6 [pii] 10.1016/S0140-6736(10)60892-6. [PubMed: 20609467]

49. Wierstra I, Alves J. FOXM1, a typical proliferation-associated transcription factor. Biological chemistry. 2007; 388(12):1257–1274. [PubMed: 18020943]

50. Raychaudhuri P, Park HJ. FoxM1: a master regulator of tumor metastasis. Cancer Res. 2011; 71(13):4329–4333. [PubMed: 21712406]

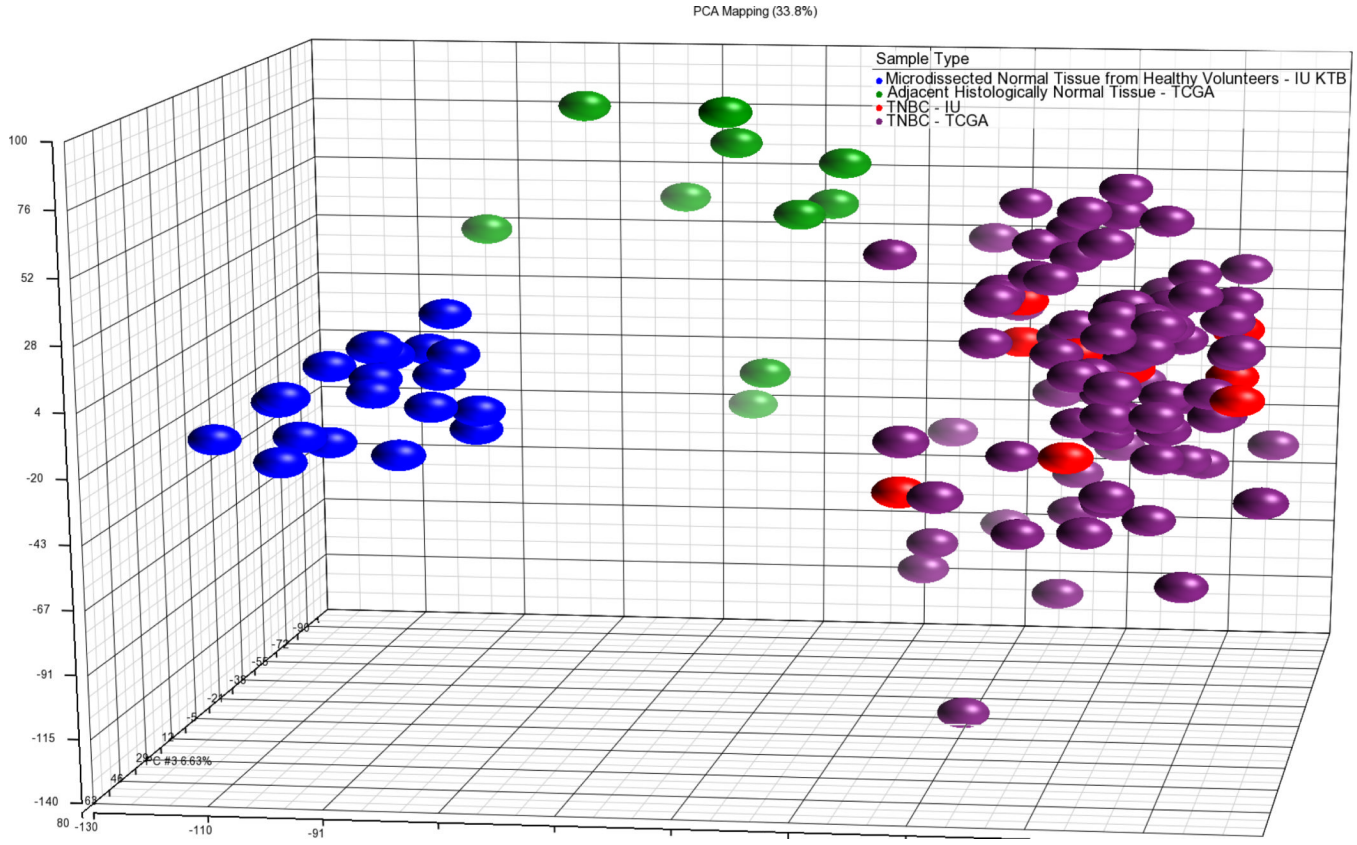
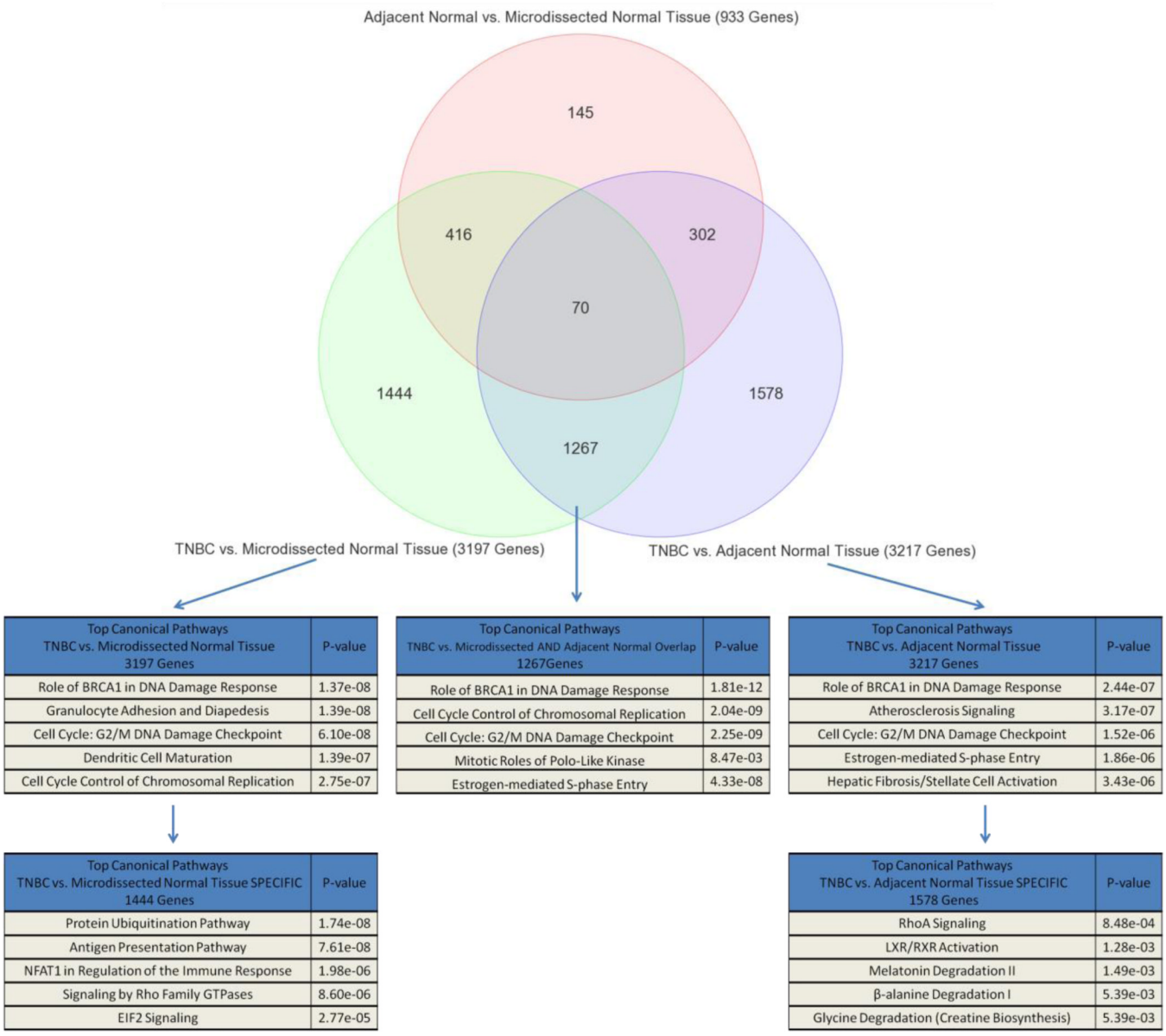


Figure 1.

Unsupervised principal components analysis (PCA) of 14,271 expressed genes demonstrating global gene expression differences between microdissected normal tissue from healthy volunteers, adjacent histologically normal tissue, and triple-negative breast cancers. The sample types cluster into three distinct groups with the TNBCs from IU and from the TCGA clustering together, demonstrating effective merging of the data. IU = Indiana University; KTB = Susan G. Komen Tissue Bank at the Indiana University Simon Cancer Center; TCGA= The Cancer Genome Atlas.



Top Canonical Pathways TNBC vs. Microdissected Normal Tissue 3197 Genes	P-value
Role of BRCA1 in DNA Damage Response	1.37e-08
Granulocyte Adhesion and Diapedesis	1.39e-08
Cell Cycle: G2/M DNA Damage Checkpoint	6.10e-08
Dendritic Cell Maturation	1.39e-07
Cell Cycle Control of Chromosomal Replication	2.75e-07

Top Canonical Pathways TNBC vs. Microdissected AND Adjacent Normal Overlap 1267Genes	P-value
Role of BRCA1 in DNA Damage Response	1.81e-12
Cell Cycle Control of Chromosomal Replication	2.04e-09
Cell Cycle: G2/M DNA Damage Checkpoint	2.25e-09
Mitotic Roles of Polo-Like Kinase	8.47e-03
Estrogen-mediated S-phase Entry	4.33e-08

Top Canonical Pathways TNBC vs. Adjacent Normal Tissue 3217 Genes	P-value
Role of BRCA1 in DNA Damage Response	2.44e-07
Atherosclerosis Signaling	3.17e-07
Cell Cycle: G2/M DNA Damage Checkpoint	1.52e-06
Estrogen-mediated S-phase Entry	1.86e-06
Hepatic Fibrosis/Stellate Cell Activation	3.43e-06

Top Canonical Pathways TNBC vs. Microdissected Normal Tissue SPECIFIC 1444 Genes	P-value
Protein Ubiquitination Pathway	1.74e-08
Antigen Presentation Pathway	7.61e-08
NFAT1 in Regulation of the Immune Response	1.98e-06
Signaling by Rho Family GTPases	8.60e-06
EIF2 Signaling	2.77e-05

Top Canonical Pathways TNBC vs. Adjacent Normal Tissue SPECIFIC 1578 Genes	P-value
RhoA Signaling	8.48e-04
LXR/RXR Activation	1.28e-03
Melatonin Degradation II	1.49e-03
β-alanine Degradation I	5.39e-03
Glycine Degradation (Creatine Biosynthesis)	5.39e-03

Figure 2.

Venn diagram of differentially expressed genes (FDR<5%, Fold-Change>+/−2) compared between each tissue type and their overlap. In order to elucidate the biological differences between these analyses, canonical pathway analysis was performed using Ingenuity Pathway Analysis. The top 5 pathways ranked by p-value are shown for the TNBC vs. Microdissected Normal and the TNBC vs. Adjacent Normal comparisons, as well as for the genes that are specific to those analyses (that did not overlap on the Venn diagram). In addition, the top 5 pathways for those genes that overlapped between TNBC vs Microdissected AND Adjacent Normal Tissues are shown.

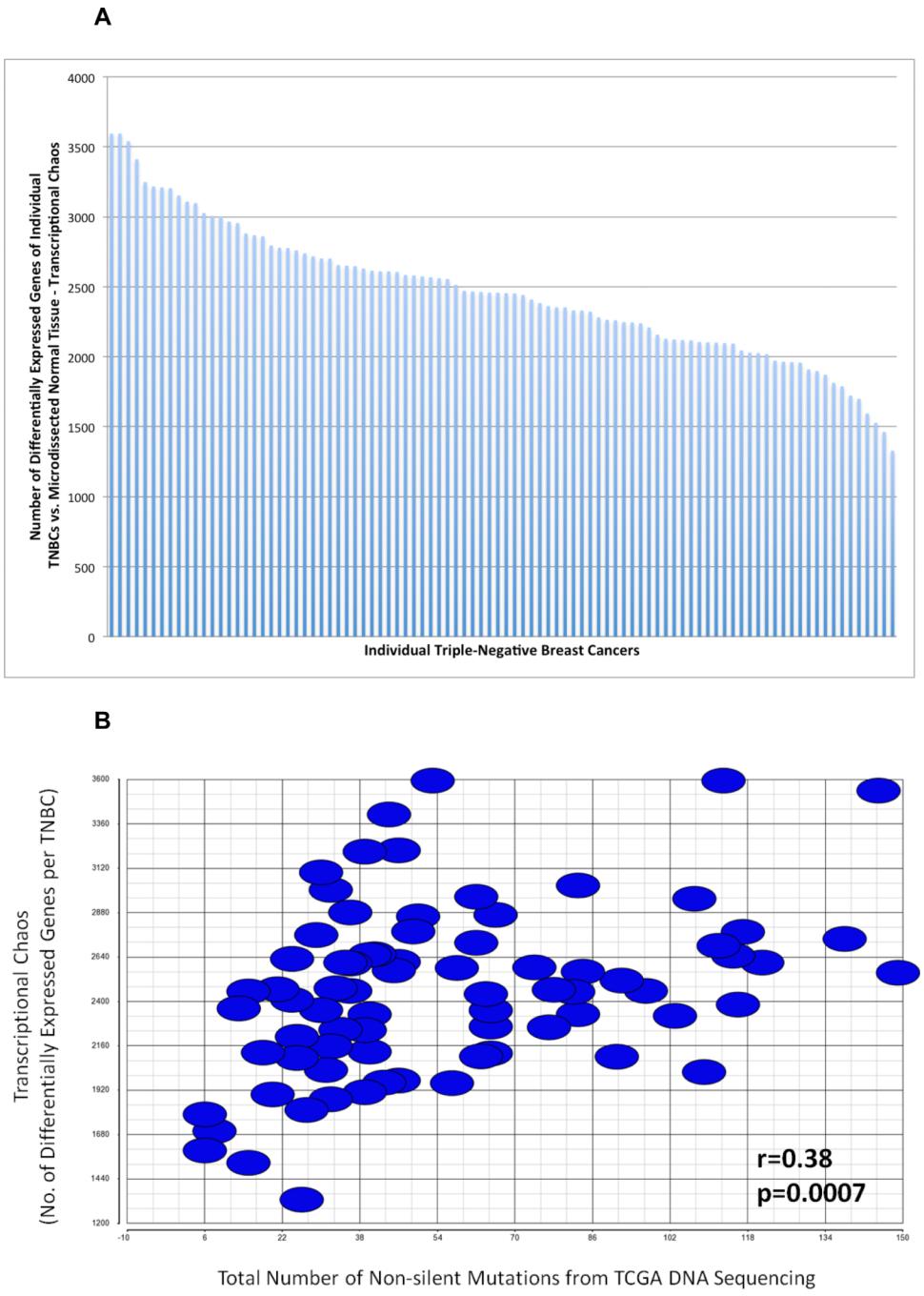


Figure 3.

(A) Waterfall plot comparing the number of differentially expressed genes for each individual TNBC vs. microdissected normal tissues (Transcriptional Chaos). A significant range is observed from 1328 – 3594 differentially expressed genes. **(B)** To provide DNA-level evidence of transcriptional chaos, mutation data from exome sequencing was available on 76 TCGA TNBCs downloaded from the TCGA Broad Firehose Application. A positive correlation is observed between transcriptional chaos and the number of non-silent DNA mutations.

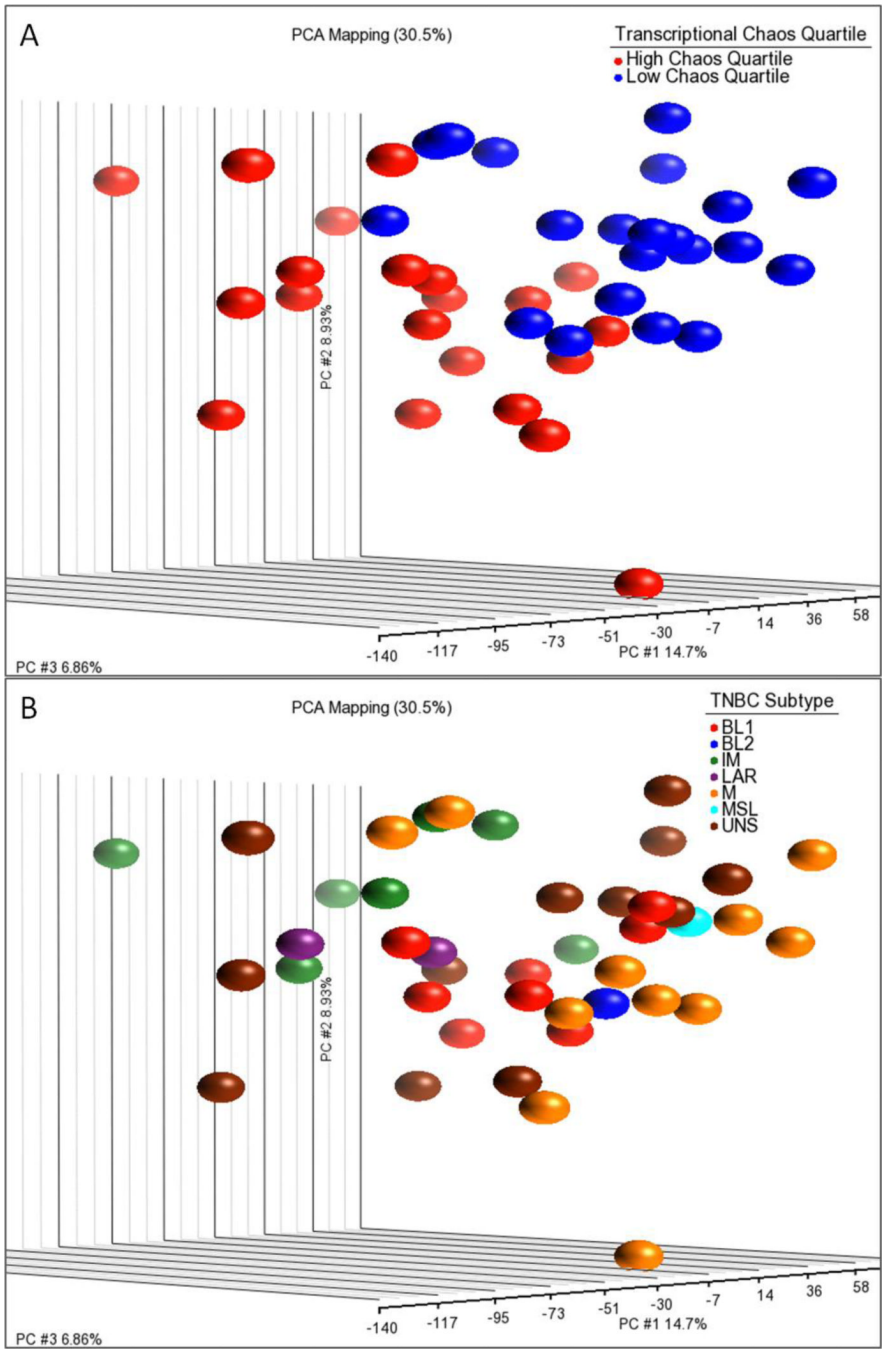


Figure 4.

TCGA TNBC samples representing the top and bottom quartiles of transcriptional chaos were plotted using an unsupervised PCA. A separation of TNBC samples is observed based on transcriptional chaos. **(A)** Samples labelled by top and bottom quartiles. **(B)** The same samples labelled by Vanderbilt TNBC molecular subtype. (BL1: basal-like 1; BL2: basal-like 2; IM: immunomodulatory; M: mesenchymal; MSL: mesenchymal stem-like; LAR: luminal androgen receptor, UNS: Unspecified)

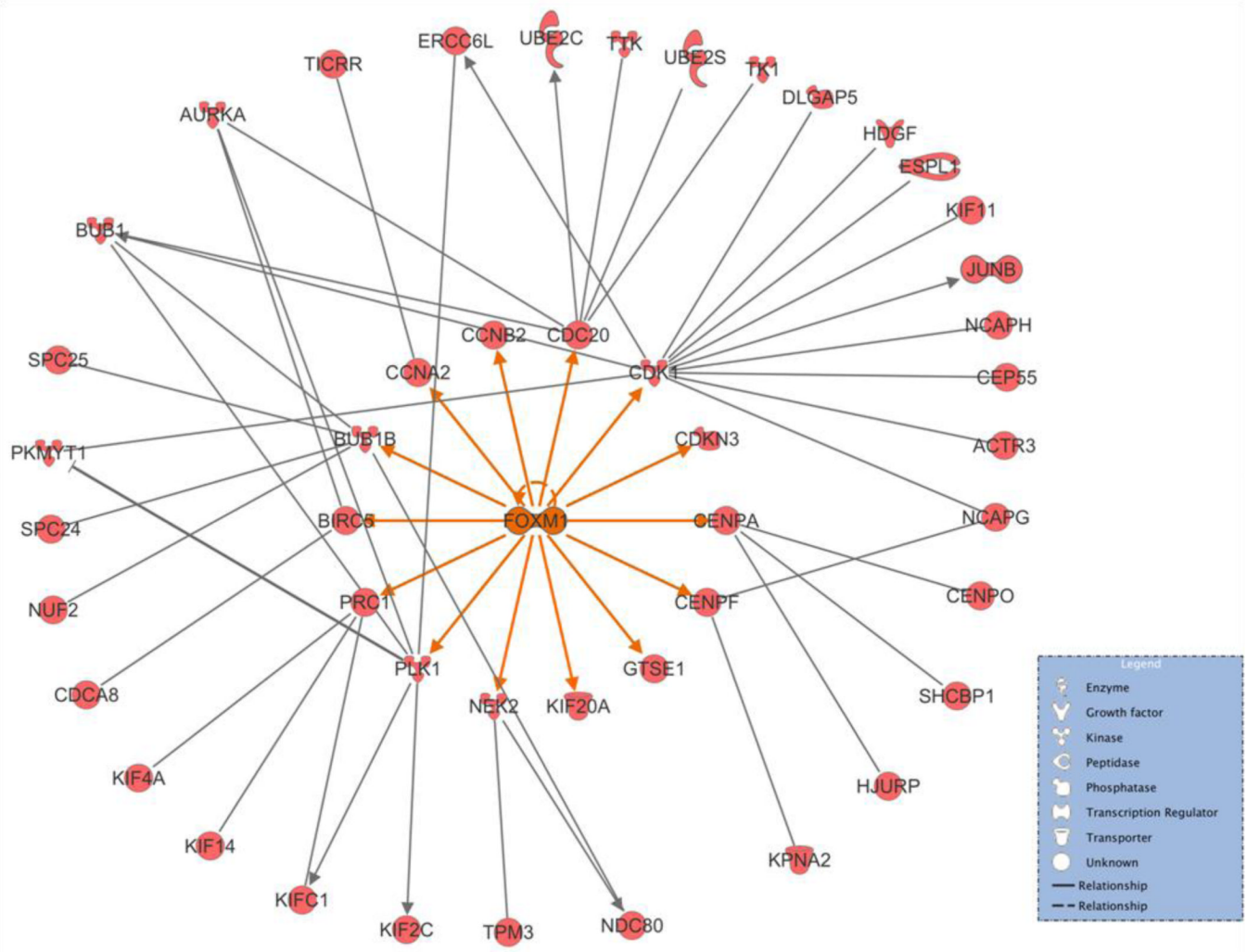


Figure 5.

Network analysis of FOXM1 within the context of the 146 TNBC core genes. Arrows in gold represent direct interactions of FOXM1 with 14 of 146 core genes. Grey arrows represent the second level of interactions adding an additional 32 of 146 core genes. In total 47 of 146 (32%) core genes are regulated by FOXM1, demonstrating that this protein is a regulator of the TNBC core gene set. (Red = Overexpression)

Table 1

Differential gene expression of drug targets clinically tested in enriched TNBC patient populations.

Treatment	Target	RNA Sequencing TNBC vs. Normal	Fold Change (p-value)	Clinical Trial Outcomes
Cetuximab	EGFR	Not Overexpressed	−2.39 (P=0.14, NS)	Negative
Gefitinib				
Imatinib	KIT	Not Overexpressed	-6.45 ($P=4.9\times10^{-3}$)	Negative
Dasatinib				
Iniparib, Olaparib, Rucaparib and others	PARP	Overexpressed	3.32 ($P=1.6\times10^{-7}$)	Some activity (in BRCA mutant TNBCs), newer PARP inhibitors currently in clinical trial