

## RESEARCH ARTICLE

# The overexpression of DNA repair genes in invasive ductal and lobular breast carcinomas: Insights on individual variations and precision medicine

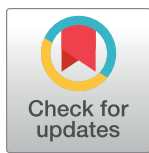
Ruwaah I. Mohamed<sup>1</sup>, Salma A. Bargal<sup>2</sup>, Asmaa S. Mekawy<sup>2,3†</sup>, Iman El-Shiekh<sup>2,3‡</sup>, Nurcan Tuncbag<sup>4</sup>, Alaa S. Ahmed<sup>2,3</sup>, Eman Badr<sup>3,5\*</sup>, Menattallah Elserafy<sup>2,3\*</sup>

**1** Center for Informatics Sciences (CIS), Nile University, Giza, Egypt, **2** Center for Genomics, Helmy Institute for Medical Sciences, Zewail City of Science and Technology, Giza, Egypt, **3** University of Science and Technology, Zewail City of Science and Technology, Giza, Egypt, **4** Graduate School of Informatics, Department of Health Informatics, Middle East Technical University, Ankara, Turkey, **5** Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt

☞ These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

\* [emostafa@zewailcity.edu.eg](mailto:emostafa@zewailcity.edu.eg) (EB); [melsarafy@zewailcity.edu.eg](mailto:melserafy@zewailcity.edu.eg) (ME)



## OPEN ACCESS

**Citation:** Mohamed RI, Bargal SA, Mekawy AS, El-Shiekh I, Tuncbag N, Ahmed AS, et al. (2021) The overexpression of DNA repair genes in invasive ductal and lobular breast carcinomas: Insights on individual variations and precision medicine. PLoS ONE 16(3): e0247837. <https://doi.org/10.1371/journal.pone.0247837>

**Editor:** Suhwan Chang, University of Ulsan College of Medicine, REPUBLIC OF KOREA

**Received:** October 27, 2020

**Accepted:** February 14, 2021

**Published:** March 4, 2021

**Copyright:** © 2021 Mohamed et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data underlying the results presented in the study are available from TCGA-BRCA project (<https://portal.gdc.cancer.gov/projects/TCGA-BRCA>).

**Funding:** The study is funded by a grant from International Centre for Genetic Engineering and Biotechnology (ICGEB) 'CRP/EGY18-05\_EC' to ME. <https://www.icgeb.org/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

In the era of precision medicine, analyzing the transcriptomic profile of patients is essential to tailor the appropriate therapy. In this study, we explored transcriptional differences between two invasive breast cancer subtypes; infiltrating ductal carcinoma (IDC) and lobular carcinoma (LC) using RNA-Seq data deposited in the TCGA-BRCA project. We revealed 3854 differentially expressed genes between normal ductal tissues and IDC. In addition, IDC to LC comparison resulted in 663 differentially expressed genes. We then focused on DNA repair genes because of their known effects on patients' response to therapy and resistance. We here report that 36 DNA repair genes are overexpressed in a significant number of both IDC and LC patients' samples. Despite the upregulation in a significant number of samples, we observed a noticeable variation in the expression levels of the repair genes across patients of the same cancer subtype. The same trend is valid for the expression of miRNAs, where remarkable variations between patients' samples of the same cancer subtype are also observed. These individual variations could lie behind the differential response of patients to treatment. The future of cancer diagnostics and therapy will inevitably depend on high-throughput genomic and transcriptomic data analysis. However, we propose that performing analysis on individual patients rather than a big set of patients' samples will be necessary to ensure that the best treatment is determined, and therapy resistance is reduced.

## Introduction

Breast cancer is the most frequent cancer in women and the second most common cancer overall. Incidence rates vary worldwide from 19.3 per 100,000 women in Eastern Africa to 89.7

**Competing interests:** The authors have declared that no competing interests exist.

per 100,000 women in Western Europe [1]. In Egypt, it is the most prevalent cancer among women (35.1%), and the second within the entire population (17.9%) [2]. Breast cancer is a widely heterogeneous disease as it is classified into multiple subtypes varying in clinical and molecular behavior, consequently leading to distinct prognosis and treatment inferences [3]. The most common histological subtypes of breast cancer are the invasive infiltrating ductal carcinoma (IDC) and lobular carcinoma (LC). Their incidence among women account for approximately 80% in case of IDC and 15% in LC [4,5].

IDC and LC differ on the molecular level and clinicopathological features. 90% of LC patients express estrogen receptor (ER) more often than IDC. In addition, 50–70% of LC patients are progesterone receptor positive (PR+), which is higher than IDC. Yet, only 10% of LC are human epidermal growth factor receptor 2 positive (HER2+) [5]. LC also displays a distinct genomic profile when compared to IDC. This includes a higher frequency of CDH1, PIK3CA, HER2, HER3, TBX3 and FOXA1 mutations in addition to more frequent loss of phosphatase and tensin homolog (PTEN) and amplifications of ESR1 [6–8]. A differential gene expression pattern between the two subtypes was also detected. Genes related to cell adhesion and invasion are among the most significant differentially expressed genes between the two subtypes [9]. LC has lower expression of E-cadherin; encoded by CDH1, which is considered the most significant discriminator between the two subtypes. E-cadherin plays a role in cell adhesion, growth and migration and is regarded as a tumor suppressor in breast cancer [10]. This suggests that the two subtypes may follow diverse mechanisms for invasive growth, where it is promoted in LC by the loss of E-cadherin, while in IDC by the overexpression of cathepsins and osteopontin, and down-regulation of thrombospondin [4]. The expression of more genes; ERBB2, p21, SORBS1, VWF, AOC3, MMRN, ITGA7, CD36, and ANXA1, were also found to be discriminative between LC and IDC [5]. On the other hand, well-differentiated IDC and LC express some genes involved in proliferation and cell cycle (cyclin D1, p16 and p27), apoptosis (MIB1 and BCL2), hypoxia response (HIF-1 alpha) and DNA damage response (MDM2) in a similar manner [11–13].

Defects in DNA repair proteins have long been associated with the development and progression of breast cancer [14]. Matta et al. has shown that DNA repair capacity in women with breast cancer reduces by 60% [15]. Around 25–40% of breast cancers have deficiency in homologous recombination (HR) and approximately 25% have Fanconi anemia-BRCA (FA-BRCA) repair pathway deficiency [16]. Moreover, 0.8–1.7% of breast cancers in women show defects in mismatch repair pathway (MMR) [17,18]. Single nucleotide polymorphisms (SNPs) in the XRCC1 and APE1 base excision repair (BER) genes have also been reported to increase the risk of breast cancer on the level of individuals. However, no population specific data was reported [19,20]. Nucleotide excision repair (NER) deficiency has also been associated with early stage breast cancer. It increases breast cancer risk in women exposed to cigarette smoke [21,22]. Furthermore, the ATM mutations account for more than 7% of breast cancer patients. It was additionally estimated that heterozygous carriers of ATM mutations have a two-fold higher risk of breast cancer [23–25].

Despite the rigorous analyses of DNA repair genes in breast cancer, the transcriptomic changes of DNA repair genes in IDC and LC are not well studied. In this study, we focused on transcriptomic analysis of IDC and LC patient tumors to reveal the expression profile of DNA repair genes in these two breast cancer subtypes. Despite showing distinct transcriptomic profiles, to our surprise, the DNA repair genes were similarly expressed in both IDC and LC. Nevertheless, we stress on the importance of taking individual variations into consideration upon analyzing DNA repair genes in cancer patients.

## Methodology

### Data retrieval

Both RNA sequencing (RNA-seq) and microRNA sequencing (miRNA-seq) data were retrieved from The Cancer Genome Atlas—Breast Invasive Carcinoma (TCGA-BRCA) project (accession date: March 05, 2020) in the format of HTSeq-Counts for RNA-seq and BCGSC miRNA Profiling text files for miRNA-seq. The focus of this study is on females with ductal and lobular neoplasms. We further selected the cases with either invasive ductal carcinoma (IDC) or lobular carcinoma (LC) using the clinical datasheet of the TCGA-BRCA project. The data included 947 cases with both RNA-seq and miRNA-seq data, and 17 cases with RNA-seq data only (total of 964 cases). The 964 cases contributed 1,063 RNA-seq samples (89 normal\_ductal, 771 IDC, and 203 LC) and 1,044 miRNA-seq samples (82 normal\_ductal, 760 IDC, and 202 LC). Note that 7 RNA-seq and 6 miRNA-seq normal lobular samples were excluded because of the small sample size. All the following analyses were done via R version 4.0 [26] and the R-scripts are all available at the GitHub repository.

### Identification of genes involved in DNA repair pathways using REACTOME database

The 310 genes involved in 12 DNA repair pathways in Homosapiens were gathered from the Reactome database [27]. Those pathways are:

1. BER Participating Molecules [R-HSA-73884].
2. DNA Bypass Participating Molecules [R-HSA-73893].
3. DNA damage Reversal Participating Molecules [R-HSA-73942].
4. DNA double strand break response Participating Molecules [R-HSA-5693606].
5. Fanconi Anemia Participating Molecules [R-HSA-6783310].
6. HDR through HRR alone Participating Molecules [R-HSA-5693567].
7. HDR through HRR and SSA Participating Molecules [R-HSA-5693567].
8. HDR through MMEJ Participating Molecules [R-HSA-5685939].
9. HDR through SSA Participating Molecules [R-HSA-5693567].
10. MMR Participating Molecules [R-HSA-5358508].
11. NER Participating Molecules [R-HSA-5696398].
12. Non-homologous end joining Participating Molecules [R-HSA-5693571]

\* Note: No proteins existed in pathway 9 only, as all the players that function in HDR through SSA, play a role in other pathways.

### Differential gene expression and miRNAs analysis

RNA-seq data obtained consisted of the expression raw counts for 60483 transcripts. The org.Hs.eg.db R package (version 3.11.4) was used to convert the Ensemble gene ID to Gene symbol resulting in 25531 genes symbols. The miRNA-seq data obtained consisted of the expression raw counts for 1881 miRNAs. In both analyses, DESeq2 R package (version 1.28.1) [28] was used to collapse the five technical replicas and run the default differential expression analysis on the datasets. DESeq2 utilizes median of ratios method to normalize gene raw counts and

estimate size factors. Negative binomial generalized linear model is used for dispersion estimation to model gene read counts. Differential analysis is performed with two-tailed Wald test. The results for comparing IDC samples to normal\_ductal samples, and IDC to LC samples were obtained with  $|\log_2FC| > 1$  and adjusted P-value  $< 0.05$ , for both DEGs and DEMs. For the DNA repair genes, the DESeqResults object had 25177 genes. Of them, only 285 out of the 310 repair genes were found (see [S3 Table](#) for the full list of DNA repair genes, functions as per Uniprot and excluded genes). Finally, Biomart R package was used to get the description of the Genes from Ensemble database (version 2.44.1) [29]. For the miRNA differential expression analysis, DNA repair targets were validated via miRTarBase 2020 and the miRNAs binding regions on the DEGs mRNAs were obtained from TargetScan [30,31].

### Generation of protein interaction subnetworks of DNA repair related pathways

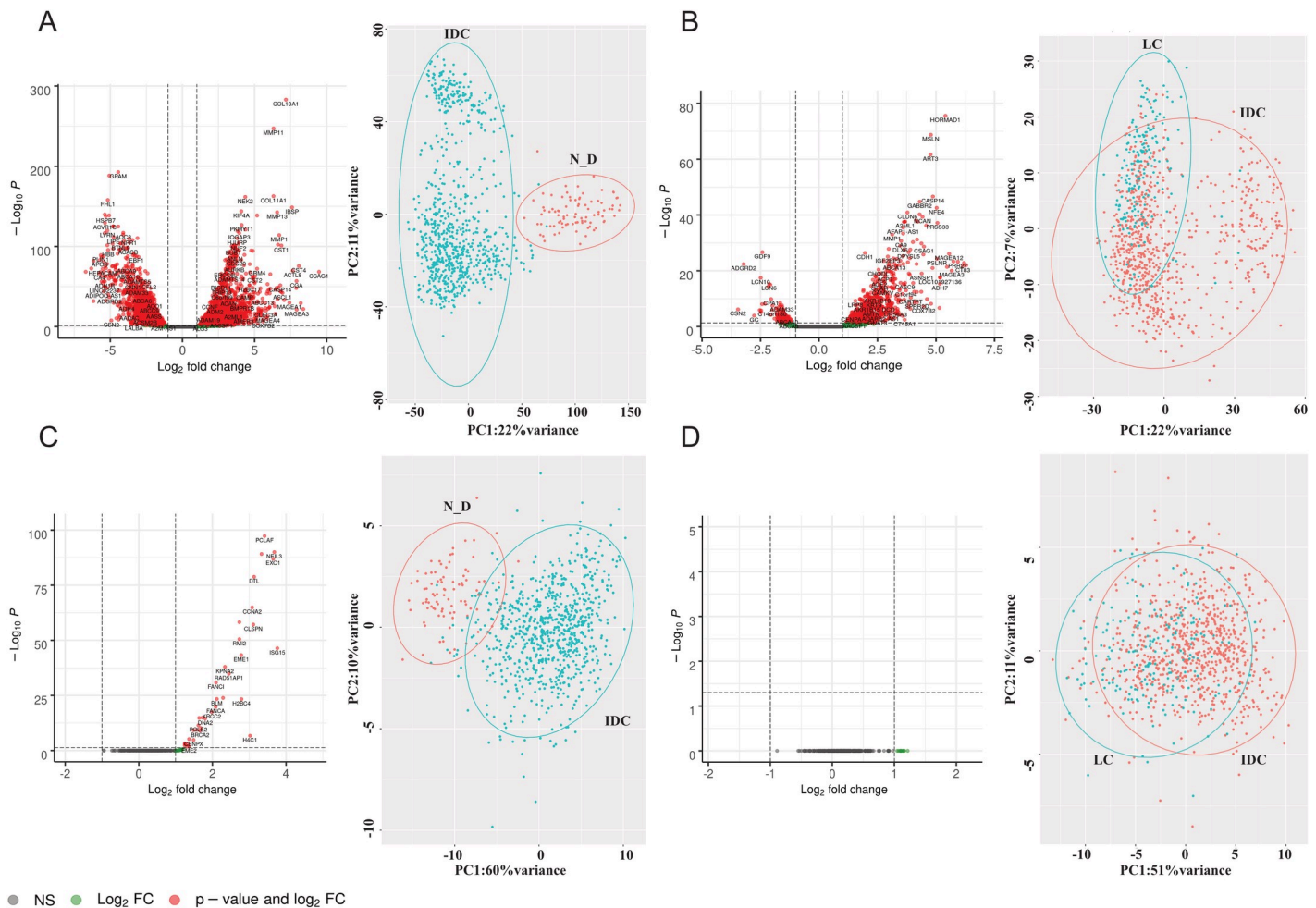
Protein-protein interactions (PPI) are retrieved from Reactome [27] and Biogrid databases [32]. In total, 18572 interactions between 4515 proteins from Reactome and 50759 interactions between 9983 proteins from Biogrid are obtained. These PPI networks are merged and resulted in a reference interactome composed of 65464 unique interactions between 11061 proteins. Proteins associated with DNA repair related pathways in Reactome, are searched for in the reference interactome. For each pathway, a subnetwork is generated. All networks are visualized in Cytoscape [33] and networks are analyzed in networkx Python package [34].

## Results

### IDC and LC have different transcriptomic profiles

In order to determine the differentially expressed genes (DEGs) between normal ductal tissues (normal\_ductal) and IDC tissues of female patients, the TCGA-BRCA dataset was utilized. The dataset includes RNA-seq data of 771 IDC and 89 normal\_ductal samples. A total of 25177 genes have been analyzed and we could identify 3854 total DEGs between normal\_ductal and IDC at the set cutoff of  $|\log_2FC| > 1$  and adjusted P-value  $< 0.05$  ([S1 Table](#)). The 3854 DEGs are displayed in red in the Volcano plot, where the positive  $\log_2$  fold change reflects overexpression in IDC when compared to normal\_ductal, and the negative  $\log_2$  fold change reflects downregulation in IDC in comparison to normal\_ductal. As shown, there are slightly more genes upregulated (2236) than downregulated (1618) ([Fig 1A](#) and [S1 Table](#)). The principle component analysis (PCA) representing the DEGs between the IDC and normal\_ductal samples showed two separate clusters for each of IDC and LC, with almost no overlap ([Fig 1A](#)).

Previous studies analyzed the differences in the transcriptomic profiles of IDC and LC. However, the small sample size was a limitation. For example, Bertucci et al. utilized both array-CGH and cDNA microarray to analyze the DEGs between 29 IDC and 21 LC samples [35]. Zhao et al. also conducted a cDNA microarray study to analyze the DEGs between 38 IDC and 21 LC samples [5]. In our study, we analyzed the DEGs between 771 IDC and 203 LC samples from the TCGA-BRCA dataset. The utilization of RNA-Seq data, which is taking over the array approaches, in addition to analyzing a larger number of samples, should increase our ability to identify more DEGs and reach more accurate conclusions [36]. Since seven normal lobular tissues were only provided in the dataset, we directly compared both breast cancer subtypes instead of utilizing the normal tissues as a reference in each case. We identified a total of 663 DEGs between IDC and LC at the set cutoff of  $|\log_2FC| > 1$  and adjusted P-value  $< 0.05$  ([S2 Table](#)). We have a significantly larger number of upregulated DEGs (591) in IDC than



**Fig 1. Analysis of DEGs and DNA repair DEGs in IDC and LC.** A) Analysis of DEGs in IDC vs Normal\_Ductal. Volcano plot of IDC vs Normal\_Ductal (25,177 genes). The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to normal\_ductal (N\_D), and the negative log<sub>2</sub> fold change reflects downregulation in IDC in comparison to normal\_ductal. PCA of all DEGs (3854 genes) between normal\_ductal vs IDC. B) Analysis of DEGs in IDC vs LC. Volcano plot of IDC vs LC (25,177 genes). The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to LC, and the negative values reflect the downregulation in IDC in comparison to LC. PCA of all DEGs (663 genes) between IDC and LC. C) Analysis of DNA repair DEGs in IDC vs Normal\_Ductal. The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to normal\_ductal, and the negative log<sub>2</sub> fold change reflects downregulation in IDC in comparison to normal\_ductal. PCA of repair DEGs (36 genes) between normal\_ductal vs IDC. D) Analysis of DNA repair DEGs in IDC vs LC; Volcano plot of IDC vs LC (285 genes). PCA of repair DEGs (36 genes) between IDC and LC. Color code for the volcano plots: Red; DEGs ( $|\log_2FC| > 1$  and adjusted P-value  $< 0.05$ ), Green; ( $|\log_2FC| > 1$  and adjusted P-value  $> 0.05$ ), grey; not significant (NS).

<https://doi.org/10.1371/journal.pone.0247837.g001>

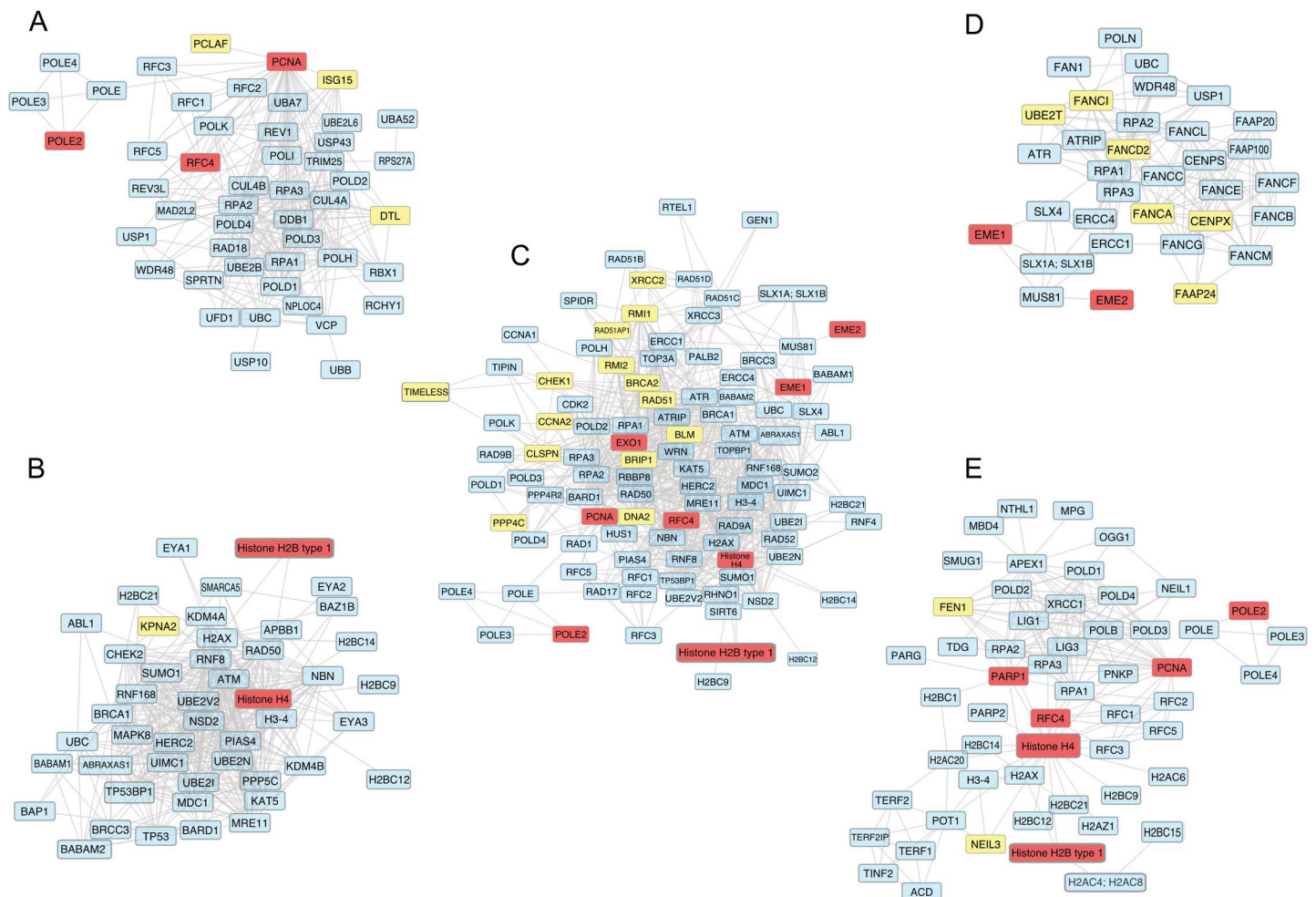
downregulated (72), as shown in red in the Volcano plot; positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to LC, and the negative log<sub>2</sub> fold change reflects the downregulation in IDC in comparison to LC (Fig 1B and S2 Table). The PCA plot shows an overlap between the samples, indicating more similarity between the two cancer subtypes in comparison to the dissimilarity between normal\_ductal and IDC (Fig 1B).

### The DNA repair genes are expressed in a similar manner in IDC and LC

DNA repair genes are known to be associated with the development and progression of breast cancer. In addition, the response of breast cancer patients to therapy is usually affected by the repair machinery in the cells [37]. However, no specific analysis for DNA repair genes in IDC and LC was previously reported to our knowledge. We focused on the analysis of DNA repair

genes in both subtypes, to identify novel DEGs and predict how their expression level can affect response to treatment. We obtained a list of 310 DNA repair genes from the Reactome database [27], then analyzed the differential expression of 285 genes between normal\_ductal and IDC, as 25 genes were excluded (see methods). The 25 excluded genes were histone genes contributing to BER (see S3 Table for the full list of DNA repair genes, functions as per UniProt and excluded genes). We could identify 36 DNA repair genes to be upregulated in IDC when compared to normal\_ductal (S4 Table and Fig 1C). The PCA plot of normal\_ductal and IDC showed a small overlap, which indicates more similarity between the samples upon focusing on the DNA repair expression profile (Fig 1C). Upon comparing the DNA repair genes between IDC and LC, no DEGs were found (Fig 1D and S5 Table). The PCA plot also shows a great overlap between IDC and LC, indicating that repair genes are expressed in a similar manner in the two subtypes (Fig 1D).

Our identified DNA repair DEGs function either in DNA bypass (Fig 2A), DNA double strand break (DSB) response (Fig 2B), Homology-directed repair (HDR) (Fig 2C), Fanconi anemia (FA) (Fig 2D), Base excision repair (BER) (Fig 2E) or in multiple repair pathways. The



**Fig 2. Protein-protein interaction networks of five repair pathways.** A) DNA bypass B) DNA DSB response C) HDR D) FA E) BER. Overexpressed genes are highlighted in red and yellow. Red indicates that the genes function in other repair pathways besides the pathway indicated. Grey edges represent protein-protein interactions.

<https://doi.org/10.1371/journal.pone.0247837.g002>

proteins that were shown to be overexpressed in both IDC and LC are highlighted in either yellow or red. Yellow indicates that the proteins function solely in the indicated DNA repair pathway and red indicates that the proteins function in multiple repair pathways (Fig 2).

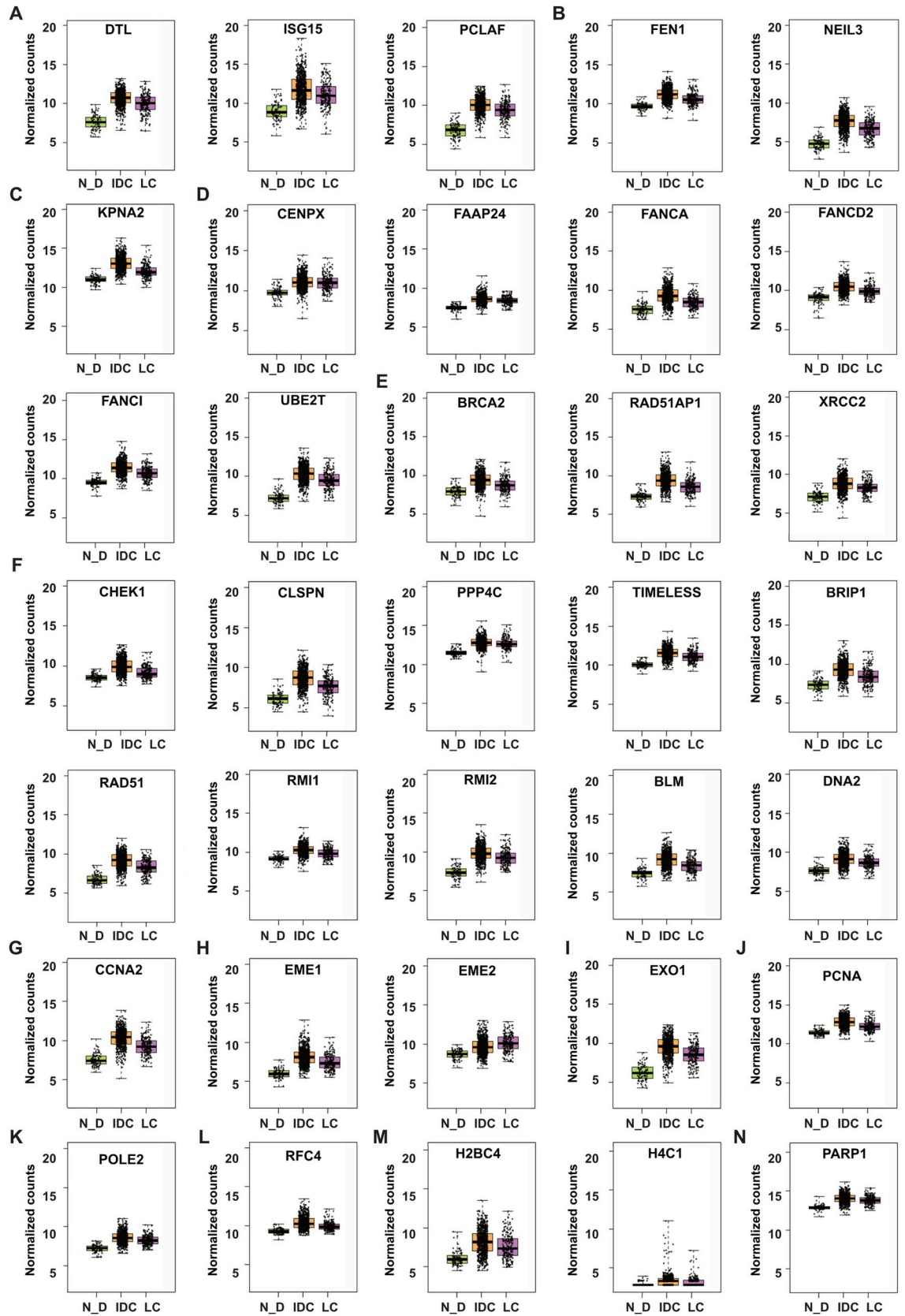
For better visualization of the expression levels of the DNA repair DEGs in different individuals, we combined the normalized counts for each DEG in normal\_ductal, IDC and LC samples in individual box plots (Fig 3). The overexpressed genes include bypass genes; DTL, ISG15 and PCLAF (Fig 3A), BER genes; FEN1 and NEIL3 (Fig 3B), KPNA2, the DNA double strand break (DSB) response (Fig 3C), FA genes; CENPX, FAAP24, FANCA, FANCD2, FANCI and UBE2T (Fig 3D). The HDR players are represented in Fig 3E–3G. BRCA2, RAD51AP1 and XRCC2 that function in HDR through both HRR & SSA are represented in Fig 3E. CHEK1, CLSPN, PPP4C, TIMELESS, BRIP1, RAD51, RMI1, RMI2, BLM and DNA2 functioning through MMEJ/alternative non-homologous end joining (alt-NHEJ) and both HRR & SSA are represented in Fig 3F and CCNA2 functioning through MMEJ is represented in Fig 3G. Finally, the overexpressed genes functioning in multiple repair pathways are EME1, EME2 (Fig 3H), EXO1 (Fig 3I), PCNA (Fig 3J), POLE2 (Fig 3K), RFC4 (Fig 3L), H2BC4, H4C1/Histone 4 (Fig 3M) and PARP1 (Fig 3N). For all repair DEGs, a remarkable variation between the normalized counts in different individuals is observed. This indicates that despite the fact that a significant number of patients overexpress the 36 genes, many patients show downregulated expression or expression levels similar to the normal\_IDC (Fig 3).

To focus on the individual variations, we plotted heatmaps to visualize the DNA repair DEGs across different samples analyzed (Figs 4 and 5). Upon comparing the normal\_ductal and IDC samples, we observed clear subtypes clusters with a very small overlap (Fig 4). Nevertheless, a remarkable variation in the expression levels of the genes in different samples was evident. The IDC samples represented in 'I' were also clustered very closely to the normal samples, indicating that their expression for DNA repair genes is more similar to normal\_ductal than IDC. However, CENPX, ISG15 and PPP4C were clearly overexpressed in the same samples, similar to other IDC samples and distinct from the normal\_ductal. Interestingly, the IDC samples represented in 'IV' show downregulation of the same genes; CENPX, ISG15 and PPP4C, similar to normal\_ductal. We also observed that samples represented in 'II' and 'III' show a significant overexpression for H4C1, which is distinct from normal\_ductal and other IDC samples (Fig 4). Overall, the data indicates clear individual variations among patients despite having a pattern for normal\_ductal and IDC, separately.

Upon comparing IDC and LC, we could not observe clear clusters for the samples, indicating a great similarity in the expression of DNA repair genes (Fig 5). Nevertheless, individual variations were also clearly observed. For example, samples represented in 'I' and 'II' show a clear overexpression of H4C1 in specific IDC patients. This indicates that this evident increase in the expression of H4C1 is specific to IDC not LC. Finally, the samples represented in 'III' also show a huge variation in expression levels of all repair DEGs regardless of the breast cancer subtype (Fig 5).

### Normal\_ductal, IDC and LC show different miRNAs expression profiles

We have analyzed the differentially expressed miRNAs (DEMs) between 82 normal\_ductal and 760 IDC samples. In addition, we analyzed the DEMs between 760 IDC and 202 LC samples. S6 and S7 Tables show 32 DEMs between normal\_ductal and IDC and 7 DEMs between IDC and LC, respectively. We focused only on the miRNAs that were validated experimentally to target our DNA repair DEGs and those with predicted target sequences to infer the possible consequences of the miRNA-gene interaction (Table 1) [30,31]. We could find that hsa-miR-375 and hsa-miR-665 were upregulated and downregulated, respectively in IDC in comparison

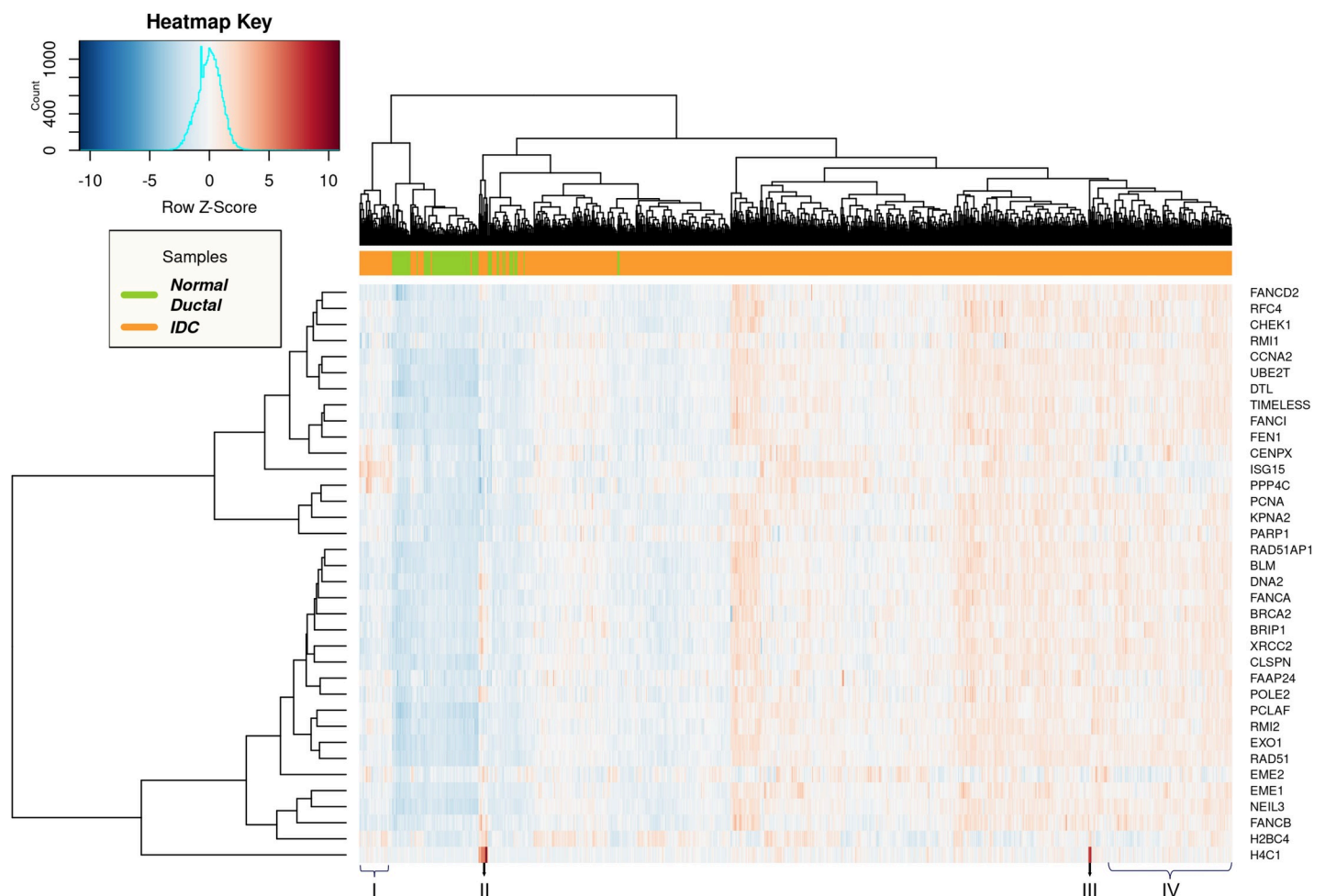




**Fig 3. Normalized counts of the DNA repair DEGs.** A) DNA bypass DEGs, B) BER DEGs, C) DNA DSB response DEG and D) FA DEGs. E-G) DEGs that function in HDR; E) through HRR&SSA F) through MMEJ and both HRR & SSA G) through MMEJ. H-N) DNA repair DEGs functioning in multiple pathways; H) FA, HDR through both HRR & SSA, I) HDR through MMEJ, HRR&SSA, MMR, J) DNA bypass, BER, NER, MMR and HDR through both HRR & SSA, K) DNA bypass, BER, NER and HDR through HRR & SSA, L) DNA bypass, BER, NER and HDR through MMEJ and both HRR&SSA, M) BER, NHEJ, DNA DSB response, HDR through both MMEJ, HRR & SSA and N) BER and NER. Boxplots represent normalized counts in normal\_ductal 'N\_D' (Green), IDC (Orange) and LC (Purple).

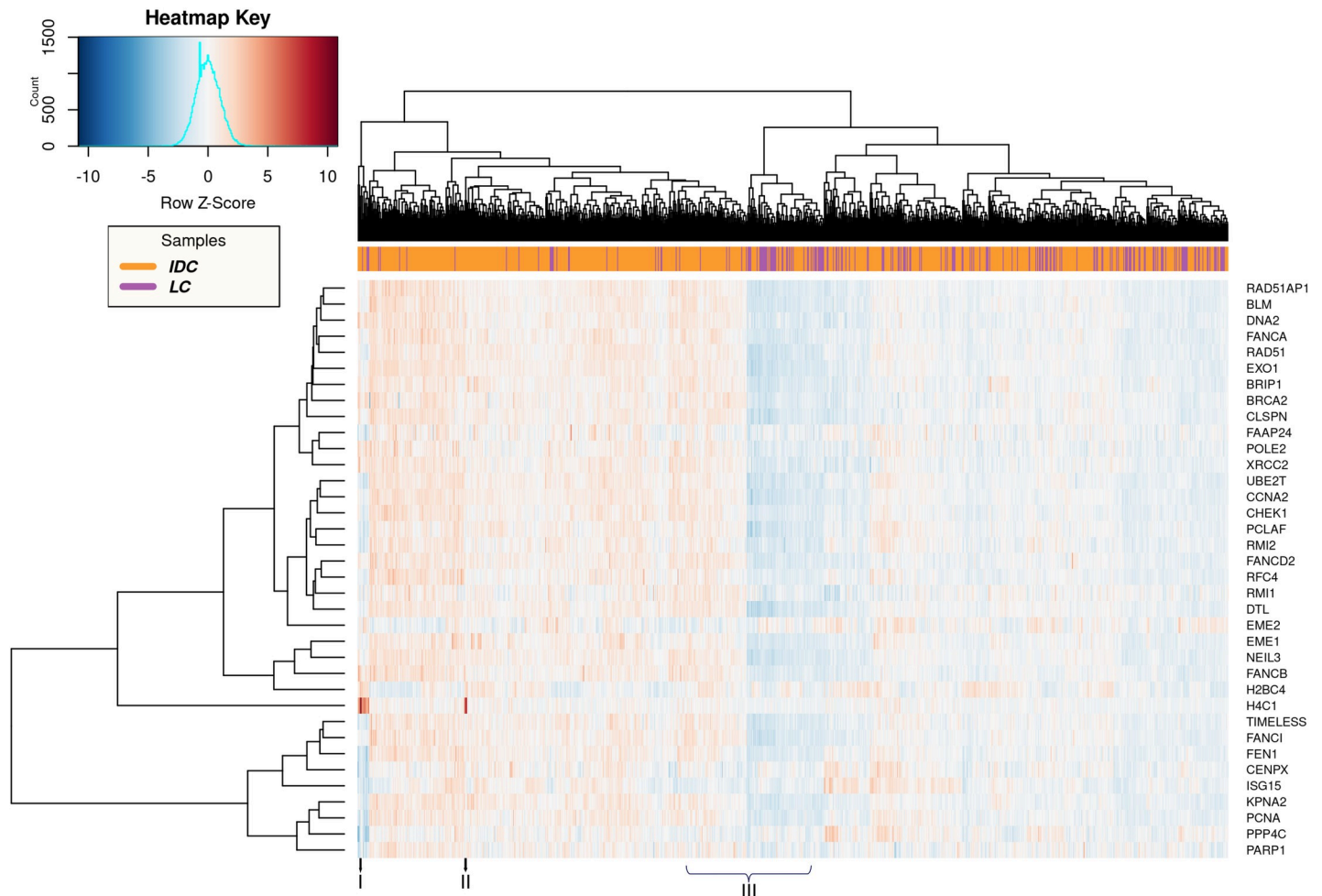
<https://doi.org/10.1371/journal.pone.0247837.g003>

to normal\_ductal (S6 Table and Table 1). No significant differences were detected for both miRNAs between IDC and LC, suggesting similar expression in both subtypes. Moreover, hsa-miR-577 was found to be upregulated in IDC in comparison to LC (S7 Table and Table 1). hsa-miR-375, hsa-miR-665 and hsa-miR-577 were validated to target EXO1, DNA2 and PARP1, respectively (Table 1) [31]. The three miRNAs target the 3'UTR of the respective mRNAs [30], which can result in either degradation, translation repression or translational upregulation in some cases [38,39]. Looking carefully at the differential expression levels of the miRNAs in different individuals, we could find a clear variation among different patients (Fig 6).



**Fig 4. Heatmap representing the 36 DNA repair DEGs in normal\_ductal vs IDC.** Normal Ductal/normal\_ductal (Green) and IDC (Orange).

<https://doi.org/10.1371/journal.pone.0247837.g004>



**Fig 5. Heatmap representing the 36 DNA repair DEGs in IDC vs LC. IDC (Orange) and LC (Purple).**

<https://doi.org/10.1371/journal.pone.0247837.g005>

### Cohort analysis did not identify BRCA1 and TP53 as DEGs, despite huge individual variations

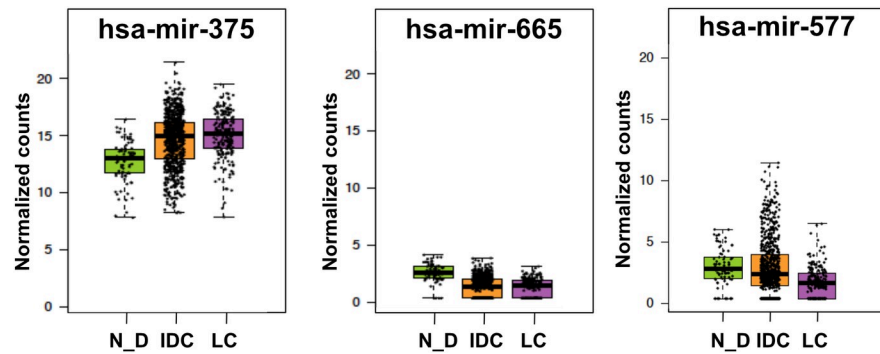
In this study, we did not use a restrictive cutoff so that we increase the number of DEGs between the samples. However, despite using a cutoff of  $|\log_2FC| > 1$  and adjusted P-value  $< 0.05$ , we could not identify some DNA repair genes of known association to breast cancer;

**Table 1. Selected DEMs that target DNA repair DEGs.**

miRNA	Target DNA repair Gene	miRNA binding region (TargetScan)	Expression level of miRNA	Expression level of DNA repair gene	Validation method
hsa-mir-375	EXO1	3'UTR	Up-regulated in tumor	Gene Up-regulated	Microarray
hsa-mir-665	DNA2	3'UTR	Down-regulated in tumor	Gene Up-regulated	NGS
hsa-mir-577	PARP1	3'UTR	Up-regulated in IDC in comparison to LC	Gene Up-regulated	NGS

The miRNA, its target DNA repair gene and the validation methods were obtained from miRTarBase 2020 [31]. The miRNAs binding regions on the DEGs mRNAs were obtained from TargetScan [30].

<https://doi.org/10.1371/journal.pone.0247837.t001>



**Fig 6. Selected DEMs that are functionally validated to target DNA repair DEGs.** Boxplots represent normalized counts in normal\_ductal 'N\_D' (Green), IDC (Orange) and LC (Purple).

<https://doi.org/10.1371/journal.pone.0247837.g006>

such as BRCA1 and TP53 [40–43]. Therefore, we took BRCA1 and TP53 as examples for demonstrating how cohort analysis may not identify certain genes as DEGs, while they could be potentially associated with IDC and LC (S1 Fig). Observing the distribution of the normalized counts in IDC and LC, it is evident that multiple patients have both genes downregulated and other patients show overexpression as well. The huge distribution of counts is also more evident in IDC, which contains a greater number of samples (S1 Fig). Therefore, we believe not considering BRCA1 and TP53 as DEGs in IDC and LC samples could be owed to the inter-individual heterogeneity rather than the irrelevance of the genes analyzed to IDC and LC (S1 Fig). This can also apply for many other genes that we did not focus on in this study.

## Discussion

Our transcriptomic analysis of RNA-Seq data of normal\_ductal and IDC have revealed 3854 DEGs. In addition, 663 genes were differentially expressed between IDC and LC. We compared our results to the previously mentioned studies; Bertucci et al. and Zhao et al., that investigated the transcriptomic differences between IDC and LC [5,35]. Consistent with both studies, CDH1 was upregulated in IDC in comparison to LC [5,35]. Moreover, CIDEA and MFAP4 were overexpressed in LC as reported in Bertucci et al. [35], PLIN1 was overexpressed and FADS2, HIST1H2AL/H2AC11, HIST1H2BJ/H2BC11 and HIST1H3D/H3C8 were downregulated in LC as reported in Zhao et al. [5]. Nevertheless, we could find that GDPD2 is downregulated in LC, which contradicts previous reports showing its upregulation [35]. In addition, contradictory to our study, both Bertucci et al. and Zhao et al. reported that ALDH1A1, FABP4 and VWF are upregulated in LC. We found FABP4 to be downregulated in LC, while the other two genes were not differentially expressed in IDC and LC. The remaining DEGs reported by either study were also not observed in our study and vice versa (S2 Table). The discrepancies in the results can be owed to our utilization of RNA-Seq data and the larger number of samples analyzed in our study, which is most likely to our advantage as the analysis should be more reliable [36]. In addition, the utilization of different patients' samples can also lie behind the differences observed.

Previous studies have focused on analyzing DNA repair genes in breast cancer. For example, a previous study has shown that 21 DNA repair genes were differentially expressed in Hispanic breast cancer female patients using DNA microarray [44]. In line with our results, overexpression of NEIL3, EME1, PCNA and RAD51 genes were reported, but the study did not report our other DEGs and vice versa. The study however, did not focus on IDC and LC and also focused on a specific population, which could explain the discrepancies in the results.

In our study, we analyzed DNA repair genes in IDC versus (vs) LC specifically. 36 DNA repair genes were found to be overexpressed in both subtypes, which are all previously associated with breast cancer.

Here, we discuss the previously reported associations. PCLAF and ISG15 overexpression are associated with poor prognosis of breast cancer, and DTL knockdown decreases breast cancer cells' proliferation and metastasis [45–48]. FEN1 overexpression increases breast cancer progression and its transcription is also activated following anti-cancer treatments. Therefore its inhibition is a potential approach to overcome resistance mechanisms [49–53]. NEIL3 is also overexpressed in breast invasive carcinoma and its upregulation positively correlates with the decrease in the survival of triple negative breast cancer (TNBC) patients [54–56]. KPNA2, which encodes for Karyopherin- $\alpha$ 2 is also highly expressed in breast cancer and its expression is associated with aggressiveness, poor outcomes in addition to chemotherapy and radiotherapy resistance [57–61]. The FA genes we identified as DEGs; CENPX, FAAP24, FANCD2, FANCI, UBE2T and FANCA are all overexpressed in breast cancer [62–65]. High FANCD2 expression correlated with poor breast cancer patients' outcomes. However, 10–20% of breast cancer patients also show loss of FANCD2 expression [66]. Finally, UBE2T knock-down also suppresses tumor growth [62].

Germ-line mutations in BRCA2 act as a major cause of hereditary breast cancer and increase the risk of its early-onset [16,67]. The BRCA2 mutation frequency was reported to be 1–3% in breast cancer [43]. RAD51AP1 is overexpressed in breast cancer stem cells population (BCSCs) and its knockout/down reduced cancer stem cells population in breast, lung, and colon cancer mouse models and improved chemotherapy and radiotherapy [68,69]. XRCC2 is a breast cancer susceptibility gene [70]. Contradictory to our data, XRCC2 was reported to be downregulated in lymph node metastatic breast cancer tissues [71].

The overexpression of CHEK1 was postulated to lead to tumor development and a risk factor in prognostics [72]. In ER- and TNBC, CHEK1 leads to shorter survival following chemotherapy, and its absence was correlated with better outcome [73]. Thus, Chk1 inhibitors such as Prexasertib might serve as a promising therapeutic agent [74,75]. Overexpression of CLSPN coding for Claspin was reported in ER- and/or PR- breast cancer [76]. Its overexpression is a radiotherapy resistance marker in metastatic lung cancer [77], which could be potentially applicable for IDC and LC as well. Timeless promotes breast cancer progression and contributes to poor prognosis [78]. ER alpha-positive (ER $\alpha$ +) breast cancer relapsed patients treated with tamoxifen overexpress Timeless, suggesting its contribution to tamoxifen resistance [79]. Inhibition of the overexpressed PPP4C in breast cancer increases cisplatin sensitivity [80]. Furthermore, RAD51 overexpression positively correlates with tumor grading in IDC, but an inverse relationship was found with estrogen-receptor status [81]. Mutations in BRIP1 increases the risk of breast and ovarian cancer, and its impairment accounts for some breast cancer familial cases [82,83]. Moreover, it is overexpressed in IDC followed by LC [84]. RMI2 is also upregulated in breast cancer [85] and SNPs in BLM and RMI1 were associated with breast carcinoma [86]. Controversially, BLM was reported to be overexpressed in aggressive clinicopathological breast cancer phenotypes, and also reported to be downregulated in other breast cancer subtypes. Therefore, BLM could be a promising biomarker for subtype identification [87]. DNA2 overexpression was reported in a breast cancer cohort with significant higher expression in basal-like breast cancer more than other subtypes, and it positively correlated to metastasis. Additionally, DNA2 partial depletion decreases breast cancer tumorigenicity [88]. CCNA2 was also reported to be overexpressed in breast cancer and utilized as a prognostic biomarker for ER+ subtype. CCNA2 overexpression is correlated with anti-estrogen tamoxifen drug resistance that is usually used for treating ER+ patients. However, CCNA2 repression to reverse the tamoxifen resistance or prevent it is still under study [89,90].

EME1 Ile350Thr variant in Southern Chinese females is significantly associated with susceptibility and early onset of breast cancer and its overexpression is suggested to lead to cisplatin resistance [91,92]. Additionally, EME2 is differentially expressed in tamoxifen resistant breast cancer cells and associated with poor outcomes in patients who did not receive radiotherapy, but not in patients who were subjected to radiotherapy [93]. EXO1 polymorphism is linked to breast cancer susceptibility and the expression level is associated with poor prognosis [94,95]. EXO1 inhibitors could diminish the repair of IR-induced DSBs, which possibly improve radiotherapy and chemotherapy [96]. PCNA is usually known as a proliferating marker in breast cancer. However, there is no evidence for correlation of its overexpression with breast cancer progression [97]. Nevertheless, PCNA expression is a biomarker for predicting high risk of relapse in patients with lymph node-negative breast cancer [98]. A unique acidic form of PCNA was revealed to be present in malignant breast cancer due to post translation modification alteration [97]. An AATT deletion in intron 18 of the POLE2 subunit of Polymerase  $\epsilon$  is associated with breast cancer, and POLE2 is suggested to contribute to lapatinib resistance in HER2+ breast cancer patients with acquired lapatinib resistance [99,100]. Furthermore, RFC4 overexpression contributes to the development of breast cancer [72]. H2BC4 and H4C1 are also overexpressed in breast cancer and the first is overexpressed in its metastatic relapse [101,102]. Finally, Poly-ADP-ribose polymerase 1 (PARP1) was shown previously to be upregulated in breast cancer, acting as an independent biomarker for poor prognosis [103].

Overall, our identified DNA repair DEGs show strong association to breast cancer. But here we report their overexpression in both IDC and LC breast cancer subtypes in particular. Rad51, NEIL3 and BLM, which were linked to invasive breast cancer and BRIP1 that was associated with IDC and LC, are the only genes to our knowledge that were previously linked to invasive breast carcinoma [81,84]. To confirm that most of the transcripts of the 36 genes are functional variants that can potentially impact cellular functions upon their overexpression, we gathered the simple somatic mutations (SSMs) data for the 36 genes in the TCGA-BRCA cases of the TCGA project (S8 Table). The data clearly indicates that SSMs are rare in the samples analyzed for SSMs; 986 out of 1098. On the contrary, the CNV gains are high for our genes. The overexpression of the genes in some of the cases analyzed can be owed to the CNV gain. Interestingly, for some genes there were a significant number of patients who suffer from CNV loss. This could explain the individual variations we have seen upon differential expression analysis.

Regarding the miRNA expression profiles in normal\_ductal, IDC and LC, we believe reaching a conclusion from cohort analysis is very difficult. Analysis of their expression in individual patients will provide much more accurate conclusions. Previous studies have also reported variable expression levels in breast cancer when it comes to the DEMs represented in Table 1. For miR-375, its downregulation suppresses epithelial-to-mesenchymal transition in invasive basal-like tumor cells and its upregulation augments cellular proliferation in ER $\alpha$ <sup>+</sup> breast cancer cells through a positive feedback loop [104–107]. miR-665 was also reported to be downregulated in breast cancer, but a recent study reported its high expression in IDC validated by microarray analysis of patients' samples and qPCR analysis of IDC cell lines (MCF-7, MDA-MB-415, and ZR-75-30). Its upregulation promotes tumor progression through increased proliferation, cell growth and inhibition of apoptosis [108,109]. A previous study identified PARP1 as a target for miR-577 causing its down regulation [110]. In this study, we find PARP1 to be upregulated in IDC in spite of the upregulation of miR-577 (Table 1). Opposite to our findings, miR-577 was previously reported to be downregulated in breast cancer. The low levels of miR-577 were associated with increased invasiveness [111]. The differential expression of miR-577 between IDC and LC could also be attributed to variability in

expression levels in the patients' samples. Interestingly, multiple IDC patients show a remarkable overexpression of the miR-577 when compared to LC patients (Fig 6).

### Implications of overexpression of DNA repair genes on precision medicine for IDC and LC patients

The implementation of precision medicine in cancer treatment has gained a great attention in recent years. Despite the advances in developing anti-cancer drugs and the clinical success and effectiveness of certain drugs, some cancer types do not respond to treatment [112]. As a matter of fact, available therapies are actually limited to some patients with certain tumors, where less than 50% of the patients show responsiveness to around 90% of the drugs. IDC and LC also differ in response to neoadjuvant, with LC being less sensitive [4,113–115]. Nevertheless, similar treatment is usually administered to stage-matched LC versus IDC. Studying the transcriptomic profile of both subtypes is thus essential to identify the genes that contribute to the development, progression and resistance of these distinct subtypes. In addition, tailor better therapy for each subtype.

DNA repair-based targeted therapy for breast cancer induces cell death through impairing DNA repair pathways and increasing the accumulation of DNA damage and breaks. Targeted therapy has less off-target side effects and greater sensitivity than chemotherapy and radiation [116]. Some of our DEGs are known to be very important drug targets for breast cancer patients. For example, PARP inhibitors (PARPi) serve as an example of the targeted therapy [117]. PARPi have shown effectiveness in tumors with deficiencies in our DEGs such as RAD51, FANCD2, FANCA, CHK1 and XRCC2 [118–120]. For BRCA1/2 mutant breast cancers, PARPi olaparib and talazoparib are now FDA-approved monotherapies [119,121,122]. Other studies also showed that their effectiveness extend to tumors without BRCA-mutations [117,123]. A potential therapeutic approach would be combining PARPi with specific inhibitors for our DEGs to achieve better response.

As previously discussed, deciphering the DNA repair expression profile for each patient is essential to understand the cellular consequences of the therapy. However, most studies follow cohort analysis using standard statistical algorithms to determine DEGs, where various normalization methods followed by negative binomial distributions or Poisson are utilized to model the gene count data. Cutoff score based on P-value generated by statistical modeling is then applied along with expression change threshold [124,125]. This method of analysis has been successful in different ways, as they could identify biomarkers and prognostic markers and determine which genes are usually overexpressed or downregulated in certain cancer types [126]. However, the drawback of this approach is that focusing on average expression levels across patient samples is of clinical relevance to many patients; but not all. Therefore, the individuals whose expression levels can be thought of as outliers in the dataset, may not be diagnosed correctly and will not respond to some of therapies that will work for others. Recently, other analysis approaches have evolved to address such limitations where the focus is on the individual level or what is known as 'single-subject analysis (SSA)'. SSA has been performed through utilizing either a specific individual sample versus a cohort of reference samples or a paired sample; tumor and control from the same subject [127]. Rankcomp is an example of the first method, where a ranking methodology and pairwise comparison is performed to ensure stable ranking against normal samples cohort. Then a Fisher's exact test is utilized to detect DEGs [128]. However, Rankcomp can result in high false discovery rate [129]. PenDA is another analysis technique developed to avoid high false discovery [129]. Some of the already developed techniques for cohort analysis have also been evaluated for paired sample analysis such as EdgeR and DESeq [130]. A more advanced approach that aims

at better specificity is single cell analysis via single-cell RNA sequencing (scRNA-seq). Unlike bulk analysis, which reports the predominant malignant clone, scRNA-seq has enabled the identification of the cellular heterogeneity and transcriptomic changes present in tumor micro-environment, either between the patients of same cancer types or even within the same patient [131,132]. These individualized approaches are very effective in determining the appropriate therapy for patients, and ensure the patient's response to treatment and reduce resistance.

## Conclusion

Overall, our data indicate that LC and IDC have different transcriptomic profiles, while they express DNA repair genes in a more similar manner. The cohort analysis performed in this study identifies 36 DNA repair genes to be overexpressed in a significant number of IDC and LC patients, which could significantly affect patients' response to specific therapy. However, the findings can be used as an indicator for the changes in transcription of a specific gene in a certain disease subtype, but not for reaching general conclusions on patients of a specific cancer type. The exact transcriptomic profile of each patient should be taken into consideration to decide for the appropriate therapy and foresee whether the patient will respond to the therapy. We suggest that a great attention should be given to the transcriptomic changes in DNA repair genes, and not only the changes on the genomic level. This should be done through utilizing the appropriate Next generation sequencing diagnostic panels, to reach the most accurate conclusions for individual patients.

## Supporting information

**S1 Fig. The expression of BRCA1 and TP53 in normal\_D, IDC and LC.** Boxplots represent normalized counts in normal\_ductal 'N\_D' (Green), IDC (Orange) and LC (Purple).  
(TIFF)

**S1 Table. DEGs in IDC vs Normal\_Ductal.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to Normal\_ductal, and the negative values reflect the downregulation in IDC in comparison to Normal\_ductal.  
(XLSX)

**S2 Table. DEGs in IDC vs LC.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to LC, and the negative values reflect the downregulation in IDC in comparison to LC.  
(XLSX)

**S3 Table. The list of DNA repair genes analyzed.** The function of each gene was determined according to Uniprot. and the DNA repair pathway the protein is involved in was determined via Reactome.  
(XLSX)

**S4 Table. DNA repair DEGs in IDC vs Normal\_Ductal.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to Normal\_ductal, and the negative values reflect the downregulation in IDC in comparison to Normal\_ductal.  
(XLSX)

**S5 Table. DNA repair DEGs in IDC vs LC.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to LC, and the negative values reflect the downregulation in IDC in comparison to LC.  
(XLS)

**S6 Table. DEMs in IDC vs Normal\_Ductal.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to Normal\_ductal, and the negative values reflect the downregulation in IDC in comparison to Normal\_ductal.

(XLSX)

**S7 Table. DEMs in IDC vs LC.** The positive log<sub>2</sub> fold change reflects overexpression in IDC when compared to LC, and the negative values reflect the downregulation in IDC in comparison to LC.

(XLSX)

**S8 Table. Simple somatic mutations and copy number variations in the TCGA-BRCA project cases.**

(XLSX)

## Acknowledgments

The results shown in this study are in whole based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>. We would like to thank Dr. Tamer Mansour, University of California–Davis, for his technical advice.

## Author Contributions

**Conceptualization:** Menattallah Elserafy.

**Formal analysis:** Ruwaa I. Mohamed, Salma A. Bargal, Asmaa S. Mekawy, Iman El-Shiekh, Eman Badr.

**Funding acquisition:** Menattallah Elserafy.

**Investigation:** Asmaa S. Mekawy, Iman El-Shiekh, Alaa S. Ahmed.

**Methodology:** Ruwaa I. Mohamed, Salma A. Bargal, Nurcan Tuncbag, Eman Badr.

**Project administration:** Menattallah Elserafy.

**Supervision:** Eman Badr, Menattallah Elserafy.

**Visualization:** Ruwaa I. Mohamed, Salma A. Bargal, Nurcan Tuncbag.

**Writing – original draft:** Asmaa S. Mekawy, Iman El-Shiekh, Menattallah Elserafy.

**Writing – review & editing:** Asmaa S. Mekawy, Iman El-Shiekh, Menattallah Elserafy.

## References

1. WHO. Cancer Breast cancer: prevention and control. World Heal Organ. 2014; 11–13. <https://doi.org/10.3989/mc.2010.49909>
2. Registry DC, Registry MC. 99 375 745. Int Agency Res Cancer, WHO. 2019; 399: 2018–2019.
3. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Frontiers in Medicine*. 2017. <https://doi.org/10.3389/fmed.2017.00227> PMID: 29276709
4. Korkola JE, DeVries S, Fridlyand J, Hwang ES, Estep ALH, Chen YY, et al. Differentiation of Lobular versus Ductal Breast Carcinomas by Expression Microarray Analysis. *Cancer Res*. 2003. PMID: 14612510
5. Zhao H, Langerød A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, et al. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell*. 2004. <https://doi.org/10.1091/mbc.e03-11-0786> PMID: 15034139
6. Barroso-Sousa R, Metzger-Filho O. Differences between invasive lobular and invasive ductal carcinoma of the breast: Results and therapeutic implications. *Therapeutic Advances in Medical Oncology*. 2016. <https://doi.org/10.1177/1758834016644156> PMID: 27482285



7. Ciriello G, Gatz ML, Beck AH, Wilkerson MD, Rhee SK, Pastore A, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*. 2015. <https://doi.org/10.1016/j.cell.2015.09.033> PMID: 26451490
8. Desmedt C, Zoppoli G, Gundem G, Pruneri G, Larsimont D, Fornili M, et al. Genomic Characterization of Primary Invasive Lobular Breast Cancer. *Journal of Clinical Oncology*. 2016. <https://doi.org/10.1200/JCO.2015.64.0334> PMID: 26926684
9. Urquidí V, Sloan D, Kawai K, Agarwal D, Tarin D, Goodison S, et al. Contrasting expression of thrombospondin-1 and osteopontin correlates with absence or presence of metastatic phenotype in an isogenic model of spontaneous human breast cancer metastasis. *Clin Cancer Res*. 2002;8. PMID: 11801541
10. Bex G, Van Roy F. The E-cadherin/catenin complex: An important gatekeeper in breast cancer tumorigenesis and malignant progression. *Breast Cancer Res*. 2001; 3. <https://doi.org/10.1186/bcr309> PMID: 11597316
11. Geradts J, Ingram CD. Abnormal expression of cell cycle regulatory proteins ductal and lobular carcinomas of the breast. *Mod Pathol*. 2000. <https://doi.org/10.1038/modpathol.3880172> PMID: 11007034
12. Soslow RA, Carlson DL, Horenstein MG, Osborne MP. A comparison of cell cycle markers in well-differentiated lobular and ductal carcinomas. *Breast Cancer Res Treat*. 2000. <https://doi.org/10.1023/a:1006479113769> PMID: 10942102
13. Coradini D, Pellizzaro C, Veneroni S, Ventura L, Daidone MG. Infiltrating ductal and lobular breast carcinomas are characterised by different interrelationships among markers related to angiogenesis and hormone dependence. *Br J Cancer*. 2002. <https://doi.org/10.1038/sj.bjc.6600556> PMID: 12402149
14. Abugable AA, Awwad DA, Fleifel D, Ali MM, El-Khamisy S, Elserafy M. Personalised medicine: Genome maintenance lessons learned from studies in yeast as a model organism. *Advances in Experimental Medicine and Biology*. 2017. [https://doi.org/10.1007/978-3-319-60733-7\\_9](https://doi.org/10.1007/978-3-319-60733-7_9) PMID: 28840557
15. Matta J, Echenique M, Negron E, Morales L, Vargas W, Gaetan FS, et al. The association of DNA Repair with breast cancer risk in women. A comparative observational study. *BMC Cancer*. 2012; 12. <https://doi.org/10.1186/1471-2407-12-490> PMID: 23088658
16. Turner N, Tutt A, Ashworth A. Hallmarks of “BRCAness” in sporadic cancers. *Nature Reviews Cancer*. Nature Publishing Group; 2004. pp. 814–819. <https://doi.org/10.1038/nrc1457> PMID: 15510162
17. Roberts ME, Jackson SA, Susswein LR, Zeinomar N, Ma X, Marshall ML, et al. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. *Genet Med*. 2018; 20: 1167–1174. <https://doi.org/10.1038/gim.2017.254> PMID: 29345684
18. Davies H, Morganella S, Purdie CA, Jang SJ, Borgen E, Russnes H, et al. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. *Cancer Res*. 2017; 77: 4755–4762. <https://doi.org/10.1158/0008-5472.CAN-17-1083> PMID: 28904067
19. Cuchra M, Mucha B, Markiewicz L, Przybyłowska-Sygut K, Pytel D, Jeziorski A, et al. The role of base excision repair in pathogenesis of breast cancer in the Polish population. *Mol Carcinog*. 2016. <https://doi.org/10.1002/mc.22436> PMID: 27870262
20. Patrono C, Sterpone S, Testa A, Cozzi R. Polymorphisms in base excision repair genes: Breast cancer risk and individual radiosensitivity. *World J Clin Oncol*. 2014. <https://doi.org/10.5306/wjco.v5.i5.874> PMID: 25493225
21. Latimer JJ, Johnson JM, Kelly CM, Miles TD, Beaudry-Rodgers KA, Lallane NA, et al. Nucleotide excision repair deficiency is intrinsic in sporadic stage I breast cancer. *Proc Natl Acad Sci U S A*. 2010; 107: 21725–21730. <https://doi.org/10.1073/pnas.0914772107> PMID: 21118987
22. Mechanic LE, Millikan RC, Player J, de Cotret AR, Winkel S, Worley K, et al. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: A population-based case-control study. *Carcinogenesis*. 2006; 27: 1377–1385. <https://doi.org/10.1093/carcin/bgi330> PMID: 16399771
23. Swift M, Morrell D, Massey RB, Chase CL. Incidence of cancer in 161 families affected by Ataxia–Telangiectasia. *N Engl J Med*. 1991; 325: 1831–1836. <https://doi.org/10.1056/NEJM199112263252602> PMID: 1961222
24. Cavaciuti E, Laugé A, Janin N, Ossian K, Hall J, Stoppa-Lyonnet D, et al. Cancer risk according to type and location of *ATM* mutation in ataxia-telangiectasia families. *Genes, Chromosom Cancer*. 2005; 42: 1–9. <https://doi.org/10.1002/gcc.20101> PMID: 15390180
25. Olsen JH, Hahnemann JM, Børresen-Dale AL, Brøndum-Nielsen K, Hammarstrom L, Kleinerman R, et al. Cancer in patients with ataxia-telangiectasia and in their relatives in the Nordic countries. *J Natl Cancer Inst*. 2001; 93: 121–127. <https://doi.org/10.1093/jnci/93.2.121> PMID: 11208881
26. R Core Team (2020). R: A language and environment for statistical computing. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020.

27. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020. <https://doi.org/10.1093/nar/gkz1031> PMID: 31691815
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
29. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt. *Nat Protoc.* 2009; 4. <https://doi.org/10.1038/nprot.2009.97> PMID: 19617889
30. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife.* 2015. <https://doi.org/10.7554/eLife.05005> PMID: 26267216
31. Huang HY, Lin YCD, Li J, Huang KY, Shrestha S, Hong HC, et al. MiRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 2020. <https://doi.org/10.1093/nar/gkz896> PMID: 31647101
32. Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* 2019. <https://doi.org/10.1093/nar/gky1079> PMID: 30476227
33. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13: 2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
34. Proceedings of the Python in Science Conference (SciPy): Exploring Network Structure, Dynamics, and Function using NetworkX. [cited 9 Oct 2020]. Available: [http://conference.scipy.org/proceedings/SciPy2008/paper\\_2/](http://conference.scipy.org/proceedings/SciPy2008/paper_2/).
35. Bertucci F, Orsetti B, Nègre V, Finetti P, Rougé C, Ahomadegbe JC, et al. Lobular and ductal carcinomas of the breast have distinct genomic and expression profiles. *Oncogene.* 2008. <https://doi.org/10.1038/onc.2008.158> PMID: 18490921
36. Tachibana C. Transcriptomics today: Microarrays, RNA-seq, and more. *Science (80-).* 2015. <https://doi.org/10.1126/science.349.6247.544>
37. Davis JD. DNA damage and breast cancer. *World J Clin Oncol.* 2011. <https://doi.org/10.5306/wjco.v2.i9.329> PMID: 21909479
38. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology.* 2018. <https://doi.org/10.3389/fendo.2018.00402> PMID: 30123182
39. Vasudevan S. Posttranscriptional Upregulation by MicroRNAs. *Wiley Interdisciplinary Reviews: RNA.* John Wiley & Sons, Ltd; 2012. pp. 311–330. <https://doi.org/10.1002/wrna.121> PMID: 22072587
40. Pharaoh PDP, Day NE, Caldas C. Somatic mutations in the p53 gene and prognosis in breast cancer: A meta-analysis. *Br J Cancer.* 1999; 80: 1968–1973. <https://doi.org/10.1038/sj.bjc.6690628> PMID: 10471047
41. Smith PD, Crossland S, Parker G, Osin P, Brooks L, Waller J, et al. Novel p53 mutants selected in BRCA-associated tumours which dissociate transformation suppression from other wild-type p53 functions. *Oncogene.* 1999; 18: 2451–2459. <https://doi.org/10.1038/sj.onc.1202565> PMID: 10229196
42. Phillips KA, Nichol K, Ozcelik H, Knight J, Done SJ, Goodwin PJ, et al. Frequency of p53 mutations in breast carcinomas from Ashkenazi Jewish carriers of BRCA1 mutations. *J Natl Cancer Inst.* 1999. <https://doi.org/10.1093/jnci/91.5.469> PMID: 10070948
43. Huszno J, Kołosza Z, Grzybowska EWA. BRCA1 mutation in breast cancer patients: Analysis of prognostic factors and survival. *Oncol Lett.* 2019; 17: 1986–1995. <https://doi.org/10.3892/ol.2018.9770> PMID: 30675265
44. Matta J, Morales L, Dutil J, Bayona M, Alvarez C, Suarez E. Differential expression of DNA repair genes in Hispanic women with breast cancer. *Mol cancer Biol.* 2013; 1: 54. <https://doi.org/10.9777/mcb.2013.10006> PMID: 25309843
45. Wang X, Jung YS, Jun S, Lee S, Wang W, Schneider A, et al. PAF-Wnt signaling-induced cell plasticity is required for maintenance of breast cancer cell stemness. *Nat Commun.* 2016. <https://doi.org/10.1038/ncomms10633> PMID: 26843124
46. Bektas N, Noetzel E, Veeck J, Press MF, Kristiansen G, Naami A, et al. The ubiquitin-like molecule interferon-stimulated gene 15 (ISG15) is a potential prognostic marker in human breast cancer. *Breast Cancer Res.* 2008; 10. <https://doi.org/10.1186/bcr2117> PMID: 18627608
47. Ueki T, Nishidate T, Park JH, Lin ML, Shimo A, Hirata K, et al. Involvement of elevated expression of multiple cell-cycle regulator, DTL/RAMP (denticleless/RA-regulated nuclear matrix associated protein), in the growth of breast cancer cells. *Oncogene.* 2008; 27: 5672–5683. <https://doi.org/10.1038/onc.2008.186> PMID: 18542055

48. Mackintosh C, Ordóñez JL, García-Domínguez DJ, Sevillano V, Llombart-Bosch A, Suzhai K, et al. 1q gain and CDT2 overexpression underlie an aggressive and highly proliferative form of Ewing sarcoma. *Oncogene*. 2012; 31: 1287–1298. <https://doi.org/10.1038/onc.2011.317> PMID: 21822310
49. Singh P, Yang M, Dai H, Yu D, Huang Q, Tan W, et al. Overexpression and hypomethylation of Flap endonuclease 1 gene in breast and other cancers. *Mol Cancer Res*. 2008. <https://doi.org/10.1158/1541-7786.MCR-08-0269> PMID: 19010819
50. Zou J, Zhu L, Jiang X, Wang Y, Wang Y, Wang X, et al. Curcumin increases breast cancer cell sensitivity to cisplatin by decreasing FEN1 expression. *Oncotarget*. 2018. <https://doi.org/10.18632/oncotarget.24109> PMID: 29541412
51. Wang J, Zhou L, Li Z, Zhang T, Liu W, Liu Z, et al. YY1 suppresses FEN1 over-expression and drug resistance in breast cancer. *BMC Cancer*. 2015. <https://doi.org/10.1186/s12885-015-1043-1> PMID: 25885449
52. Guo E, Ishii Y, Mueller J, Srivatsan A, Gahman T, Putnam CD, et al. FEN1 endonuclease as a therapeutic target for human cancers with defects in homologous recombination. *Proc Natl Acad Sci U S A*. 2020. <https://doi.org/10.1073/pnas.2009237117> PMID: 32719125
53. He L, Zhang Y, Sun H, Jiang F, Yang H, Wu H, et al. Targeting DNA Flap Endonuclease 1 to Impede Breast Cancer Progression. *EBioMedicine*. 2016. <https://doi.org/10.1016/j.ebiom.2016.11.012> PMID: 27852524
54. Hildrestrand GA, Neurauter CG, Diep DB, Castellanos CG, Krauss S, Bjørås M, et al. Expression patterns of Neil3 during embryonic brain development and neoplasia. *BMC Neurosci*. 2009. <https://doi.org/10.1186/1471-2202-10-45> PMID: 19426544
55. Tran OTN, Tadesse S, Chu C, Kidane D. Overexpression of NEIL3 associated with altered genome and poor survival in selected types of human cancer. *Tumor Biol*. 2020. <https://doi.org/10.1177/1010428320918404> PMID: 32364878
56. Shinmura K, Kato H, Kawanishi Y, Igarashi H, Goto M, Tao H, et al. Abnormal expressions of DNA glycosylase genes NEIL1, NEIL2, and NEIL3 are associated with somatic mutation loads in human cancer. *Oxid Med Cell Longev*. 2016. <https://doi.org/10.1155/2016/1546392> PMID: 27042257
57. Dankof A, Fritzsche FR, Dahl E, Pahl S, Wild P, Dietel M, et al. KPNA2 protein expression in invasive breast carcinoma and matched peritumoral ductal carcinoma in situ. *Virchows Arch*. 2007; 451: 877–881. <https://doi.org/10.1007/s00428-007-0513-5> PMID: 17899179
58. Gluz O, Wild P, Meiler R, Diallo-Danebrock R, Ting E, Mohrmann S, et al. Nuclear karyopherin  $\alpha 2$  expression predicts poor survival in patients with advanced breast cancer irrespective of treatment intensity. *Int J Cancer*. 2008; 123: 1433–1438. <https://doi.org/10.1002/ijc.23628> PMID: 18561322
59. Dahl E, Kristiansen G, Gottlob K, Klamann I, Ebner E, Hinzmann B, et al. Molecular profiling of laser-microdissected matched tumor and normal breast tissue identifies karyopherin  $\alpha 2$  as a potential novel prognostic marker in breast cancer. *Clin Cancer Res*. 2006. <https://doi.org/10.1158/1078-0432.CCR-05-2090> PMID: 16818692
60. Cui X, Jing X, Wu X, Xu J, Liu Z, Huo K, et al. Analyses of DNA Methylation Involved in the Activation of Nuclear Karyopherin Alpha 2 Leading to Identify the Progression and Prognostic Significance Across Human Breast Cancer. 2020. <https://doi.org/10.2147/CMAR.S261290> PMID: 32801900
61. Song K-H, Jung S-Y, Park J-I, Ahn J, Park JK, Um H-D, et al. Molecular Sciences Inhibition of Karyopherin- $\alpha 2$  Augments Radiation-Induced Cell Death by Perturbing BRCA1-Mediated DNA Repair. 2019. <https://doi.org/10.3390/ijms20112843> PMID: 31212646
62. Ueki T, Park JH, Nishidate T, Kijima K, Hirata K, Nakamura Y, et al. Ubiquitination and downregulation of BRCA1 by ubiquitin-conjugating enzyme E2T overexpression in human breast cancer cells. *Cancer Res*. 2009. <https://doi.org/10.1158/0008-5472.CAN-09-1809> PMID: 19887602
63. Fang C Bin, Wu HT, Zhang ML, Liu J, Zhang GJ. Fanconi Anemia Pathway: Mechanisms of Breast Cancer Predisposition Development and Potential Therapeutic Targets. *Frontiers in Cell and Developmental Biology*. 2020. <https://doi.org/10.3389/fcell.2020.00160> PMID: 32300589
64. Singh TR, Saro D, Ali AM, Zheng XF, Du C hu, Killen MW, et al. MHF1-MHF2, a Histone-Fold-Containing Protein Complex, Participates in the Fanconi Anemia Pathway via FANCM. *Mol Cell*. 2010; 37: 879–886. <https://doi.org/10.1016/j.molcel.2010.01.036> PMID: 20347429
65. Yan Z, Delannoy M, Ling C, Dae D, Osman F, Muniandy PA, et al. A Histone-Fold Complex and FANCM Form a Conserved DNA-Remodeling Complex to Maintain Genome Stability. *Mol Cell*. 2010; 37: 865–878. <https://doi.org/10.1016/j.molcel.2010.01.039> PMID: 20347428
66. Van Der Groep P, Hoelzel M, Buerger H, Joenje H, De Winter JP, Van Diest PJ. Loss of expression of FANCD2 protein in sporadic and hereditary breast cancer. *Breast Cancer Res Treat*. 2008; 107: 41–47. <https://doi.org/10.1007/s10549-007-9534-7> PMID: 17333336

67. Akashi-Tanaka S, Watanabe C, Takamaru T, Kuwayama T, Ikeda M, Ohyama H, et al. BRCAness predicts resistance to taxane-containing regimens in triple negative breast cancer during neoadjuvant chemotherapy. *Clin Breast Cancer*. 2015; 15: 80–85. <https://doi.org/10.1016/j.clbc.2014.08.003> PMID: 25445419
68. Chudasama D, Bo V, Hall M, Anikin V, Jeyaneethi J, Gregory J, et al. Identification of cancer biomarkers of prognostic value using specific gene regulatory networks (GRN): A novel role of RAD51AP1 for ovarian and lung cancers. *Carcinogenesis*. 2018; 39: 407–417. <https://doi.org/10.1093/carcin/bgx122> PMID: 29126163
69. Bridges A, Ramachandran S, Pathania R, Parwal U, Lester A, Rajpurohit P, et al. RAD51AP1 deficiency reduces tumor growth by targeting stem cell self-renewal. *Cancer Res*. 2020; 80: canres.3713.2019. <https://doi.org/10.1158/0008-5472.CAN-19-3713> PMID: 32665355
70. Kluzniak W, Wokolorczyk D, Rusak B, Huzarski T, Gronwald J, Stempa K, et al. Inherited variants in XRCC2 and the risk of breast cancer. *Breast Cancer Res Treat*. 2019; 178: 657–663. <https://doi.org/10.1007/s10549-019-05415-5> PMID: 31463769
71. Bashir N, Sana S, Mahjabeen I, Kayani MA. Association of reduced XRCC2 expression with lymph node metastasis in breast cancer tissues. *Fam Cancer*. 2014; 13: 611–617. <https://doi.org/10.1007/s10689-014-9745-0> PMID: 25159888
72. Wu M, Pang J-S, Sun Q, Huang Y, Hou J-Y, Chen G, et al. The clinical significance of CHEK1 in breast cancer: a high-throughput data analysis and immunohistochemical study. *Int J Clin Exp Pathol*. 2019;12. PMID: 31933717
73. Al-Kaabi MM, Alshareeda AT, Jerjees DA, Muftah AA, Green AR, Alsubhi NH, et al. Checkpoint kinase1 (CHK1) is an important biomarker in breast cancer having a role in chemotherapy response. *Br J Cancer*. 2015. <https://doi.org/10.1038/bjc.2014.576> PMID: 25688741
74. Thompson R, Eastman A. The cancer therapeutic potential of Chk1 inhibitors: How mechanistic studies impact on clinical trial design. *Br J Clin Pharmacol*. 2013. <https://doi.org/10.1111/bcp.12139> PMID: 23593991
75. Zhou ZR, Yang ZZ, Wang SJ, Zhang L, Luo JR, Feng Y, et al. The Chk1 inhibitor MK-8776 increases the radiosensitivity of human triple-negative breast cancer by inhibiting autophagy. *Acta Pharmacol Sin*. 2017. <https://doi.org/10.1038/aps.2016.136> PMID: 28042876
76. Verlinden L, Bempt I, Vanden Eelen G, Drijkoningen M, Verlinden I, Marchal K, et al. The E2F-regulated Gene Chk1 is highly expressed in triple-negative estrogen receptor-/progesterone receptor-/HER-2- breast carcinomas. *Cancer Res*. 2007; 67: 6574–6581. <https://doi.org/10.1158/0008-5472.CAN-06-3545> PMID: 17638866
77. Choi SH, Yang H, Lee SH, Ki JH, Nam DH, Yoo HY. TopBP1 and Claspin contribute to the radioresistance of lung cancer brain metastases. *Mol Cancer*. 2014; 13: 1–8. <https://doi.org/10.1186/1476-4598-13-1> PMID: 24387052
78. Chi L, Zou Y, Qin L, Ma W, Hao Y, Tang Y, et al. TIMELESS contributes to the progression of breast cancer through activation of MYC. *Breast Cancer Res*. 2017; 19: 53. <https://doi.org/10.1186/s13058-017-0838-1> PMID: 28464854
79. Tozlu-Kara S, Roux V, Andrieu C, Vendrell J, Vacher S, Lazar V, et al. Oligonucleotide microarray analysis of estrogen receptor  $\alpha$ -positive postmenopausal breast carcinomas: Identification of HRPAP20 and TIMELESS as outstanding candidate markers to predict the response to tamoxifen. *J Mol Endocrinol*. 2007; 39: 305–318. <https://doi.org/10.1677/JME-07-0001> PMID: 17909269
80. Wang B, Zhao A, Sun L, Zhong X, Zhong J, Wang H, et al. Protein phosphatase PP4 is overexpressed in human breast and lung tumors. *Cell Research*. Nature Publishing Group; 2008. pp. 974–977. <https://doi.org/10.1038/cr.2008.274> PMID: 19160544
81. Maacke H, Opitz S, Jost K, Hamdorf W, Henning W, Krüger S, et al. Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. *Int J Cancer*. 2000; 88: 907–913. [https://doi.org/10.1002/1097-0215\(20001215\)88:6<907::aid-ijc11>3.0.co;2-4](https://doi.org/10.1002/1097-0215(20001215)88:6<907::aid-ijc11>3.0.co;2-4) PMID: 11093813
82. Rafnar T, Gudbjartsson DF, Sulem P, Jonasdottir A, Sigurdsson A, Jonasdottir A, et al. Mutations in BRIP1 confer high risk of ovarian cancer. *Nat Genet*. 2011; 43: 1104–1107. <https://doi.org/10.1038/ng.955> PMID: 21964575
83. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J Clin Oncol*. 2007. <https://doi.org/10.1200/JCO.2006.09.1066> PMID: 17416853
84. Gupta I, Ouhtit A, Al-Ajmi A, Rizvi SGA, Al-Riyami H, Al-Riyami M, et al. BRIP1 overexpression is correlated with clinical features and survival outcome of luminal breast cancer subtypes. *Endocr Connect*. 2018; 7: 65–77. <https://doi.org/10.1530/EC-17-0173> PMID: 29138235

85. Zhan W, Liu Y, Gao Y, Gong R, Wang W, Zhang R, et al. RMI2 plays crucial roles in growth and metastasis of lung cancer. *Signal Transduction and Targeted Therapy*. Springer Nature; 2020. p. 188. <https://doi.org/10.1038/s41392-020-00295-4> PMID: 32883952
86. Broberg K, Huynh E, Engström KS, Björk J, Albin M, Ingvar C, et al. Association between polymorphisms in RMI1, TOP3A, and BLM and risk of cancer, a case-control study. *BMC Cancer*. 2009. <https://doi.org/10.1186/1471-2407-9-140> PMID: 19432957
87. Arora A, Abdel-Fatah TMA, Agarwal D, Doherty R, Moseley PM, Aleskandarany MA, et al. Transcriptomic and protein expression analysis reveals clinicopathological significance of bloom syndrome heliase (BLM) in breast cancer. *Mol Cancer Ther*. 2015; 14: 1057–1065. <https://doi.org/10.1158/1535-7163.MCT-14-0939> PMID: 25673821
88. Peng G, Dai H, Zhang W, Hsieh HJ, Pan MR, Park YY, et al. Human nuclease/helicase DNA2 alleviates replication stress by promoting DNA end resection. *Cancer Res*. 2012; 72: 2802–2813. <https://doi.org/10.1158/0008-5472.CAN-11-3152> PMID: 22491672
89. Gao T, Han Y, Yu L, Ao S, Li Z, Ji J. CCNA2 Is a Prognostic Biomarker for ER+ Breast Cancer and Tamoxifen Resistance. St-Pierre Y, editor. *PLoS One*. 2014; 9: e91771. <https://doi.org/10.1371/journal.pone.0091771> PMID: 24622579
90. Castellano I, Chiusa L, . . . AV-A of, 2013 undefined. A simple and reproducible prognostic index in luminal ER-positive breast cancers. Oxford Univ Press.
91. Zhao J, Liu L, Zhang A, Chen Q, Fang W, Zeng L, et al. Effect of EME1 exon variant Ile350Thr on risk and early onset of breast cancer in southern Chinese women. *J Biomed Res*. 2013; 27: 27. <https://doi.org/10.7555/JBR.27.20130013> PMID: 23720674
92. Tomoda Y, Katsura M, Okajima M, Hosoya N, Kohno N, Miyagawa K. Functional evidence for Eme1 as a marker of cisplatin resistance. *Int J Cancer*. 2009. <https://doi.org/10.1002/ijc.24268> PMID: 19267403
93. Post AEM, Bussink J, Sweep FCGJ, Span PN. Changes in DNA Damage Repair Gene Expression and Cell Cycle Gene Expression Do Not Explain Radioresistance in Tamoxifen-Resistant Breast Cancer. *Oncol Res*. 28: 33–40. <https://doi.org/10.3727/096504019X15555794826018> PMID: 31046897
94. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet*. 2015; 47: 373–380. <https://doi.org/10.1038/ng.3242> PMID: 25751625
95. Wang HC, Chiu CF, Tsai RY, Kuo YS, Chen HS, Wang RF, et al. Association of genetic polymorphisms of EXO1 gene with risk of breast cancer in Taiwan. *Anticancer Res*. 2009. PMID: 19846925
96. Erdal E, Haider S, Rehwinkel J, Harris AL, Mchugh PJ. A prosurvival DNA damage-induced cytoplasmic interferon response is mediated by end resection factors and is limited by Trex1. 2017. <https://doi.org/10.1101/gad.289769.116> PMID: 28279982
97. Bechtel PE, Hickey RJ, Schnaper L, Sekowski JW, Long BJ, Freund R, et al. A Unique Form of Proliferating Cell Nuclear Antigen Is Present in Malignant Breast Cells1. *CANCER Res*. 1998.
98. Schönborn I, Minguillon C, Möhner M, Ebeling K. PCNA as a potential prognostic marker in breast cancer. *The Breast*. 1994; 3: 97–102. [https://doi.org/10.1016/0960-9776\(94\)90007-8](https://doi.org/10.1016/0960-9776(94)90007-8)
99. Zhou Q, Effati R, Talvinen K, Pospiech H, Syväoja JE, Collan Y. Genomic changes of the 55 kDa subunit of DNA polymerase  $\epsilon$  in human breast cancer. *Cancer Genomics and Proteomics*. 2008. PMID: 19129559
100. Bao S, Chen Y, Yang F, Sun C, Yang M, Li W, et al. Screening and Identification of Key Biomarkers in Acquired Lapatinib-Resistant Breast Cancer. *Front Pharmacol*. 2020; 11: 577150. <https://doi.org/10.3389/fphar.2020.577150> PMID: 33013420
101. Xie W, Zhang J, Zhong P, Qin S, Zhang H, Fan X, et al. Expression and potential prognostic value of histone family gene signature in breast cancer. *Exp Ther Med*. 2019. <https://doi.org/10.3892/etm.2019.8131> PMID: 31772649
102. Campone M, Campion L, Roché H, Gouraud W, Charbonnel C, Magrangers F, et al. Prediction of metastatic relapse in node-positive breast cancer: Establishment of a clinicogenomic model after FEC100 adjuvant regimen. *Breast Cancer Res Treat*. 2008. <https://doi.org/10.1007/s10549-007-9673-x> PMID: 17659439
103. Siraj AK, Pratheeshkumar P, Parvathareddy SK, Divya SP, Al-Dayel F, Tulbah A, et al. Overexpression of PARP is an independent prognostic marker for poor survival in Middle Eastern breast cancer and its inhibition can be enhanced with embelin co-treatment. *Oncotarget*. 2018. <https://doi.org/10.18632/oncotarget.26470> PMID: 30647872
104. Hong S, Noh H, Teng Y, Shao J, Rehmani H, Ding HF, et al. SHOX2 is a direct miR-375 target and a novel epithelial-to-mesenchymal transition inducer in breast cancer cells. *Neoplasia (United States)*. 2014. <https://doi.org/10.1016/j.neo.2014.03.010> PMID: 24746361

105. Luo D, Wilson JM, Harvel N, Liu J, Pei L, Huang S, et al. A systematic evaluation of miRNA: MRNA interactions involved in the migration and invasion of breast cancer cells. *J Transl Med*. 2013. <https://doi.org/10.1186/1479-5876-11-57> PMID: 23497265
106. Munagala R, Aqil F, Vadhanam M V., Gupta RC. MicroRNA “signature” during estrogen-mediated mammary carcinogenesis and its reversal by ellagic acid intervention. *Cancer Lett*. 2013. <https://doi.org/10.1016/j.canlet.2013.06.012> PMID: 23791885
107. Rocha Simonini PDS, Breiling A, Gupta N, Malekpour M, Youns M, Omranipour R, et al. Epigenetically deregulated microRNA-375 is involved in a positive feedback loop with estrogen receptor  $\alpha$  in breast cancer cells. *Cancer Res*. 2010. <https://doi.org/10.1158/0008-5472.CAN-10-1318> PMID: 20978187
108. Si H, Sun X, Chen Y, Cao Y, Chen S, Wang H, et al. Circulating microRNA-92a and microRNA-21 as novel minimally invasive biomarkers for primary breast cancer. *J Cancer Res Clin Oncol*. 2013. <https://doi.org/10.1007/s00432-012-1315-y> PMID: 23052693
109. Zhao XG, Hu JY, Tang J, Yi W, Zhang MY, Deng R, et al. miR-665 expression predicts poor survival and promotes tumor metastasis by targeting NR4A3 in breast cancer. *Cell Death Dis*. 2019. <https://doi.org/10.1038/s41419-019-1705-z> PMID: 31209222
110. Wang YH, Zhang XY, Han YQ, Yan F, Wu R. MicroRNA-577 inhibits cardiomyocyte apoptosis induced by myocardial infarction via targeting PARP1. *Eur Rev Med Pharmacol Sci*. 2019. [https://doi.org/10.26355/eurrev\\_201911\\_19451](https://doi.org/10.26355/eurrev_201911_19451) PMID: 31773708
111. Yin C, Mou Q, Pan X, Zhang G, Li H, Sun Y. MiR-577 suppresses epithelial-mesenchymal transition and metastasis of breast cancer by targeting Rab25. *Thorac Cancer*. 2018. <https://doi.org/10.1111/1759-7714.12612> PMID: 29524309
112. Goswami S, Walle T, Cornish AE, Basu S, Anandhan S, Fernandez I, et al. Immune profiling of human tumors identifies CD73 as a combinatorial target in glioblastoma. *Nature Medicine*. 2020. <https://doi.org/10.1038/s41591-019-0694-x> PMID: 31873309
113. Jain S, Fisher C, Smith P, Millis RR, Rubens RD. Patterns of metastatic breast cancer in relation to histological type. *Eur J Cancer*. 1993. [https://doi.org/10.1016/0959-8049\(93\)90053-i](https://doi.org/10.1016/0959-8049(93)90053-i) PMID: 8297656
114. Borst MJ, Ingold JA. Metastatic patterns of invasive lobular versus invasive ductal carcinoma of the breast. *Surgery*. 1993. <https://doi.org/10.5555/uri:pii:003960609390249D> PMID: 8211676
115. Newman LA. A Prospective Trial of Preoperative Chemotherapy in Resectable Breast Cancer: Predictors of Breast-Conservation Therapy Feasibility. *Ann Surg Oncol*. 2002. <https://doi.org/10.1007/BF02573059> PMID: 11923128
116. Rose M, Burgess JT, O’Byrne K, Richard DJ, Bolderson E. PARP Inhibitors: Clinical Relevance, Mechanisms of Action and Tumor Resistance. *Front Cell Dev Biol*. 2020. <https://doi.org/10.3389/fcell.2020.564601> PMID: 33015058
117. McCann KE, Hurvitz SA. Advances in the use of PARP inhibitor therapy for breast cancer. *Drugs in Context*. 2018. <https://doi.org/10.7573/dic.212540> PMID: 30116283
118. Mani C, Jonnalagadda S, Lingareddy J, Awasthi S, Gmeiner WH, Palle K. Prexasertib treatment induces homologous recombination deficiency and synergizes with olaparib in triple-negative breast cancer cells. *Breast Cancer Res*. 2019. <https://doi.org/10.1186/s13058-019-1192-2> PMID: 31492187
119. Gilmore E, McCabe N, Kennedy RD, Parkes EE. DNA Repair Deficiency in Breast Cancer: Opportunities for Immunotherapy. *J Oncol*. 2019. <https://doi.org/10.1155/2019/4325105> PMID: 31320901
120. Pommier Y, Huang SH, Das BB, Renaud A, Zhang Y, Takeda SH, et al. 284 Differential Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Eur J Cancer*. 2012; 48: 87. [https://doi.org/10.1016/s0959-8049\(12\)72082-8](https://doi.org/10.1016/s0959-8049(12)72082-8)
121. Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N Engl J Med*. 2017. <https://doi.org/10.1056/NEJMoa1706450> PMID: 28578601
122. Litton JK, Rugo HS, Ettl J, Hurvitz SA, Gonçalves A, Lee KH, et al. Talazoparib in patients with advanced breast cancer and a germline BRCA mutation. *N Engl J Med*. 2018. <https://doi.org/10.1056/NEJMoa1802905> PMID: 30110579
123. Keung MY, Wu Y, Badar F, Vadgama J V. Response of Breast Cancer Cells to PARP Inhibitors Is Independent of BRCA Status. *J Clin Med*. 2020. <https://doi.org/10.3390/jcm9040940> PMID: 32235451
124. Lin Y, Golovnina K, Chen ZX, Lee HN, Negron YLS, Sultana H, et al. Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual *Drosophila melanogaster*. *BMC Genomics*. 2016; 17. <https://doi.org/10.1186/s12864-015-2353-z> PMID: 26732976
125. Li X, Cooper NGF, O’Toole TE, Rouchka EC. Choice of library size normalization and statistical methods for differential gene expression analysis in balanced two-group comparisons for RNA-seq studies. *BMC Genomics*. 2020; 21. <https://doi.org/10.1186/s12864-020-6502-7> PMID: 31992223

126. Kutomi G, Mizuguchi T, Satomi F, Maeda H, Shima H, Kimura Y, et al. Current status of the prognostic molecular biomarkers in breast cancer: A systematic review. *Oncology Letters*. Spandidos Publications; 2017. pp. 1491–1498. <https://doi.org/10.3892/ol.2017.5609> PMID: 28454281
127. Vitali F, Li Q, Schissler AG, Berghout J, Kenost C, Lussier YA. Developing a “personalome” for precision medicine: Emerging methods that compute interpretable effect sizes from single-subject transcriptomes. *Brief Bioinform*. 2017. <https://doi.org/10.1093/bib/bbx149> PMID: 29272327
128. Wang H, Sun Q, Zhao W, Qi L, Gu Y, Li P, et al. Individual-level analysis of differential expression of genes and pathways for personalized medicine. *Bioinformatics*. 2015; 31. <https://doi.org/10.1093/bioinformatics/btu522> PMID: 25165092
129. Richard M, Decamps C, Chuffart F, Brambilla E, Rousseaux S, Khochbin S, et al. PenDA, a rank-based method for personalized differential analysis: Application to lung cancer. *PLoS Comput Biol*. 2020; 16. <https://doi.org/10.1371/journal.pcbi.1007869> PMID: 32392248
130. Rachid Zaim S, Kenost C, Berghout J, Vitali F, Zhang HH, Lussier YA. Evaluating single-subject study methods for personal transcriptomic interpretations to advance precision medicine. *BMC Med Genomics*. 2019;12. <https://doi.org/10.1186/s12920-018-0455-6> PMID: 30704475
131. Benson M. Clinical implications of omics and systems medicine: Focus on predictive and individualized treatment. *J Intern Med*. 2016. <https://doi.org/10.1111/joim.12412> PMID: 26891944
132. Gustafsson M, Gawel DR, Alfredsson L, Baranzini S, Bjorkander J, Blomgran R, et al. A validated gene regulatory network and GWAS identifies early regulators of T cell-associated diseases. *Sci Transl Med*. 2015. <https://doi.org/10.1126/scitranslmed.aad2722> PMID: 26560356