

## Genetic Profiling of Breast Cancer with and Without Preexisting Metabolic Disease



Wenjiang Jing<sup>\*,1</sup>, Ling Li<sup>†,1</sup>, Xiumei Zhang<sup>‡,1</sup>, Shouxin Wu<sup>§</sup>, Jiangman Zhao<sup>§</sup>, Qunxing Hou<sup>§</sup>, Haotian Wu<sup>§</sup>, Wu Ma<sup>\*</sup>, Shuheng Li<sup>¶</sup>, Huimin Liu<sup>§</sup> and Binhui Yang<sup>#</sup>

\*Department of Oncology, The Aviation Hanzhong 3201 Hospital, Xi'an Jiao Tong University, Hanzhong 723000, Shaanxi, China; <sup>†</sup>Department of Oncology, Tengzhou Central People's Hospital, Jining Medical University, Tengzhou 277500, Shandong, China; <sup>‡</sup>Department of Pathology, People's Hospital of Xinghua City, Xinghua 225700, Jiangsu, China; <sup>§</sup>Biotecon Medical Diagnostics Co., Ltd., Zhangjiang Center for Translational Medicine, Shanghai 201203, China; <sup>¶</sup>Department of Thyroid and Breast Surgery, Tengzhou Central People's Hospital, Jining Medical University, Tengzhou 277500, Shandong, China; <sup>#</sup>Department of Orthopaedic, The Aviation Hanzhong 3201 Hospital, Xi'an Jiao Tong University, Hanzhong 723000, Shaanxi, China

### Abstract

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death among women. Various mechanisms are involved in the initiation and progression of breast cancer. Metabolic dysregulation has been associated with increasing breast cancer incidence and mortality. However, little is known about how metabolic disease regulates the development and progression of breast cancer at the molecular level. Here, using a hybridization capture-based panel including 124 cancer-associated genes, we performed targeted next-generation sequencing of tumor tissues and matched blood samples from 20 postmenopausal patients with primary breast cancer, in which 6 cases suffered from preexisting metabolic disorders including hypertension, type 2 diabetes, and coronary heart disease. We took only the protein-altering variants and identified 170 somatic mutations of 59 genes. Among these, 40 mutated genes were found in the metabolic disease group, and 33 mutated genes were found in the non-metabolic disease group. Importantly, nonsynonymous mutations of 26 genes (MSH3, BRAF, MLH3, MTOR, DDR2, ALK, etc.) were uniquely present in the metabolic disease group. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes enrichment analysis were performed to investigate biological functions and key pathways of somatic mutations. TP53, PIK3CA, and PTEN were the top three commonly mutated genes at a higher frequency compared with the Cancer Genome Atlas (TCGA) data, and several novel but infrequent mutations in other genes were also found. Although further studies are required to validate these variants, our results are the first to suggest a specific molecular profile of breast cancer with preexisting metabolic disease.

*Translational Oncology* (2020) 13, 245–253

Address all correspondence to: Huimin Liu, Biotecon Medical Diagnostics Co., Ltd., Zhangjiang Center for Translational Medicine, 180 Zhangheng Road, Shanghai 201203, China. E-mail: [hmliu@biotecon.com](mailto:hmliu@biotecon.com) or Binhui Yang, Department of Orthopaedic, The Aviation Hanzhong 3201 Hospital, Xi'an Jiao Tong University, 783 Tianhan Avenue, Hanzhong 723000, Shaanxi, China. E-mail: [ybh3201@163.com](mailto:ybh3201@163.com).

<sup>1</sup>Equal contributors: Wenjiang Jing, Ling Li and Xiumei Zhang contributed equally to this work and should be regarded as co-first authors.

Received 30 June 2019; Revised 18 September 2019; Accepted 19 September 2019

© 2019 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1936-5233/19

<https://doi.org/10.1016/j.tranon.2019.09.008>

## Introduction

Breast cancer alone is expected to account for 30% of all new cancers in women in Western countries, and the overall incidence has been consistently increasing ever since 1980s [1]. China has a low incidence of breast cancer, but since the 1990s, its incidence has increased more than twice as fast as global rates, particularly in urban areas [2]. Breast cancer is now the most frequently diagnosed cancer and is the sixth leading cause of cancer-related death in Chinese women [3]. Reproductive and hormonal factors, including a long menstrual life, increased age at first live birth, nulliparity, and limited breastfeeding, are associated with a modestly increased risk of breast cancer in the Chinese population [4–7]. Obesity and low levels of physical activity also probably contribute to the increasing incidence of breast cancer in China [8]. In accordance with the receptor status of the breast cancer cells, estrogen receptor (ER)-positive tumors include luminal types A and B, whereas ER-negative tumors include subtypes in which human epidermal growth factor receptor 2 (HER2, also known as ERBB2) is overexpressed and a basallike subtype that is triple negative for ER, the progesterone receptor, and HER2 [9]. The four main breast cancer subtypes are caused by different subsets of genetic and epigenetic abnormalities [10]. To date, using different technology platforms, molecular studies of breast cancer have identified diverse subtype-specific mutations and signaling pathway changes [11–15]. However, the epidemiological characteristics and genetic background are different between Chinese and Western population [2,16]. Further studies are needed to completely characterize the molecular architecture of breast cancer in Chinese women.

Chronic diseases including cancer, diabetes, and cardiovascular diseases are emerging as one of the greatest threats to human health in the 21st century. Higher body fatness is not only associated with a higher risk of hypertension, type 2 diabetes, and coronary heart disease but also with postmenopausal breast cancer [17]. Especially, metabolic reprogramming is fundamental for the development, rapid proliferation, and survival of cancer cells [18]. Hypertension has been implicated as a risk factor for breast cancer among postmenopausal women [19,20]. Furthermore, numerous studies have shown that patients with type 2 diabetes are at a greater risk of developing breast cancer [21–25]. Similarly, the link between type 2 diabetes and breast cancer appears to be most evident in postmenopausal women [21]. The prevalence of prior cardiac disease and of cardiovascular risk factors such as coronary heart disease, myocardial infarction, diabetes mellitus, arterial hypertension, hyperlipoproteinemia, and obesity rises with increasing age at breast cancer diagnosis [26]. The presence of preexisting cardiac diseases and risk factors adversely affects breast cancer survival, independent of specific therapy-related cardiotoxic effects [27]. Moreover, cross-talk between estrogen and the Notch pathway in breast cancer and coronary heart disease suggests a strong association between these two kinds of disease [28,29]. In a word, a wealth of studies have found that preexisting metabolic diseases are significantly associated with increasing breast cancer incidence and mortality. However, the molecular mechanisms underpinning this relationship are yet to be elucidated.

In this study, we established a hybridization capture-based panel including 124 cancer-associated genes and performed targeted next-generation sequencing (NGS) of tumor tissues and matched blood samples from 20 postmenopausal patients with primary breast cancer. We identified somatic mutations and biological consequences

associated with breast cancer with and without preexisting metabolic disease including hypertension, type 2 diabetes, and coronary heart disease. Our sequencing results were further compared with TCGA data to clarify the difference between them. To our knowledge, we present the first demonstration of the molecular profile of breast cancer with and without preexisting metabolic disease. These results provide further evidence for a critical role of the body metabolism during the initiation and progression of breast cancer.

## Materials and Methods

### Patients and Samples

We recruited 20 postmenopausal patients with primary breast cancer who were divided into metabolic disease group and non-metabolic disease group in accordance with whether suffering from preexisting metabolic disorders including hypertension, type 2 diabetes, and coronary heart disease (Table 1). There were no significant differences in clinical characteristics of breast cancer between two groups (Table 2). Fresh frozen tumor tissues and matched blood samples were collected from each patient. 10 patients were recruited at the Aviation Hanzhong 3201 Hospital (Shaanxi, China) between January 2017 and December 2017. 10 patients were recruited at the Tengzhou Central People's Hospital (Shandong, China) between January 2017 and December 2017. The experiments were performed with the understanding and written consent of each patient, and the investigation was performed in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* [30]. The present study was approved by the Medical Ethics Committee of Aviation Hanzhong 3201 Hospital, Xi'an Jiao Tong University, and Tengzhou Central People's Hospital, Jining Medical University.

### Panel Information

We established a hybridization capture-based NGS panel, Biotecan PanCancer Detection (BTC-PCD), which is capable of detecting

**Table 1.** Demographic Characteristics.

Characteristics	Metabolic Disease	Non-metabolic Disease
Age		
≤50	0	0
>50	6	14
Gender		
Male	0	0
Female	6	14
Family History		
Yes	0	0
No	6	14
Smoking		
Yes	0	0
No	6	14
Drinking		
Yes	0	0
No	6	14
Menopausal Status		
Premenopausal	0	0
Postmenopausal	6	14
Preexisting Metabolic Disease		
Hypertension	4	0
Type 2 diabetes	2	0
Coronary heart disease	1	0

*Metabolic disease*, the group in which patients with breast cancer suffered from preexisting metabolic disorders. *Non-metabolic disease*, the group in which patients with breast cancer were free from preexisting metabolic disorders.

**Table 2.** Clinical Characteristics.

Characteristics	Metabolic Disease	Non-metabolic Disease	P-value
<b>Invasive Ductal Carcinoma</b>			
Yes	6	14	
No	0	0	
<b>Tumor Size (cm)</b>			
≤2.0	1	5	
2.1–5.0	4	6	
>5.0	1	3	0.597
<b>Stage</b>			
I-II	4	6	
III-IV	2	8	0.329
<b>Lymph Nodes</b>			
Negative	2	5	
Positive	4	9	0.919
<b>ER Status</b>			
Negative	4	7	
Positive	2	7	0.492
<b>PR Status</b>			
Negative	4	9	
Positive	2	5	0.919
<b>HER2 Status</b>			
Negative	3	6	
Positive	3	8	0.769
<b>Ki67 Value</b>			
<14%	0	0	
≥14%	6	14	

ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.

protein-coding mutations in 124 cancer-associated genes including cancer genetic risk genes, targeted drugs (approved by FDA, clinical trials) and chemotherapy-associated genes, and prognosis genes. BTC-PCD panel was designed referring to cancer-related database, clinical guidelines, and high-quality references.

### DNA Extraction and Quality Control

Genomic DNA (gDNA) from fresh tissue was extracted by QIAamp DNA Mini Kit, and gDNA from blood by QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). Quantity and purity of gDNA were assessed by Qubit® 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA), and NanoDrop ND-1000 (Thermo Fisher Scientific, Wilmington, DE, USA). Fragmentation status was evaluated by the Agilent 2200 TapeStation system using the Genomic DNA ScreenTape assay (Agilent Technologies, Santa Clara, CA, USA) which was able to produce a DNA Integrity Number. An additional quality control (QC) step to assess fresh frozen tissue DNA integrity was performed using a multiplex polymerase chain reaction (PCR) approach. Briefly, 30 ng of gDNA were amplified using three different-size set of primers of glyceraldehyde-3-phosphate dehydrogenase gene (200–400 base pair), and the concentration of PCR products was determined by Agilent 2100 Bioanalyzer instrument (Agilent Technologies). Then, to estimate fresh frozen tissue gDNA fragmentation, we evaluated an Average Yield Ratio value, calculated by yield ratio of each amplicon compared with a reference DNA (Promega Madison, WI, USA).

### Library Preparation, Hybridization Capture, and Sequencing

A total of 300 ng of each gDNA sample based on Qubit quantification were mechanically fragmented on an E220-focused ultrasonicator Covaris (Covaris, Woburn, MA, USA). 200 ng of sheared gDNA were used to perform end repair, A-tailing, and adapter ligation with KAPA library preparation kits (Kapa Biosystems Inc. Wilmington, MA, USA) following the manufacturer instruc-

tions. Subsequently, the libraries were captured using xGen Lock-down Probe Pools (Integrated DNA Technologies, Coralville, IA, USA) and amplified. After QC and quantification by Agilent 2100 Bioanalyzer (Agilent Technologies) and Qubit® 2.0 Fluorometer (Invitrogen), the libraries were sequenced on an Illumina Next 500 platform (Illumina Inc, San Diego, CA, USA) High Output mode, 2 × 75 cycles.

### Bioinformatics Analysis

Clean data were obtained after filtering the low-quality reads, including reads with adapter sequences, reads with proportion of N more than 10%, and reads with low-quality base numbers more than 5. Reads were aligned to the reference human genome (UCSC hg19) using the Burrows-Wheeler Aligner v.0.7.12 [31,32]. Then, the Picard and Genome Analysis Toolkit (GATK v.3.2) method was adopted for duplicate removal, local realignment, and base quality recalibration and generated the quality statistics, including mapped reads, mean mapping quality, and mean coverage [33,34]. Finally, the GATK HaplotypeCaller was used for single nucleotide variation (SNV) and Insertion/Deletion (InDel) calling.

Variants were annotated using the ANNOVAR software tool. Annotations for mutation function (synonymous, nonsynonymous, stop-gain, frameshift, and unknown), mutation location (exonic, intronic, and untranslated region), amino acid changes, 1000 Genomes Project data, and dbSNP reference number were performed.

Somatic SNVs and InDels of tumors compared with matched blood samples were named using MuTect v.1.1.4 and VarScan2 v.2.3.9 software. The mutations with variant allele frequency >5% were defined as high confidence mutations. Then, gene mutation data were downloaded from TCGA database, and comparative analysis was performed using the sequencing data produced in the present study.

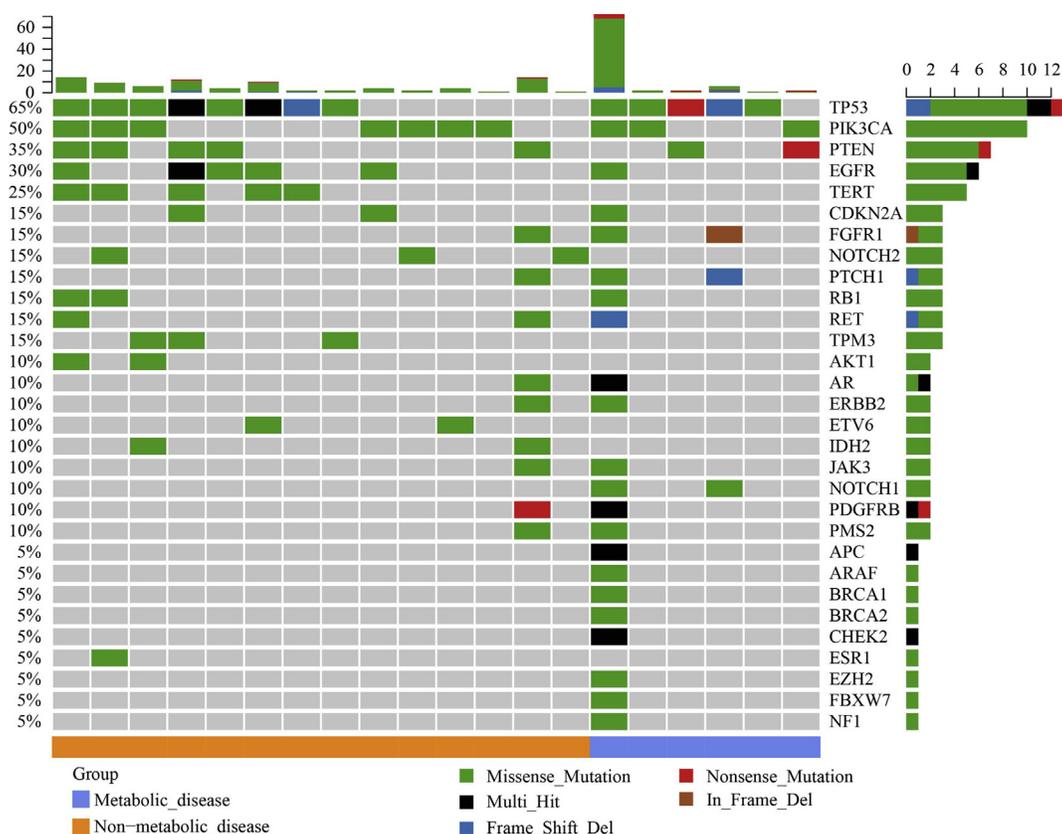
### Statistical Analysis

The mutation landscape and lollipop plot were created by the MafTools (v.1.8.0) in R software. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed to investigate the biological importance of the somatic mutated genes using the ClusterProfiler (v.3.10.1) in R software [35]. The cutoff of *p*-values <0.05 and FDR <0.05 were used to assess the significance of enrichment terms. For comparison of the clinical characteristics between two groups, Chi-square test was used.

## Results

### Demographic and Clinical Characteristics of Samples

20 postmenopausal patients with primary breast cancer suffering invasive ductal carcinoma were recruited (Tables 1 and 2). Among these patients, 14 cases made up the non-metabolic disease group and 6 cases were in the metabolic disease group presenting with established metabolic disorders including hypertension (4), type 2 diabetes (2), and coronary heart disease (1). In the metabolic disease group, one woman concurrently suffered from essential hypertension and type 2 diabetes before she was diagnosed with breast cancer. All patients were of Han nationality, and there was no statistical bias in demographic characteristics including age, gender, smoking, drinking, and family history. Fresh frozen tumor tissues were collected after surgery and histologically confirmed by two pathologists. There were no significant differences in receptor status, tumor size, TNM stage, and lymph node metastasis between two groups. Peripheral blood



**Figure 1.** Frequency and types of mutations in top 30 genes identified by targeted sequencing. Different mutations and subtypes are colored in the middle panel. Frequencies of mutations are shown on left (%). Top panel shows the number of nonsynonymous mutations in each sample, and right panel shows the number of samples with nonsynonymous mutations. Yellow bar represents non—metabolic disease group, whereas light blue bar stands for metabolic disease group at the bottom. Missense mutation, a point mutation in which a single nucleotide change results in a codon that codes for a different amino acid. Nonsense mutation, a point mutation that results in a premature stop codon or a nonsense codon. Frameshift del, a variant caused by deletion that moves the coding sequence out of frame. In-frame del, a deletion that keeps the sequence in frame. Multihit, contains multiple types of the aforementioned nonsynonymous mutations.

samples from these 20 patients were also collected and sequenced by BTC-PCD in the present study.

### Landscape of Somatic Mutations

We performed targeted NGS of tumor tissues and matched blood samples to study the somatic mutations of breast cancer with and without preexisting metabolic disease. After annotation of somatic SNVs and InDels, we focused only on protein-altering variants. As a result, we identified 170 somatic mutations of 59 genes in total 20 patients, which could be classified into five types including missense mutation, nonsense mutation, frameshift del, in-frame del, and multihit. Top 30 genes with somatic mutations identified by BTC-PCD were showed in Figure 1. The five most frequently mutated genes were TP53 (65%), PIK3CA (50%), PTEN (35%), EGFR (30%), and TERT (25%).

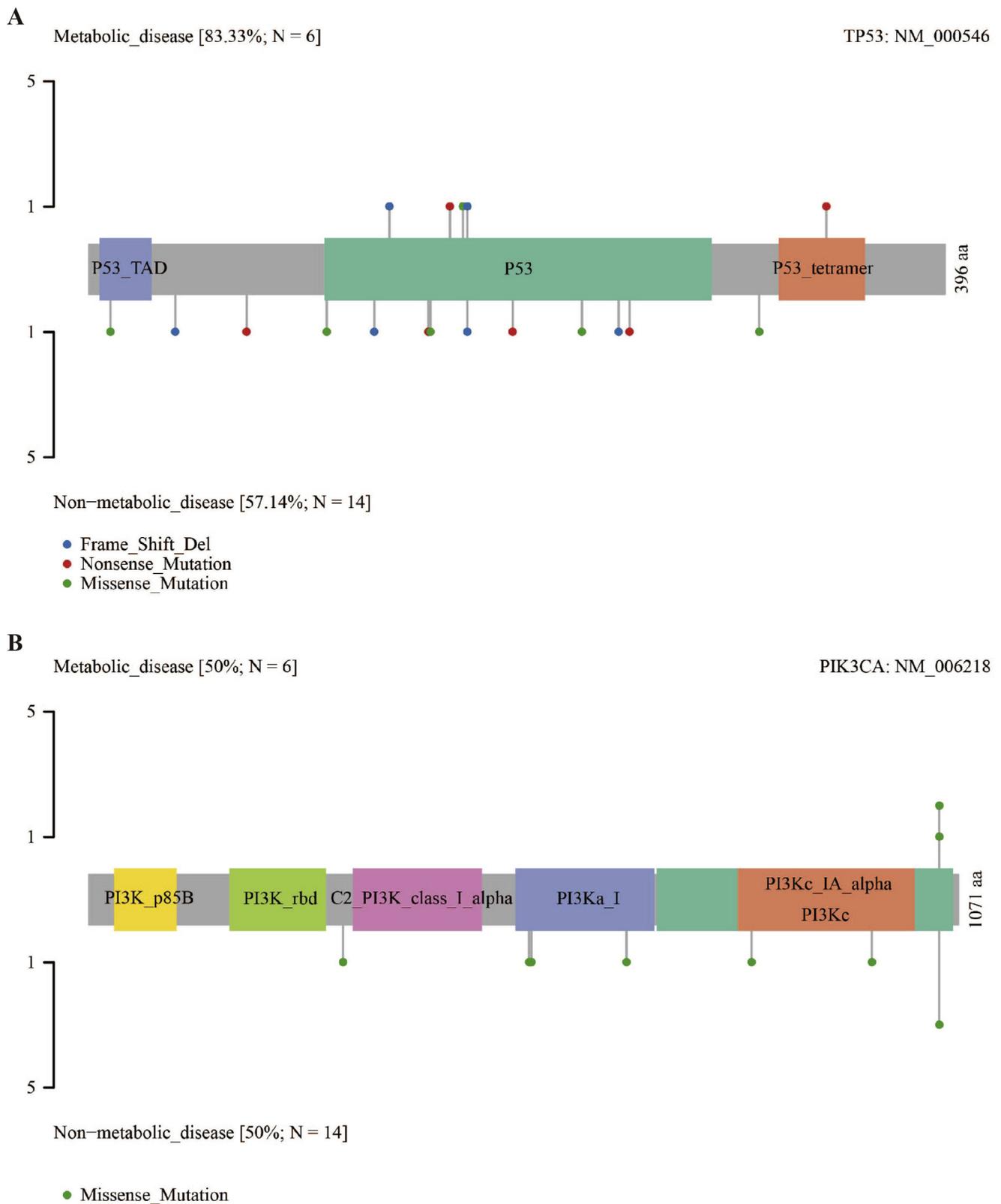
Most of the genes were present with different frequency and sites of mutations between two groups. For example, TP53 and PIK3CA showed different amino acid alterations in two groups (Figure 2). 83.33% patients harbored mutations in TP53 in the metabolic disease group, whereas only 57.14% patients in the non—metabolic\_disease group. Although half of patients harbored mutations in PIK3CA in two groups, their mutation frequency and sites were significantly different.

### Analysis of Somatic Mutation Differences Between Two Groups

Preexisting metabolic diseases are significantly associated with increasing breast cancer incidence and mortality. However, little is known about the mechanisms at the molecular level. To further investigate the difference of somatic mutations between two groups, we plotted Venn diagram of genes with somatic mutations (Figure 3). Among 59 somatic mutated genes, 40 mutated genes were found in the metabolic disease group and 33 mutated genes were found in the non—metabolic disease group. Importantly, somatic mutations of 26 genes including MSH3, BRAF, MLH3, MTOR, DDR2, ALK, etc. were uniquely present in the metabolic disease group, whereas somatic mutations of 19 genes including NRAS, TPM3, TERT, CD74, NOTCH2, OBSCN, etc. only emerged in the non—metabolic disease group. 14 somatic mutated genes including RET, AR, PIK3CA, PDGFRB, PMS2, EGFR, FGFR1, CDKN2A, PTCH1, RB1, TP53, ERBB2, JAK3, and PTEN concurrently emerged in two groups.

### Biological Functions and Key Pathways Analysis of Somatic Mutations in Two Groups

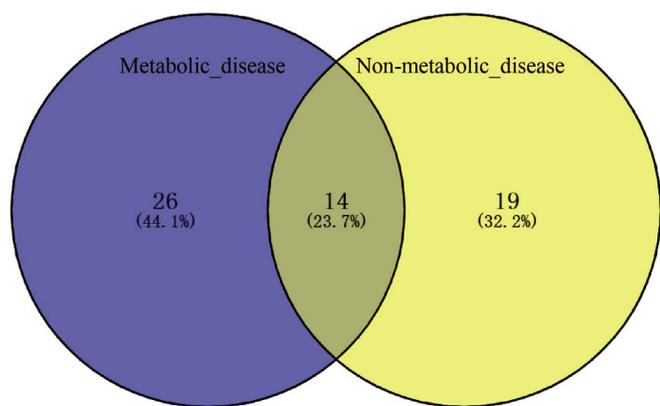
To better understand the biological consequences of the above--described somatic mutations, we performed GO and KEGG



**Figure 2.** Amino acid alterations of TP53 (A) and PIK3CA (B) in two groups. Structure domains are marked in different colors. Lollipops show the locations of protein-altering variants. The proportion of patient-harbored nonsynonymous mutations in each group is shown in the square brackets.

enrichment analysis using the ClusterProfiler (v.3.10.1) in R software (Figures 4 and 5). GO enrichment results showed that functional categories were most involved in extracellular signal-regulated kinase

1 (ERK1) and ERK2 signaling in the metabolic disease group (Figure 4A) and small RNA processing in the non-metabolic disease group (Figure 5A). Top 15 pathways were depicted in accordance



**Figure 3.** Venn diagram of genes with somatic mutations in metabolic disease and non-metabolic disease groups.

with gene count and  $p$ -value enriched by KEGG, and most pathways involved were cancer-related (Figures 4B and 5B). Furthermore, 26 genes uniquely present in the metabolic disease group were prevalently distributed in proteoglycans pathway, epidermal growth factor receptor (EGFR) signaling, and PI3K-Akt signaling. Comparatively, 19 genes that specifically emerged in the non-metabolic disease group were more enriched in pathways associated with EGFR signaling, hormone signaling, and PD-1/PD-L1 signaling.

#### Comparison of Somatic Mutated Genes Identified by TCGA Cohort and BTC-PCD Cohort

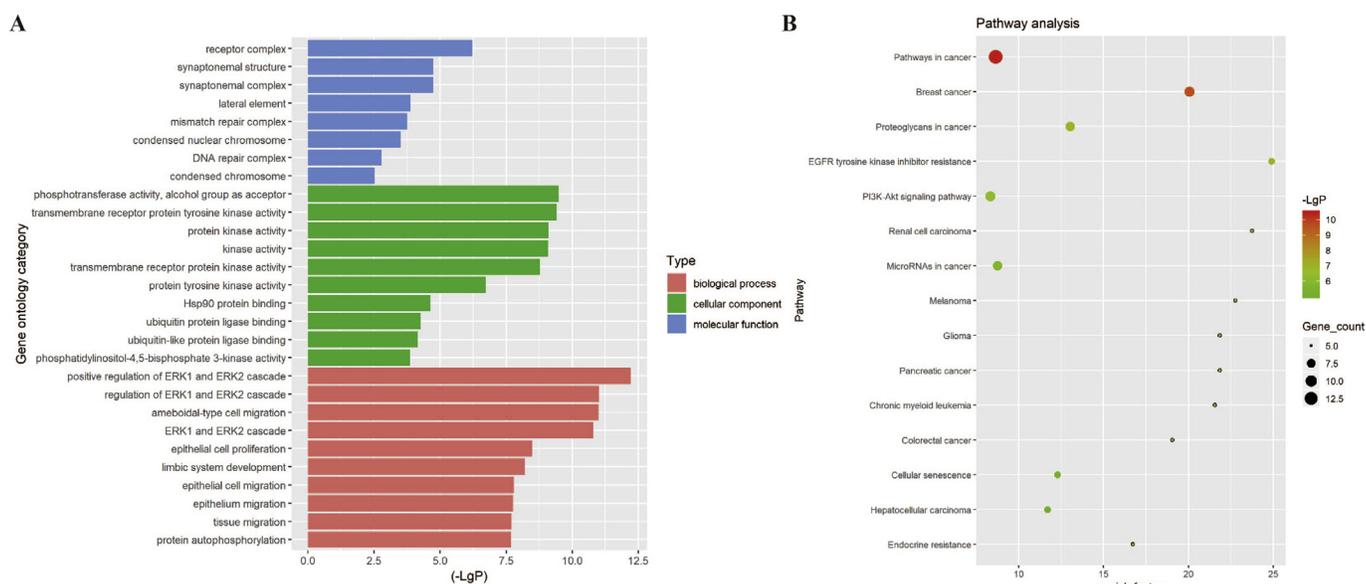
Comparative analysis of somatic mutated genes was performed between BTC-PCD cohort and TCGA cohort (Figure 6). The original data of breast cancer downloaded from TCGA database were firstly aligned to BTC-PCD panel and excluded genes that were absent in our panel. Thus, TCGA cohort used in our study only contained 83 genes which were covered by BTC-PCD panel. We

observed important differences between two cohorts. 47 somatic mutated genes identified by TCGA cohort were covered by BTC-PCD cohort, including TP53, PIK3CA, PTEN, OBSCN, NF1, AKT1, etc (Figure 6A). Overall, genes identified by BTC-PCD cohort were at a higher mutation frequency compared with that in TCGA cohort (Figure 6B). 12 novel somatic mutated genes were identified in this study, including TERT (25%), CDKN2A (15%), FGFR1 (15%), IDH2 (10%), BRCA1 (5%), BRCA2 (5%), CDA (5%), FGFR3 (5%), VHL (5%), STK11 (5%), SMARCB1 (5%), and NRAS (5%).

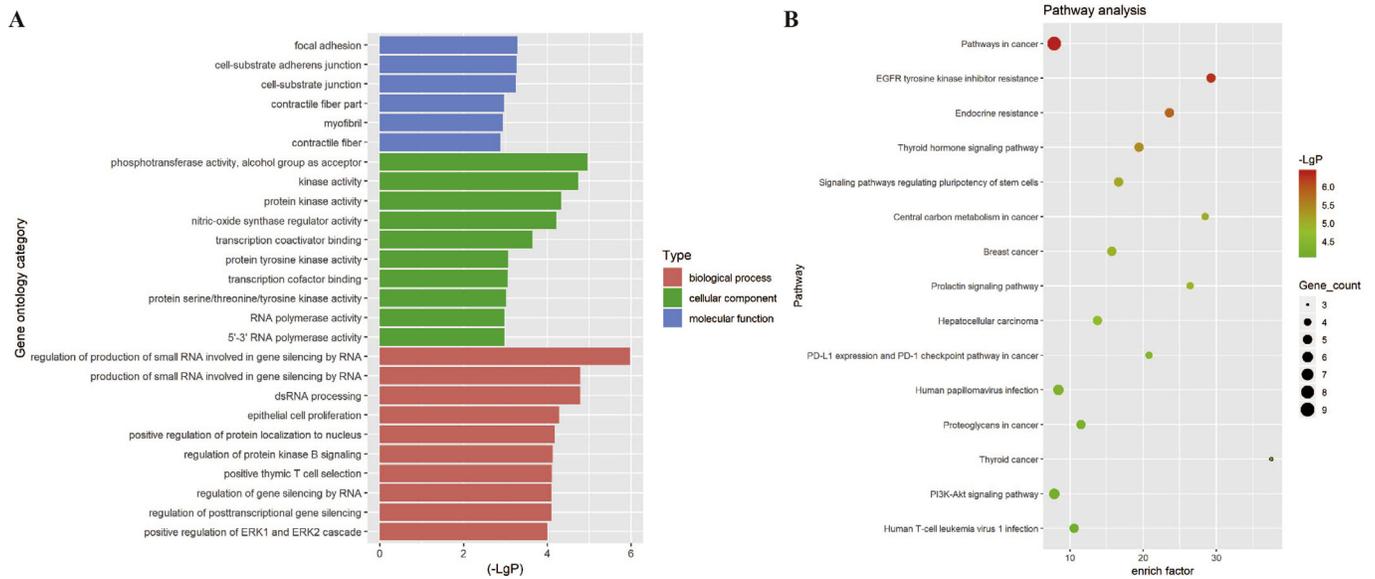
#### Discussion

Altered metabolic features are observed quite generally across many types of cancer cells, and reprogrammed metabolism is considered a hallmark of cancer [36,37]. Breast cancer is the most frequent type of cancer in women worldwide [38]. A lot of studies have found that preexisting metabolic diseases, including hypertension, type 2 diabetes, and coronary heart disease, are significantly associated with increasing breast cancer incidence and mortality. However, the molecular mechanism to address this relationship is still obscure. In this study, we investigated the role of preexisting metabolic disease in regulating the development and progression of breast cancer at the molecular level.

Somatic mutations that occur in tumor cell genomes play a vital role in cancer development. We used germline DNA from blood samples as a reference for detecting somatic SNVs and InDels in tumor tissues of 20 postmenopausal patients with primary breast cancer. We totally identified 170 somatic mutations of 59 genes. Intriguingly, 26 genes including MSH3, BRAF, MLH3, MTOR, DDR2, ALK, KIT, KDR, FBXW7, APC, ROS1, CDK6, MET, SMO, EZH2, TSC1, NOTCH1, MDM2, PTPN11, BRCA2, TSC2, NF1, BRCA1, CHEK2, ARAF, and VHL were uniquely mutated in the metabolic disease group (Figure 3). The clinical significances of these somatic mutated genes were very important. For



**Figure 4.** Biological functions and key pathways analysis of somatic mutations in the metabolic disease group. A. GO analysis of somatic mutations in the metabolic disease group. B. KEGG analysis of somatic mutations in the metabolic disease group. Gene count represents the number of mutated genes enriched in each pathway. GO, Gene ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes.

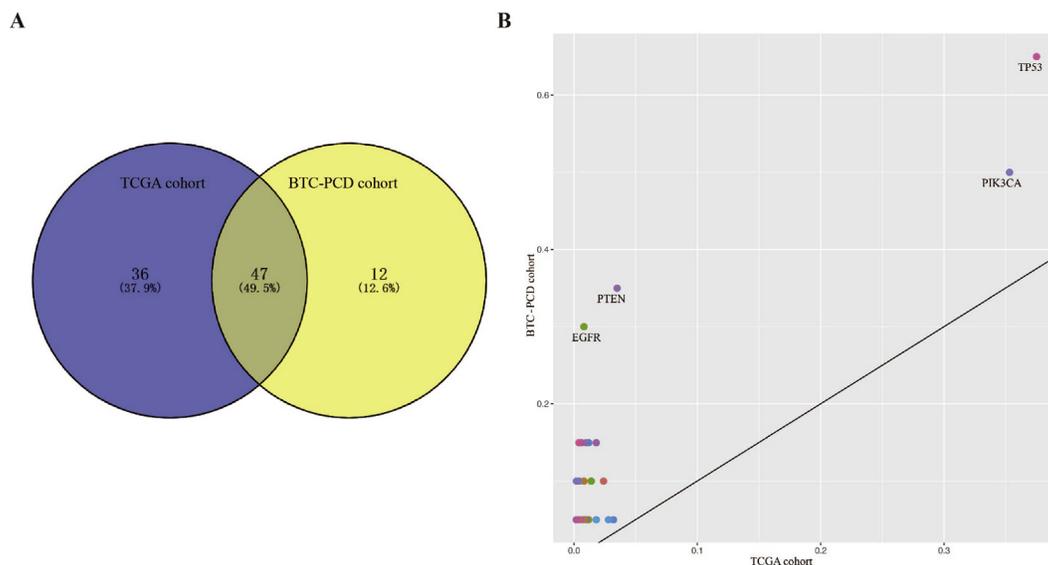


**Figure 5.** Biological functions and key pathways analysis of somatic mutations in the non-metabolic disease group. A. GO analysis of somatic mutations in the non-metabolic disease group. B. KEGG analysis of somatic mutations in the non-metabolic disease group.

example, previous works have demonstrated that MSH3 germline variants were associated with breast cancer risk and radiosensitivity in patients with breast cancer [39,40]. A somatic variant caused by deletion was found in our study. This is the first demonstration of MSH3 as somatic gene in breast cancer with preexisting metabolic disease. Besides, somatic mutations in BRAF, MLH3, MTOR, DDR2, ALK, FBXW7, APC, ROS1, CDK6, MET, SMO, EZH2, TSC1, NOTCH1, MDM2, TSC2, NF1, CHEK2, and ARAF genes in human breast tumor were reported by TCGA [10]. In the present study, we further found that these genes were specially mutated in the metabolic disease group. These results may help to better understand the underlying mechanism of breast cancer with preexisting metabolic

disease, providing promising therapeutic targets in this specific subgroup. Further studies are required to verify whether these somatic mutations cause a different pathogenetic mechanism or correlate with patients' outcome.

GO and KEGG enrichment analysis were performed to understand the biological consequences of somatic mutations in two groups. GO enrichment showed that functional categories were most involved in ERK1 and ERK2 signaling in the metabolic disease group (Figure 4A). Conventional mitogen-activated protein kinases (MAPKs) include p38 MAPK, ERK1/2, c-Jun N-terminal kinases/stress-activated protein kinases, and ERK5 [41]. Activation of ERK1/2 pathways is important for breast cancer-cell proliferation [42,43].



**Figure 6.** Comparison of somatic mutated genes identified by TCGA cohort and BTC-PCD cohort. A. Venn diagram of somatic mutated genes identified by TCGA cohort and BTC-PCD cohort. B. Comparison of the frequency of 47 somatic mutated genes identified by TCGA cohort and BTC-PCD cohort. TCGA, The Cancer Genome Atlas. BTC-PCD, Biotecan PanCancer Detection.

Furthermore, KEGG enrichment demonstrated that 26 genes uniquely present in the metabolic disease group were prevalently distributed in proteoglycans pathway, EGFR signaling, and PI3K-Akt signaling (Figure 4B). Hyperactivation of PI3K-Akt signaling pathway is considered as a hallmark in a wide spectrum of human cancers [44–46]. A number of studies have demonstrated that PI3K-Akt signaling involves in regulation of metabolism, growth, survival, angiogenesis, and metastasis of tumor cell [47,48]. Thus, ERK1/ERK2 signaling and PI3K-Akt signaling may play a critical role in the initiation and progression of breast cancer with preexisting metabolic disease, and how they contribute to pathogenetic mechanism could be the subject of future exploration.

Finally, we compared somatic mutated genes identified by TCGA cohort and BTC-PCD cohort. 83 somatic mutated genes reported by TCGA database were covered by our BTC-PCD panel (Figure 6). Among these, 47 somatic mutated genes identified by TCGA cohort were detected by BTC-PCD cohort at a higher mutation frequency. In particular, 12 novel somatic mutated genes including TERT, CDKN2A, FGFR1, IDH2, BRCA1, BRCA2, CDA, FGFR3, VHL, STK11, SMARCB1, and NRAS were identified in the present study, which were not reported by TCGA before. Actually, many of them, such as telomerase reverse transcriptase (TERT) promoter hotspot mutations and cyclin-dependent kinase inhibitor 2A (CDKN2A) variants, have been found in breast cancer, which were highly consistent with the BTC-PCD findings [49,50]. These differences could be resulted from the limited sample size, population, or bioinformatics tools differences and need further investigation.

Taken together, we investigated the genetic profile of breast cancer with and without preexisting metabolic disease including hypertension, type 2 diabetes, and coronary heart disease. By targeted NGS of tumor tissues and matched blood samples, we identified 170 somatic mutations of 59 genes in 20 postmenopausal primary patients with breast cancer. We observed significant differences in somatic mutations and biological consequences between the metabolic disease group and the non-metabolic disease group. How precisely the regulation occurs and/or whether preexisting metabolic disease interacts with other known pathways regulating breast cancer-cell proliferation and growth remains to be determined in future studies.

### Funding

This study was funded by grants from the Supporting Fund for Teachers' Research of Jining Medical University (grant no. JY2016KJ041Y) and the Shandong Province Medical and Health Technology Development Project (grant no. 2016WS0627).

### Availability of Data and Materials

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

### Authors' Contributions

BY, XZ, WM, SL conceived the study and designed the experiments; HL performed experiments, analyzed the data, and wrote the manuscript; WJ, LL collected samples and analyzed the clinical data; SW, JZ, QH, HW performed bioinformatics analysis. All authors read and approved the final manuscript.

### Ethics Approval and Consent to Participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Medical Ethics Committee of Aviation Hanzhong 3201 Hospital and Tengzhou Central People's Hospital, and with the 1964 Helsinki declaration

and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the present study.

### Consent to Publish

Informed consent for the publication of this study was obtained from all individual participants.

### Competing Interests

The authors declare that they have no competing interests.

### Acknowledgments

Not applicable.

### References

- [1] Siegel RL, Miller KD and Jemal A (2019). Cancer statistics, 2019. *CA Cancer J Clin* **69**(1), 7–34.
- [2] Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, Chen WQ, Shao ZM and Goss PE (2014). Breast cancer in China. *Lancet Oncol* **15**(7), e279–e289.
- [3] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J (2016). Cancer statistics in China, 2015. *CA Cancer J Clin* **66**(2), 115–132.
- [4] Gao YT, Shu XO, Dai Q, Potter JD, Brinton LA, Wen W, Sellers TA, Kushi LH, Ruan Z and Bostick RM, et al (2000). Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. *Int J Cancer* **87**(2), 295–300.
- [5] Yanhua C, Geater A, You J, Li L, Shaoqiang Z, Chongsuvivatwong V and Sriplung H (2012). Reproductive variables and risk of breast malignant and benign tumours in Yunnan province, China. *Asian Pac J Cancer Prev* **13**(5), 2179–2184.
- [6] Zhang Q, Liu LY, Wang F, Mu K and Yu ZG (2012). The changes in female physical and childbearing characteristics in China and potential association with risk of breast cancer. *BMC Public Health* **12**, 368.
- [7] Bao PP, Shu XO, Gao YT, Zheng Y, Cai H, Deming SL, Ruan ZX, Su Y, Gu K and Lu W, et al (2011). Association of hormone-related characteristics and breast cancer risk by estrogen receptor/progesterone receptor status in the shanghai breast cancer study. *Am J Epidemiol* **174**(6), 661–671.
- [8] Shu O, Jin F, Dai Q, Shi JR, Potter JD, Brinton LA, Hebert JR, Ruan Z, Gao YT and Zheng W (2001). Association of body size and fat distribution with risk of breast cancer among Chinese women. *Int J Cancer* **94**(3), 449–455.
- [9] Lin SX, Chen J, Mazumdar M, Poirier D, Wang C, Azzi A and Zhou M (2010). Molecular therapy of breast cancer: progress and future directions. *Nat Rev Endocrinol* **6**(9), 485–493.
- [10] Cancer Genome Atlas N (2012). Comprehensive molecular portraits of human breast tumours. *Nature* **490**(7418), 61–70.
- [11] Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Bloom MJ, Jervis S, van Leeuwen FE, Milne RL and Andrieu N, et al (2017). Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *J Am Med Assoc* **317**(23), 2402–2416.
- [12] Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S and Wedge DC, et al (2016). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**(7605), 47–54.
- [13] Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, Wang X, Qiao JW, Cao S and Petralia F, et al (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**(7605), 55–62.
- [14] Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, Aas T, Alexandrov LB, Larsimont D and Davies H, et al (2015). Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med* **21**(7), 751–759.
- [15] Liu YR, Jiang YZ, Xu XE, Yu KD, Jin X, Hu X, Zuo WJ, Hao S, Wu J and Liu GY, et al (2016). Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. *Breast Cancer Res* **18**(1), 33.

- [16] Jiang YZ, Yu KD, Bao J, Peng WT and Shao ZM (2014). Favorable prognostic impact in loss of TP53 and PIK3CA mutations after neoadjuvant chemotherapy in breast cancer. *Cancer Res* **74**(13), 3399–3407.
- [17] Nimptsch K and Pischon T (2015). Body fatness, related biomarkers and cancer risk: an epidemiological perspective. *Horm Mol Biol Clin Investig* **22**(2), 39–51.
- [18] DeBerardinis RJ and Chandel NS (2016). Fundamentals of cancer metabolism. *Sci Adv* **2**(5):e1600200.
- [19] Staessen JA, Wang J, Bianchi G and Birkenhäger WH (2003). Essential hypertension. *Lancet* **361**(9369), 1629–1641.
- [20] Han H, Guo W, Shi W, Yu Y, Zhang Y, Ye X and He J (2017). Hypertension and breast cancer risk: a systematic review and meta-analysis. *Sci Rep* **7**, 44877.
- [21] Xue F and Michels KB (2007). Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. *Am J Clin Nutr* **86**(3), s823–s835.
- [22] Martin SD and McGee SL (2018). Metabolic reprogramming in type 2 diabetes and the development of breast cancer. *J Endocrinol* **237**(2), R35–R46.
- [23] Muck BR, Trotnow S and Hommel G (1975). Cancer of the breast, diabetes and pathological glucose tolerance. *Arch Gynakol* **220**(1), 73–81.
- [24] Talamini R, Franceschi S, Favero A, Negri E, Parazzini F and La Vecchia C (1997). Selected medical conditions and risk of breast cancer. *Br J Cancer* **75**(11), 1699–1703.
- [25] Qi , He P, Yao H, Song R, Ma C, Cao M, Cui B and Ning G (2019 Nov). Cancer risk among patients with type 2 diabetes: a real-world study in Shanghai, China. *J Diabetes* **11**(11), 878–883.
- [26] Patnaik JL, Byers T, Diguiseppe C, Denberg TD and Dabelea D (2011). The influence of comorbidities on overall survival among older women diagnosed with breast cancer. *J Natl Cancer Inst* **103**(14), 1101–1111.
- [27] Schmitz KH, Prosnitz RG, Schwartz AL and Carver JR (2012). Prospective surveillance and management of cardiac toxicity and health in breast cancer survivors. *Cancer* **118**(8 Suppl), 2270–2276.
- [28] Rizzo P, Miao H, D'Souza G, Osipo C, Song LL, Yun J, Zhao H, Mascarenhas J, Wyatt D and Antico G, et al (2008). Cross-talk between notch and the estrogen receptor in breast cancer suggests novel therapeutic approaches. *Cancer Res* **68**(13), 5226–5235.
- [29] Fortini F, Vieceli Dalla Sega F, Caliceti C, Lambertini E, Pannuti A, Peiffer DS, Balla C and Rizzo P (2019). Estrogen-mediated protection against coronary heart disease: the role of the Notch pathway. *J Steroid Biochem Mol Biol* **189**, 87–100.
- [30] Gandevia B and Tovell A (1964). Declaration of Helsinki. *Med J Aust* **2**, 320–321.
- [31] Li H and Durbin R (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**(5), 589–595.
- [32] Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G and Rhead B, et al (2013). The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res* **41**(Database issue), D64–D69.
- [33] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S and Daly M, et al (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**(9), 1297–1303.
- [34] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA and Hanna M, et al (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**(5), 491–498.
- [35] Yu G, Wang LG, Han Y and He QY (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**(5), 284–287.
- [36] Hanahan D and Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* **144**(5), 646–674.
- [37] Pavlova NN and Thompson CB (2016). The emerging hallmarks of cancer metabolism. *Cell Metab* **23**(1), 27–47.
- [38] Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* **127**(12), 2893–2917.
- [39] Miao HK, Chen LP, Cai DP, Kong WJ, Xiao L and Lin J (2015). MSH3 rs26279 polymorphism increases cancer risk: a meta-analysis. *Int J Clin Exp Pathol* **8**(9), 11060–11067.
- [40] Mangoni M, Bisanzio S, Carozzi F, Sani C, Biti G, Livi L, Barletta E, Costantini AS and Gorini G (2011). Association between genetic polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT genes and radiosensitivity in breast cancer patients. *Int J Radiat Oncol Biol Phys* **81**(1), 52–58.
- [41] Arthur JS and Ley SC (2013). Mitogen-activated protein kinases in innate immunity. *Nat Rev Immunol* **13**(9), 679–692.
- [42] Ventura C, Venturino A, Miret N, Randi A, Rivera E, Núñez M and Cocca C (2015). Chlorpyrifos inhibits cell proliferation through ERK1/2 phosphorylation in breast cancer cell lines. *Chemosphere* **120**, 343–350.
- [43] Liu HJ, Ooms LM, Srijakotre N, Man J, Vieusseux J, Waters JE, Feng Y, Bailey CG, Rasko JE and Price JT, et al (2016). PtdIns(3,4,5)P3-dependent Rac exchanger 1 (PREX1) Rac-guanine nucleotide exchange factor (GEF) activity promotes breast cancer cell proliferation and tumor growth via activation of extracellular signal-regulated kinase 1/2 (ERK1/2) signaling. *J Biol Chem* **291**(33), 17258–17270.
- [44] Martini M, De Santis MC, Braccini L, Gulluni F and Hirsch E (2014). PI3K/AKT signaling pathway and cancer: an updated review. *Ann Med* **46**(6), 372–383.
- [45] Chen H, Zhou L, Wu X, Li R, Wen J, Sha J and Wen X (2016). The PI3K/AKT pathway in the pathogenesis of prostate cancer. *Front Biosci (Landmark Ed)* **21**, 1084–1091.
- [46] Corti F, Nichetti F, Raimondi A, Niger M, Prinzi N, Torchio M, Tamborini E, Perrone F, Pruneri G and Di Bartolomeo M, et al (2019). Targeting the PI3K/AKT/mTOR pathway in biliary tract cancers: a review of current evidences and future perspectives. *Cancer Treat Rev* **72**, 45–55.
- [47] Koundouros N and Poulogiannis G (2018). Phosphoinositide 3-kinase/Akt signaling and redox metabolism in cancer. *Front Oncol* **8**, 160.
- [48] Liu X, Xu Y, Zhou Q, Chen M, Zhang Y, Liang H, Zhao J, Zhong W and Wang M (2018). PI3K in cancer: its structure, activation modes and role in shaping tumor microenvironment. *Future Oncol* **14**(7), 665–674.
- [49] Shimoi T, Yoshida M, Kitamura Y, Yoshino T, Kawachi A, Shimomura A, Noguchi E, Yunokawa M, Yonemori K and Shimizu C, et al (2018). TERT promoter hotspot mutations in breast cancer. *Breast Cancer* **25**(3), 292–296.
- [50] Aftab A, Shahzad S, Hussain HMJ, Khan R, Irum S and Tabassum S (2019). CDKN2A/P16INK4A variants association with breast cancer and their in-silico analysis. *Breast Cancer* **26**(1), 11–28.