Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2024**; 29(1): 2
https://doi.org/10.31083/j.fbl2901002

Original Research

Optical Genome Mapping Reveals the Landscape of Structural Variations and Their Clinical Significance in HBOC-Related Breast Cancer

Yanan Cheng[1,2], Li Dong[1,2], Dechao Bu[3], Lei Han[1,2], Yi Zhao[3], Juntian Liu[4], Xiaojing Guo[5], Hanli Xu[6,*], Jinpu Yu[1,2,*]

[1]Cancer Molecular Diagnostics Core, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, 300060 Tianjin, China
[2]Key Laboratory of Cancer Prevention and Therapy, Key Laboratory of Cancer Immunology and Biotherapy, Key Laboratory of Breast Cancer Prevention and Therapy, Tianjin Medical University, Ministry of Education, 300060 Tianjin, China
[3]Research Center for Ubiquitous Computing Systems, Institute of Computing Technology, Chinese Academy of Sciences, 100190 Beijing, China
[4]Cancer Prevention Center, Second Department of Breast Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, 300060 Tianjin, China
[5]Department of Breast Pathology and Lab, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, 300060 Tianjin, China
[6]College of Life Sciences and Bioengineering, School of Physical Science and Engineering, Beijing Jiaotong University, 100044 Beijing, China
*Correspondence: xuhanli@bjtu.edu.cn (Hanli Xu); jyu@tmu.edu.cn (Jinpu Yu)
Academic Editor: Peter Brenneisen
Submitted: 13 July 2023 Revised: 18 August 2023 Accepted: 25 August 2023 Published: 9 January 2024

Abstract

Background: Structural variations (SVs) are common genetic alterations in the human genome. However, the profile and clinical relevance of SVs in patients with hereditary breast and ovarian cancer (HBOC) syndrome (germline *BRCA1/2* mutations) remains to be fully elucidated. **Methods**: Twenty HBOC-related cancer samples (5 breast and 15 ovarian cancers) were studied by optical genome mapping (OGM) and next-generation sequencing (NGS) assays. **Results**: The SV landscape in the 5 HBOC-related breast cancer samples was comprehensively investigated to determine the impact of intratumor SV heterogeneity on clinicopathological features and on the pattern of genetic alteration. SVs and copy number variations (CNVs) were common genetic events in HBOC-related breast cancer, with a median of 212 SVs and 107 CNVs per sample. The most frequently detected type of SV was insertion, followed by deletion. The 5 HBOC-related breast cancer samples were divided into SV^{high} and SV^{low} groups according to the intratumor heterogeneity of SVs. SV^{high} tumors were associated with higher Ki-67 expression, higher homologous recombination deficiency (HRD) scores, more mutated genes, and altered signaling pathways. Moreover, 60% of the HBOC-related breast cancer samples displayed chromothripsis, and 8 novel gene fusion events were identified by OGM and validated by transcriptome data. **Conclusions**: These findings suggest that OGM is a promising tool for the detection of SVs and CNVs in HBOC-related breast cancer. Furthermore, OGM can efficiently characterize chromothripsis events and novel gene fusions. SV^{high} HBOC-related breast cancers were associated with unfavorable clinicopathological features. SVs may therefore have predictive and therapeutic significance for HBOC-related breast cancers in the clinic.

Keywords: HBOC; breast cancer; optical genome mapping; structural variation; *BRCA1/2*; chromothripsis

1. Introduction

Breast cancer is one of the most common malignant tumor types worldwide [1]. Over the past decade, the study of breast cancer genomics has been greatly assisted by advances in large-scale next-generation sequencing (NGS) technology [2,3]. Structural variations (SVs) are large-size genetic variations in the human genome, and include insertion, deletion, duplication, inversion, and translocation. SVs have been associated with different traits and with various diseases, including breast cancer [4–6]. They contribute to gene fusion, oncogene amplification, tumor suppressor gene deletion and other complex alterations leading to evolution of the cancer genome. These alterations have the potential to impact large stretches of DNA sequence, thereby disrupting genes and regulatory elements [7–9]. Many SVs are closely linked to tumorigenesis and have been used for tumor subtyping and diagnosis, as well as for effective targeted therapy [10–14]. SVs are very prevalent in breast cancer cell lines [12]. Deletions in breast cancer genomes can also affect enhancers and thus contribute to oncogenesis [15]. SVs have been reported in the breast cancer cell line SK-BR-3 and in patient-derived organoids. Detailed maps of the breast cancer genome have been established by integrating SV profiles, thereby revealing how SVs can disrupt the genome and also shedding light on the complex mechanisms involved in evolution of the cancer genome [8,16].

Hereditary breast and ovarian cancer (HBOC) syndrome is most commonly characterized by deleterious germline mutations in the *BRCA1* or *BRCA2* genes. Advances in NGS technology have led to the discovery of several non-*BRCA* genes that are also responsible for HBOC syndrome, such as mutations in the *PALB2*, *ATM*, *BRIP1*, *RAD51D* and *RAD51C* genes, etc. [17]. The HBOC syndrome is estimated to cause 5–10% of all breast cancers [18,19]. *BRCA1* and *BRCA2* are tumor suppressor genes that play a crucial role in the cell by rehabilitating damaged DNA in the homologous recombination repair (HRR) pathway [18]. Breast cancers with abnormal DNA repair functions (homologous recombination deficiency, HRD) are more likely to exhibit genomic instability, including abnormal SVs [20]. However, there is still only limited knowledge regarding structural abnormalities and SV heterogeneity in HBOC-related cancers, with more research required on this topic.

However, the genome-wide detection of SVs remains challenging. Karyotype analysis is the traditional approach used for identifying SVs, but is limited by the poor quality of mitotic chromosome metaphases and the low resolution of this technique [21]. More recently, NGS technology has greatly enhanced the resolution and throughput of genetic analysis and facilitated the discovery of SVs. However, because of the short read-length, NGS has difficulty with SVs located in repetitive regions and regions with high or low GC content [12,22]. Moreover, algorithms that are used to interpret SVs from NGS data with short read-lengths have a high false-negative rate [6,23]. Third-generation, long-read, single-molecule sequencing technologies from Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) have proven more reliable in identifying SVs, with substantial improvements in both sensitivity and specificity. However, these technologies still have a relatively low accuracy and also require deep sequencing to detect SVs [6,16,23].

Recent advances in long-read sequencing technology have opened the possibility of more precise and sensitive detection of SVs [24]. Optical genome mapping (OGM) is a nanochannel-based genome mapping technology that can detect multiple classes of SVs with high resolution. OGM can generate kilobase- to megabase-size genomic maps by automatically tagging single molecules with fluorophores, thereby providing images of labeled and linearized ultra-high molecular weight (UHMW, >150 kbp) DNA molecules [21,24]. Accurate and precise labelling, together with the rare variant analysis pipeline for solid tumors, allows the detection of low-level mosaic SVs by comparing single molecules directly with the human genome. In addition, a separate, coverage-based algorithm allows the detection of large copy number variations (CNVs) and aneuploidies [21].

The OGM technique was used in this study to investigate SVs in HBOC-related breast cancers with *BRCA1* or *BRCA2* mutation. This should help to determine the impact of intratumor heterogeneity in SVs on clinicopathological features and on the genetic alteration profile. To our knowledge, this is the first study that uses OGM technology to comprehensively analyze SVs in HBOC-related breast cancer with *BRCA1* or *BRCA2* mutation.

2. Materials and Methods

2.1 Patient Enrollment

OGM analysis was performed on 20 HBOC syndrome patients, comprising 5 breast cancer patients and 15 ovarian cancer cases. These were identified from more than 200 breast and ovarian cancer patients from our earlier studies [25,26]. The 5 HBOC-related breast cancer patients were comprehensively investigated in the present study. All 5 patients were diagnosed with invasive carcinoma (stage I/II) and completed adjuvant therapy following surgery. Immunohistochemical staining was performed for Ki-67, estrogen receptor (ER), progesterone receptor (PR), and receptor tyrosine-protein kinase erbB-2 (HER2). Breast cancer samples were identified as luminal A, luminal B, HER-2 overexpression, or triple negative breast cancer (TNBC) subtypes according to the protein expression levels of ER, PR and HER2. Ki-67 expression was classified as either Ki-67^{low} ($\leq$0.3) or Ki-67^{high} (>0.3). Table 1 shows the clinicopathological data, *BRCA1/BRCA2* status, and other risk gene status for the 5 breast cancer patients. **Supplementary Table 1** shows the clinicopathological data, *BRCA1/BRCA2* and other risk gene status for the 15 HBOC-related ovarian cancers. This project was approved by the Ethics Committee of Tianjin Medical University Cancer Institute and Hospital (No. Ek2018050). All experiments were performed in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained from each patient.

2.2 Optical Genome Mapping

UHMW genomic DNA (gDNA) was extracted from flash-frozen tissue stored at –80 °C. This was performed using the Bionano Prep Animal Tissue DNA Kit (Bionano Genomics, San Diego, CA, USA) as recommended by the manufacturer. The gDNA was left to homogenise overnight at room temperature. The next day, DNA molecules were labelled using the DLS (Direct Label and Stain) DNA Labeling Kit (Bionano Genomics, San Diego, CA, USA) as recommended by the manufacturer. A Proteinase K solution (Qiagen, Germantown, MD, USA) was then used to inactivate the enzyme, and successive membrane adsorption steps were used for cleanup. The DNA backbone was counterstained overnight before quantification, and the labelled gDNA solution was then loaded onto a Bionano Saphyr chip and scanned on the Bionano Saphyr instrument (Bionano Genomics, San Diego, CA, USA).

2.3 Calling of Structural Variants and Variant Filtering

Genome analysis and the calling of structural variations was performed using the rare variant pipeline (RVP) in Bionano Solve (v3.7, Bionano Genomics, San Diego, CA, USA). The variant hg19 DLE-1 SV mask, which blocks difficult-to-map regions and common artifacts, was turned on for data filtering. The following recommended confidence scores were then applied: insertion, 0; deletion, 0; inversion, 0.7; duplication, –1; intra- and inter-translocation, 0.05; copy number, 0.99 (low stringency, filter set to 0). All SVs and CNVs detected in each sample were exported in SMAP files.

2.4 Calculation of the HRD Score Using OGM Data

The HRD score combines three independent measures of genomic instability, namely genome-wide loss of heterozygosity (LOH), telomeric allelic imbalance (TAI), and large-scale state transitions (LST). These were calculated on the basis of the label and coverage of the ultra-long DNA molecule. LOH calculation was based on large deletions, which counts the number of regions representing one parental allele longer than 15 Mb but shorter than the whole chromosome. TAI represents the number of regions with CN gain, CN loss, and LOH that extend to one of the subtelomeres, but do not cross the centromere, and are $>$10 Mb in size. LST represents the number of chromosomal breakpoints (change in copy number or allelic content) between adjacent regions that are $>$10 Mb but not whole chromosome.

2.5 Evaluation of Chromothripsis Using OGM Data

The chromothripsis status was inferred by visual scoring according to the CNV. The number of switches between copy-number states was counted for each chromosome. Chromosomes containing 10 or more switches within 50 Mb were scored as chromothripsis-positive with high confidence. Chromosomes with 8 to 9, or 6 to 7 switches within 50 Mb were scored as chromothripsis-positive with intermediate and low confidence, respectively [27].

2.6 Evaluation of Intratumoral Heterogeneity (ITH) Using OGM Data

The variant allele frequencies (VAFs) for all SVs within a sample were divided by the maximum VAF, thus normalizing them within the range of 0 to 1. Subsequently, the 0–1 range was partitioned into a series of windows with a width of 0.05, and the proportion of SVs falling within each window was calculated. Finally, the Shannon diversity formula was applied to these proportions in order to compute the ITH value as follows: $\mathrm{ITH} = -\sum(P_i)(\ln P_i)$, where P_i is the proportion of SVs falling within window i.

2.7 Whole Exome Sequencing and Data Processing

For each of the five HBOC-related breast cancer samples, DNA was extracted from paired tumor/normal tissues using the Invitrogen PureLink Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, USA) as recommended by the manufacturer. DNA enrichment and library preparation were carried out using the Agilent SureSelect Human All Exon V6 kit (Agilent Technologies, Santa Clara, CA, USA) according to routine protocols. Libraries were sequenced on the Illumina NovaSeq6000 platform (Illumina, San Diego, CA, USA) with 150-bp paired-end runs.

Quality control of raw reads was performed using fastp (v0.23.1) and Trimmomatic (v0.32) to trim reads with adapters and to remove low-quality reads. Clean reads were aligned to the reference genome using BWA (v0.7.16a), and the aligned reads were then analyzed to identify somatic variants, including SNPs and Indels using GATK (v4.1.6.0). Somatic variants were annotated using ANNOVAR (v2015Mar22) and summarized by R (v4.3.1) package maftools to gain insight into the genomic landscape of each cancer sample.

2.8 RNA Sequencing and Data Processing

Total RNA was extracted from three HBOC-related breast cancer samples (P3, P4, P5) using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) and routine protocols. Transcriptome libraries were made using the Illumina TruSeq RNA sample preparation Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Libraries were sequenced on the Illumina NovaSeq6000 platform (Illumina, San Diego, CA, USA) to generate 150-bp paired-end reads.

Quality control of raw reads was performed using (v0.23.1) and Trimmomatic (v0.32) to trim reads with adapters and to remove low-quality reads. Clean reads were then aligned to the reference genome using Hisat2 (v9.2.1), and the reference-based assembly of transcripts was performed using StringTie (v1.3.3b) to quantify the expression level of exons, transcripts and genes. Gene expression levels [log2(FPKM+1)] in tumor samples were compared to those in normal samples using one-tailed T-test.

2.9 Enrichment of Signaling Pathways

All COSMIC (https://cancer.sanger.ac.uk/cosmic) genes found by WES to contain a mutation underwent KEGG signaling pathway enrichment analysis using the KOBAS (v3.0) as previously reported [28] (http://bioinfo.org/kobas). Similarly, the COSMIC genes found by OGM to contain a SV were analyzed by KEGG enrichment. Cancer- and treatment-related signaling pathways that showed statistical significance was visualized.

2.10 Statistical Analysis

Categorical variables were compared using Fisher's exact test or the chi-square test, while continuous variables were compared using the Wilcoxon rank sum test or T-test. Statistical analysis was performed using SPSS 23 software (IBM SPSS statistics, Chicago, IL, USA). $p < 0.05$ or $p < 0.01$ were considered to represent a statistically significant result, as indicated.

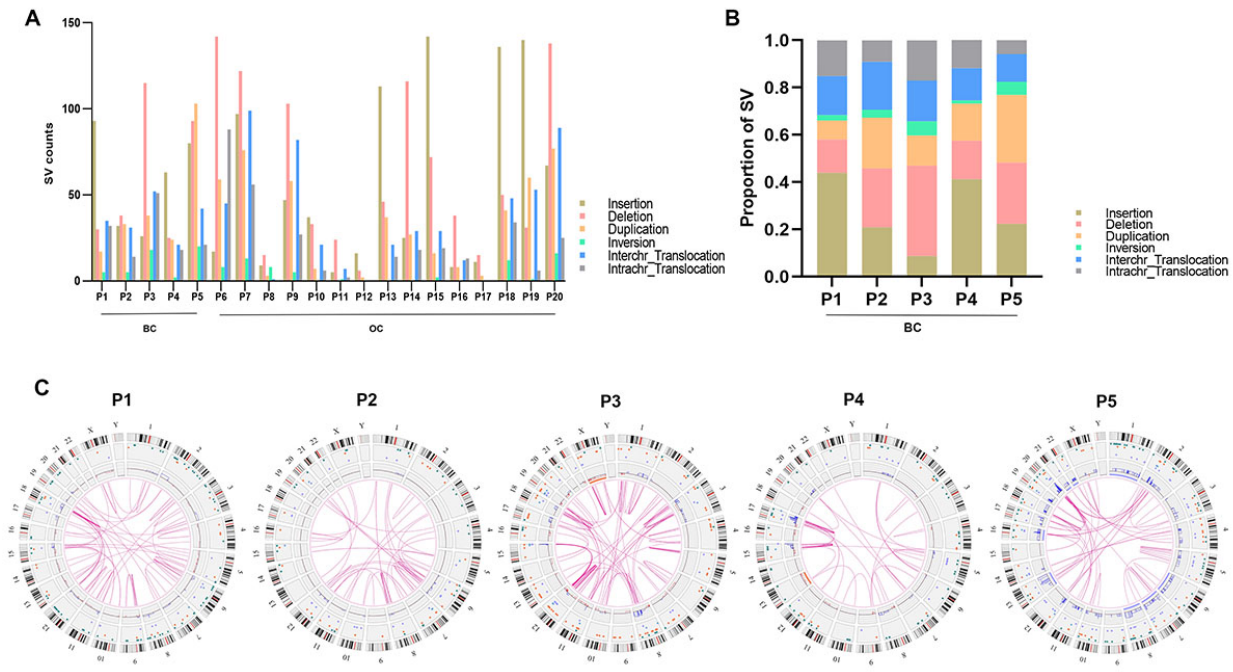

Fig. 1. Structural variations (SVs) detected by optical genome mapping (OGM) in hereditary breast and ovarian cancer (HBOC) tissue samples. (A) SV counts in 20 HBOC tissue samples. (B) SV counts in 5 HBOC-related breast cancer tissue samples. (C) Circle plots showing the SVs and copy number variations (CNVs) in 5 HBOC-related breast cancer tissue samples. BC, breast cancer; OC, ovarian cancer.

3. Results

3.1 Data Quality and SV/CNV Calling of OGM Data

We first evaluated the technical performance of OGM analysis in 20 HBOC-related cancer samples. This gave the following results (median and interquartile range): N50 $\geq$150 kbp of 253.85 kbp (219.38–312 kbp), total DNA ($\geq$150 kbp) of 1221.2 Gbp (363.61–1496.09 Gbp), map rate of 83.35% (51.3–89.8%), effective coverage of 318.97$\times$ (79.95–394.05$\times$), label density of 15.43 (14.15–18.13), PLV of 2.45% (2.14–3.6%), and NLV of 12.25% (8–18.48%). Thus, all OGM raw data was qualified and met the requirement for calling SVs (**Supplementary Table 2**).

A median of 223 SVs per sample were detected in 20 HBOC-related cancer samples using the recommended confidence filter settings with control database ($\leq$1%) filter. These included a median of 42 insertions, 42 deletions, 30 duplications, 3.5 inversions, 30 inter-chromosomal translocations and 18 intra-chromosomal translocations (Fig. 1A). A median of 212 SVs per sample were detected in the 5 HBOC-related breast cancers filtered using the recommend confidence filter settings with control database ($\leq$1%), with a median of 63 insertions, 38 deletions, 33 duplications, 5 inversions, 35 inter-chromosomal translocations and 20 intra-chromosomal translocations (Fig. 1B,C, Table 2). The most frequent SV type was insertion, followed by deletion. SV patterns differed between each sample. The most frequent SV type was insertion in samples P1 and P4, deletion in samples P2 and P3, and duplication in sample P5. Inversion was the least frequent SV type in all samples (Fig. 1B). A median of 107 CNVs per sample were detected, with a median of 61 non-masked gains and 29 non-masked losses (Fig. 1C, Table 2). Therefore, SVs and CNVs were both very common in HBOC-related breast cancer.

3.2 Number of SVs Associated with Ki-67 Expression

We next examined the number of SVs in breast cancer subtypes. TNBC contained 93 insertions in P1 and 80 in P5, whereas luminal B breast cancer showed 32 insertions in P2, 26 in P3, and 63 in P4. TNBC showed a trend for more inversions than luminal B breast cancer. The number of other SV types between the breast cancer subtypes was similar (Fig. 2A). Furthermore, the total number of SVs in Ki-67^{high} breast cancer samples was 212 (P1), 300 (P3), and 359 (P5), whereas in Ki-67^{low} samples it was 153 (P2) and 153 (P4). Therefore, the Ki-67^{high} samples showed a trend for more total SVs than Ki-67^{low} samples. Similar trends were seen for inter-chromosomal translocations (Ki-67^{high}: 35 (P1), 52 (P3), 42 (P5) vs. Ki-67^{low}: 31 (P2), 21 (P4)), as well as for intra-chromosomal translocations (Ki-67^{high}: 32 (P1), 51 (P3), 21 (P5) vs. Ki-67^{low}: 14 (P2), 18 (P4)) (Fig. 2B). We therefore defined breast cancer samples with more SVs and with high Ki-67 expression as SV^{high} HBOC-related breast cancer (n = 3), and those with less SVs and low Ki-67 expression as SV^{low} HBOC-related breast cancer (n = 2).

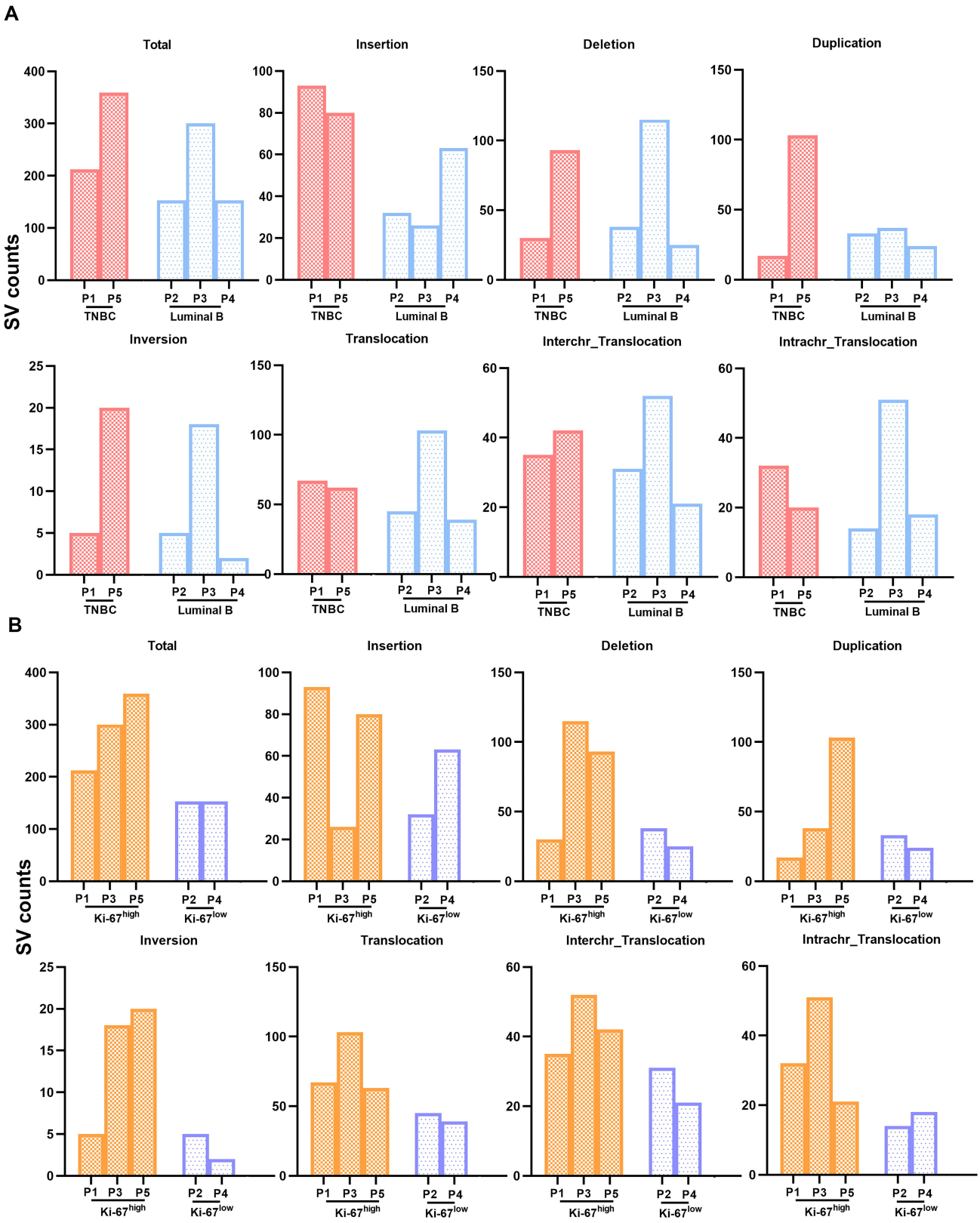

Fig. 2. The relationship between SV counts and luminal subtype/Ki-67 expression in HBOC-related breast cancers, as detected by OGM. (A) SV counts in triple negative and in luminal B breast cancers. (B) SV counts in breast cancers with different levels of Ki-67 expression.

Table 1. Clinicopathologic feature of five breast cancer samples.

ID	Age	Pathlogy	Stage	Lymph node metastasis	Neoadjuvant therapy	Adjuvant systemic therapy	Relapse	ER status	PR status	HER2 status	HER2 FISH	Ki-67	p53	CDK5/6	Luminal subtype	*BRCA1/2* germline mutation	Oher risk genes germline mutation
P1	53	invasive carcinoma	II	Yes	No	Yes	No	<1%	<1%	1+	NA	50%	80%	50%	Triple-negative	*BRCA1*, p.Glu1836fs	NA
P2	35	invasive carcinoma	I	No	Yes	Yes	No	1%	<1%	1+	NA	10%	<1%	10%	Luminal B	*BRCA1*, p.Leu1306fs	NA
P3	44	invasive carcinoma	II	Yes	No	Yes	No	90%	20%	2+	Negative, heterogeneity	55%	5%	<1%	Luminal B	*BRCA2*, p.Arg2520Ter	NA
P4	28	invasive carcinoma	II	No	No	Yes	No	90%	15%	2+	Negative	30%	<1%	<1%	Luminal B	*BRCA1*, p.Gln1281Ter	NA
P5	64	invasive carcinoma	II	No	No	Yes	No	<1%	<1%	0	NA	70%	>90%	40%	Triple-negative	*BRCA1*, p.Asp942fs	NA

ER, estrogen receptor; PR, progesterone receptor; HER2, receptor tyrosine-protein kinase erbB-2; NA, not applicable.

Table 2. SVs and CNVs summary of the five HBOC-related breast cancers in OGM.

SV calls using the recommend confidence filter settings without control filter						
	P1	P2	P3	P4	P5	Total
Insertion	691	610	628	641	639	3209
Deletion	565	619	705	530	640	3059
Duplication	93	116	124	77	163	573
Inversion	89	77	77	58	87	388
Interchr_Translocation	35	31	52	21	42	181
Intrachr_Translocation	32	14	51	18	20	135
Total	1505	1467	1637	1345	1591	7545
SV calls using the recommend confidence filter settings with control ≤1% filter						
	P1	P2	P3	P4	P5	Total
Insertion	93	32	26	63	80	293
Deletion	30	38	115	25	93	301
Duplication	17	33	38	24	103	214
Inversion	5	5	18	2	20	50
Interchr_Translocation	35	31	52	21	42	181
Intrachr_Translocation	32	14	51	18	21	135
Total	212	153	300	153	359	1174
CNV calls (non-masked only) and Aneuploidy (non-masked only)						
	P1	P2	P3	P4	P5	Total
Gain	40	19	61	71	245	436
Loss	29	0	63	36	2	130
Total	69	19	124	107	247	566
Aneuploidy Gain	0	0	0	0	8	8
Aneuploidy Loss	0	0	1	1	0	2
Total	0	0	1	1	8	10

SVs, structural variations; CNVs, copy number variations; HBOC, hereditary breast and ovarian cancer; OGM, optical genome mapping.

Table 3. Fusions identified by OGM and validated by RNA-Seq.

Sample ID	Fusion	LeftBreakpoint	RightBreakpoint	SV type
P3	*C12orf76*::*ZFAT*	chr12:110486168	chr8:135669980	interchr_fusion
P4	*CLTC*::*DHX40*	chr17:57768072	chr17:57676098	dup
P5	*CKMT2-AS1*::*PDE4D*	chr5:80597106	chr5:58489362	intrachr_fusion
P4	*NAV1*::*PKP1*	chr1:201687883	chr1:201282294	dup_split
P5	*EXOC6*::*NSMCE4A*	chr10:94733989	chr10:123727321	intrachr_fusion
P5	*RBM38*::*RAE1*	chr20:55968389	chr20:55929088	dup
P5	*TOM1L2*::*TRPC4AP*	chr17:17875576	chr20:33642834	interchr_fusion
P5	*EIF3L*::*TRIOBP*	chr22:38254747	chr22:38130406	dup

3.3 SV^{high} HBOC-Related Breast Cancers Show Higher HRD Scores

Next, three core indexes for HRD (LOH, TAL, LST) were calculated using OGM data. The median HRD score for the three SV^{high} HBOC-related breast cancers was higher than in the two SV^{low} samples (129 vs. 69). For the SV^{high} and SV^{low} samples, the median LOH score was 14 vs. 4, the median TAI value was 15 vs. 5, and the median LST score was 97 vs. 60, respectively (Fig. 3A). SVs and gene mutations were also detected in DNA damage repair (DDR) genes. SVs were frequently detected in HR genes other than *BRCA1/2* genes, and were present in both the SV^{high} and SV^{low} samples (Fig. 3B). In total, 80% (4/5) of the HBOC-related breast cancer samples carried DDR mutations. The median number of DDR gene mutations (3 vs. 2.5, Fig. 2B) and the median ITH score (2.771 vs. 2.367, Fig. 3C) were both slightly higher in SV^{high} samples compared to SV^{low} samples, although neither reached statistical significance, possibly due to the small sample size. Therefore, a higher number of SVs in HBOC-related breast cancer was associated with higher Ki-67 expression and higher HRD scores.

3.4 SV^{high} HBOC-Related Breast Cancer has More Mutated Genes and Altered Signaling Pathways

WES data analysis was performed to compare genetic mutations between SV^{high} and SV^{low} HBOC-related breast cancer samples. This revealed that missense mutations were the most frequent variant in both groups. The median number of total mutations per sample (1145 vs 946.5) and of missense mutations per sample (906 vs 719.5) were both slightly higher in the SV^{high} group compared to the SV^{low} group (Fig. 4A,B). Furthermore, the median number of mutated genes per sample was higher in the SV^{high} group (642.7 vs. 573, $p = 0.053$) (Fig. 4C). The median number of SV genes per sample was also higher in the SV^{high} group (590 vs. 295.5, Fig. 4D).

We next examined whether the COSMIC genes with SVs and distinct mutations (**Supplementary Tables 3–6**) were enriched in tumor-related signaling pathways (Fig. 4E, **Supplementary Table 7**). The SV^{high} group was more enriched in SV-related signaling pathways than the SV^{low} group (20 vs. 13). Seven signaling pathways were exclusively enriched in the SV^{high} group, namely the p53, VEGF, Jak-STAT, Hippo, TGF-beta, NF-kappa B and endocrine resistance pathways (Fig. 4E). In addition, more mutation-related signaling pathways were enriched in the SV^{high} group (20 vs. 6), including signaling pathways associated with cancer therapy such as the EGFRi resistance pathway, the PD-L1&PD-1 pathway, platinum drug resistance pathway, and endocrine resistance pathway. Therefore, SV^{high} HBOC-related breast cancers exhibit more mutated genes and altered signaling pathways.

3.5 SV^{high} HBOC-Related Breast Cancers Show More Chromothripsis Events

Three regions showed copy number gains in all 5 HBOC-related breast cancer samples, including chromosome (Chr) 1q25-32 gain, Chr 1q41-44 gain and Chr 8q22-q24 gain (harboring *MYC* oncogene) (Fig. 5A). Chr 1q25.3-q44 gain, Chr6p24.3 gain and Chr 8q22.3-q24.3 gain (harboring *MYC* oncogene) were recurrent CNV regions in SV^{low} samples (Fig. 5A). More recurrent CNV regions were detected in SV^{high} samples, including seven CNV gain regions and one CNV loss region: Chr1q23.1-q32.3 gain, Chr1q41-q44 gain, Chr2p25.3-q25.1gain, chr3q22.1-q26.3gain (harboring *PIK3CA* oncogene), chr8q12.1-q24.3 gain (harboring *MYC* oncogene, HRR gene *NBN*), chr17q21.3-q24.3 gain (harboring *RNF43* oncogene, HRR gene *BRIP1*), chr18q21.2-q23 gain (harboring cell proliferation gene *BCL2*), chrX q21.31 loss (Fig. 5A). Therefore, SV^{high} HBOC-related breast cancers exhibited more CNVs involved in tumorigenesis.

We further evaluated chromothripsis events in all 5 HBOC-related breast cancer samples according to CNV visual scoring as previously reported [27]. This revealed that 60% (3/5) of HBOC-related breast cancer samples showed chromothripsis (P3, P4, P5). The higher CNV visual score referred to more severe chromothripsis in tumor tissue (Fig. 5C). The most severe chromothriptic events were seen in P5. The circle plot (Fig. 5B) and CNV plot of the whole genome (Fig. 5D) for P5 revealed abnormal CNV changes and multiple chromosomes affected by chromothripsis. Of note, the tumor suppressor gene *PTEN* was lost in chromothriptic chromosome 10. Transcriptome sequencing showed that *PTEN* gene expression

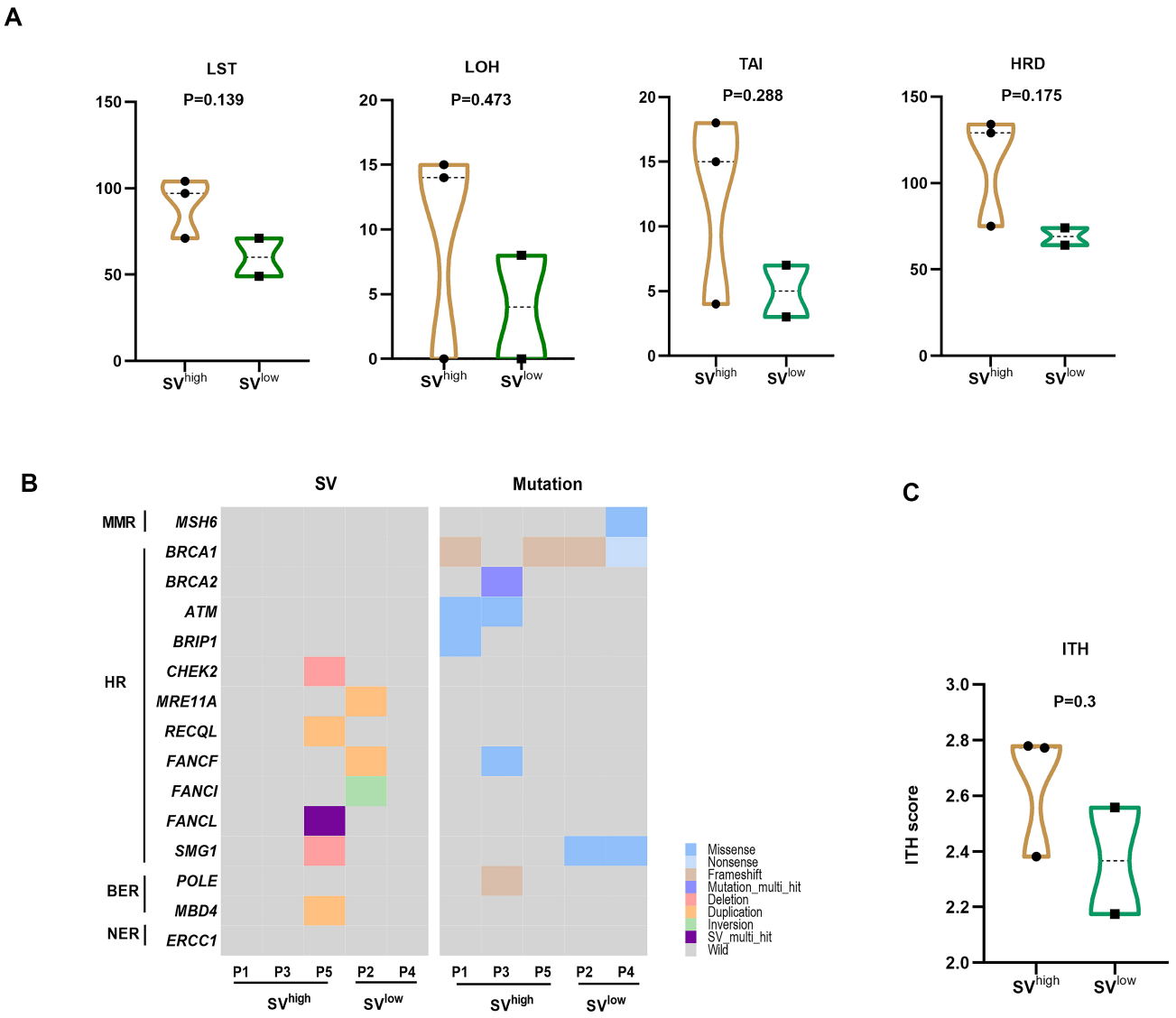

Fig. 3. Associations between SV counts and homologous recombination deficiency (HRD), mutations in DNA damage repair (DDR) genes, and intratumoral heterogeneity (ITH) in HBOC-related breast cancers. (A) HRD core indexes in SV^{high} and SV^{low} groups. (B) SVs and mutations in DNA damage repair (DDR) genes in SV^{high} and SV^{low} groups. (C) ITH in SV^{high} and SV^{low} groups.

was significantly down-regulated in breast cancer tissues at the mRNA level compared to normal adjacent tissues (p = 0.009) (Fig. 5E). The Golgi-associated gene *GOLPH3* was duplicated in chromothriptic chromosome 5. Transcriptome sequencing confirmed that *GOLPH3* expression in breast cancer tissues was significantly up-regulated at the mRNA level compared with normal adjacent tissues (p = 0.014) (Fig. 5F). Both these genes are closely associated with platinum drug resistance, and hence it is interesting to note that patient P5 may show early relapse after platinum-based chemotherapy in the clinic.

3.6 SV^{high} HBOC-Related Breast Cancers Show More Novel Gene Fusions

OGM detected 8 novel gene fusions in HBOC-related breast cancers. These were validated by RNA-Seq (Table 3). 50% fusions were due to interchr_trans or intrachr_trans, 50% fusions were due to duplication. 75% (6/8) fusions were detected in SV^{high} HBOC-related breast cancer samples. Three novel fusions are described here in detail (Fig. 6A–C). *CKMT2-AS1*::*PDE4D* was detected in SV^{high} sample P5 (Fig. 6A) and was formed by the fusion of lncRNA *CKMT2-AS1* with functional gene *PDE4D* as an intrachr_fusion. Transcriptome sequencing showed that expression of *PDE4D* in tumor tissues was significantly increased compared with normal adjacent tissues (p = 0.013). *C12orf76*::*ZFAT* was detected in SV^{high} sample P3 and was formed by an interchr_fusion of the functional genes *C12orf76* and *ZFAT*. The expression of *ZFAT* in tumor tissues was significantly increased compared to normal adjacent tissues (Fig. 6B) (p = 0.0009). *CLTC*::*DHX40* was detected in SV^{low} sample P4 for the first time (Fig. 6C).

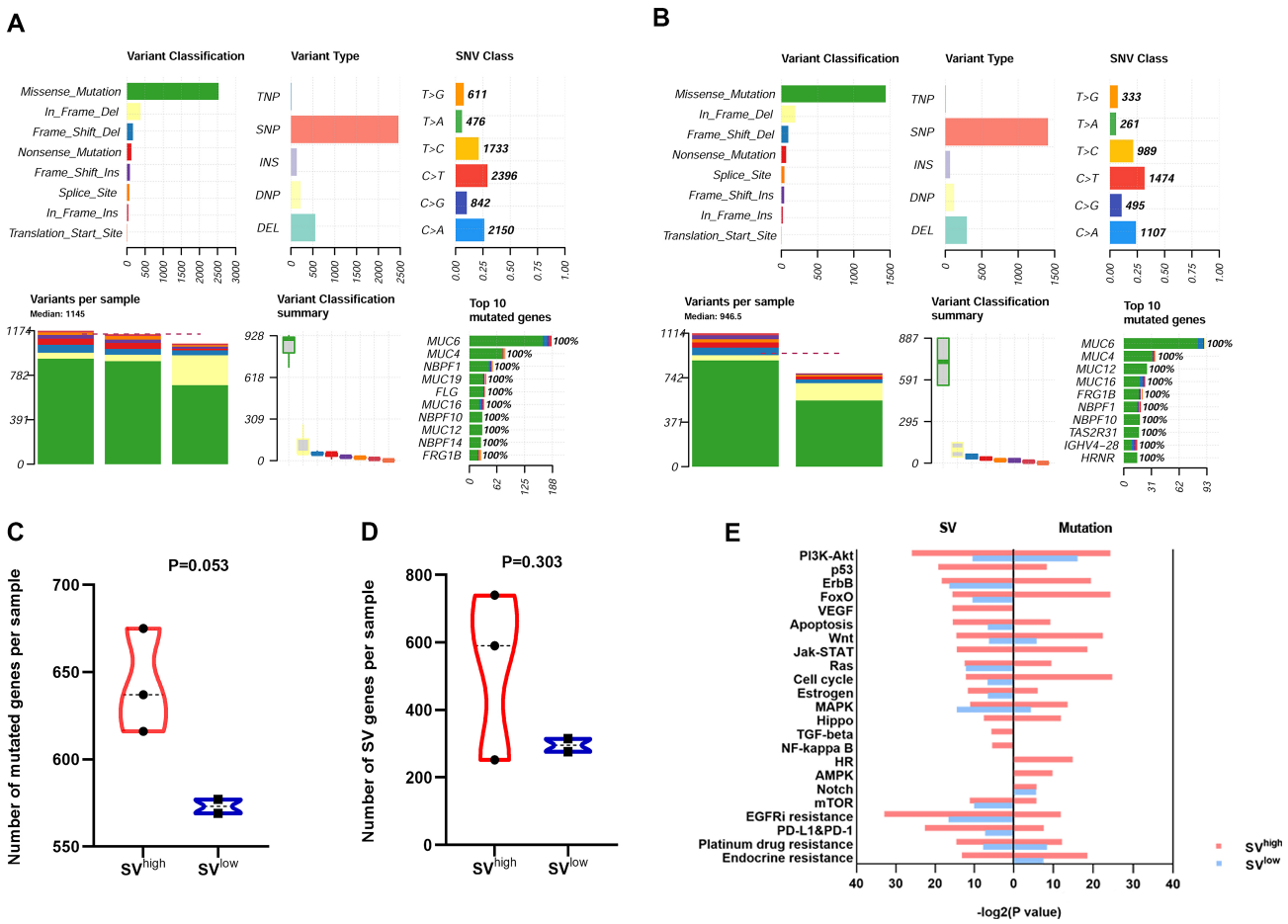

Fig. 4. Comparison of genetic mutations and altered signaling pathways between SV^{high} and SV^{low} HBOC-related breast cancers. (A) Mutations in SV^{high} samples. (B) Mutations in SV^{low} samples. (C) Number of mutated genes per sample in SV^{high} and SV^{low} groups. (D) Number of SV genes per sample in the SV^{high} and SV^{low} groups. (E) Altered signaling pathways at the SV and mutation level in the SV^{high} and SV^{low} groups.

The fusion was caused by duplication, and both *CLTC* and *DHX40* were significantly up-regulated in tumor tissues compared with normal adjacent tissues (p = 0.00036, p = 0.0014).

4. Discussion

The aim of this study was to assess the utility of OGM for detecting SVs in HBOC-related breast cancers. OGM is an advanced technology that uses high-resolution imaging techniques to map the structure and organization of the genome, thus providing valuable insights into SVs. It can identify SVs that are often missed by traditional sequencing methods, allowing a more comprehensive understanding of the genome. Furthermore, OGM is a label-free technique, making it a non-destructive and cost-effective tool compared to other sequencing technologies. It also allows direct visualization and analysis of genomic features, thereby eliminating the need for time-consuming DNA amplification or labeling. OGM has demonstrated its utility in cancer research and in helping to advance precision medicine [21,24]. In the present research, the landscape of SVs in HBOC-related breast cancer samples was comprehensively investigated by OGM to determine how the intratumor heterogeneity of SVs impacts clinicopathological features and the genetic alteration profile. OGM was also able to detect gene fusions with high accuracy, and several novel gene fusions were identified and validated.

Our results indicate that SVs are very common in HBOC-related breast cancers. Manual inspection of all the SVs revealed a redundancy rate of 7.8%, which was mainly focused on deletion. However, this SV type showed no difference in trend between groups and did not affect the results for the association between number of SVs and luminal type or Ki-67 expression, or the group based on the SV count. Hence, no further analysis was performed and the data is shown in **Supplementary Table 8** for reference (**Supplementary Table 8**). The number of SVs was found to be associated with Ki-67 expression, which is a nuclear marker of cell proliferation and an important indicator of tumor cell activity. Breast cancers that express high levels of Ki-67 are associated with faster tumor growth and worse clinical outcome [29]. It has been suggested that SVs could

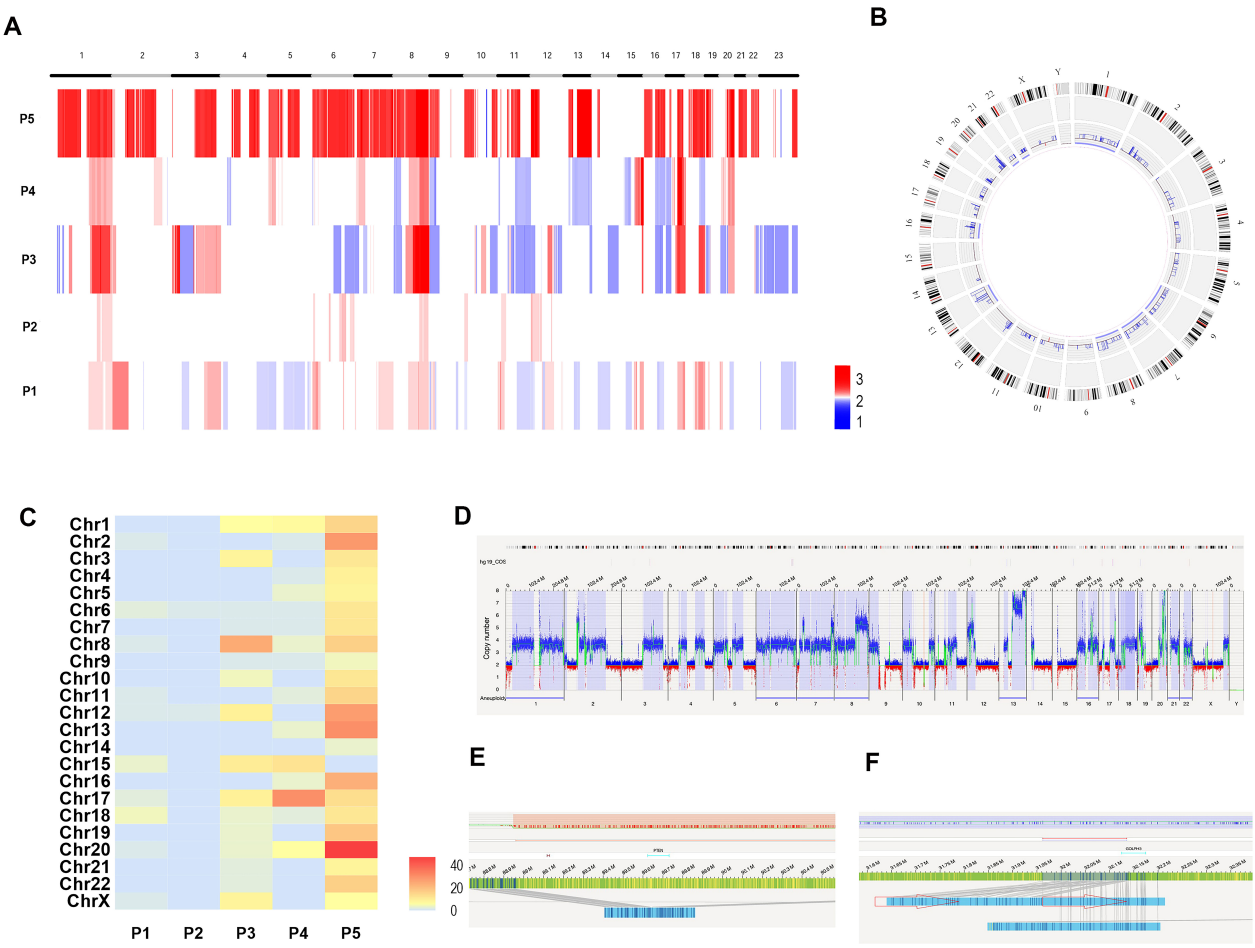

Fig. 5. Copy Number Variations (CNVs) and chromothripsis detected by OGM in HBOC-related breast cancers. (A) The CNV landscape of the 5 samples. (B) The circle plot shows the abnormal CNV changes in P5. (C) The heat map shows the chromothripsis visual scores for the 5 samples. (D) The whole-genome CNV profiles illustrates the abnormal CNV changes in P5. (E) Schematic showing loss of the tumor suppressor gene *PTEN* in chromothriptic chromosome 10 in P5. (F) Schematic showing duplication of the Golgi-associated protein gene *GOLPH3* in chromothriptic chromosome 5 in P5.

provide additional prognostic information for patients [30]. HBOC-related breast cancers with SV^{high} could therefore imply poor prognosis for *BRCA1/2* mutated patients and indicate heterogeneity for tumor invasiveness, drug resistance, and recurrence.

BRCA1/2 mutations are the main cause of HRD in the clinic. The HRD scores in the 5 HBOC-related breast cancer patients examined in this study ranged from 64 to 134, demonstrating they all had homologous recombination deficiency [31]. We also observed that SV^{high} patients showed a trend for higher HRD scores and more DDR mutations, implying these HBOC-related breast cancer patients could be sensitive to PARP inhibitors [32]. Moreover, SV^{high} patients showed a trend for higher ITH scores, which is a biomarker for immunotherapy [33]. Given that PD-L1 expression and the PD-1 checkpoint pathway were enriched in SV^{high} patients, it is also reasonable to conclude that such patients may benefit from PARP inhibitors combined with immunotherapy.

We also explored the genetic alteration profile of HBOC-related breast cancers. SV^{high} HBOC-related breast cancers showed an overall trend for more genetic mutations and more mutated genes. The open access BioPortal/TCGA databases do not provide SV data, but CNV and SNV data can be obtained. For validation purposes, we evaluated the correlation between CNV load and SNV in TCGA. Positive correlations were found between the CNV load and SNV count (Spearman correlation analysis, $p = 5.459672 \times 10^{-45}$, rho = 0.432) (**Supplementary Fig. 1A**), as well as between CNV load and the number of mutated genes (Spearman correlation analysis, $p = 1.485089 \times 10^{-44}$, rho = 0.43) (**Supplementary Fig. 1B**). These results suggest that samples with a high CNV load have high SNV counts and a high number of mutated genes. Furthermore, it has been reported that more SVs were associated with more mutations in ovarian and prostate cancers [34]. Our results are consistent with the above trends reported in the literature on the correlation between SVs and mutation. In

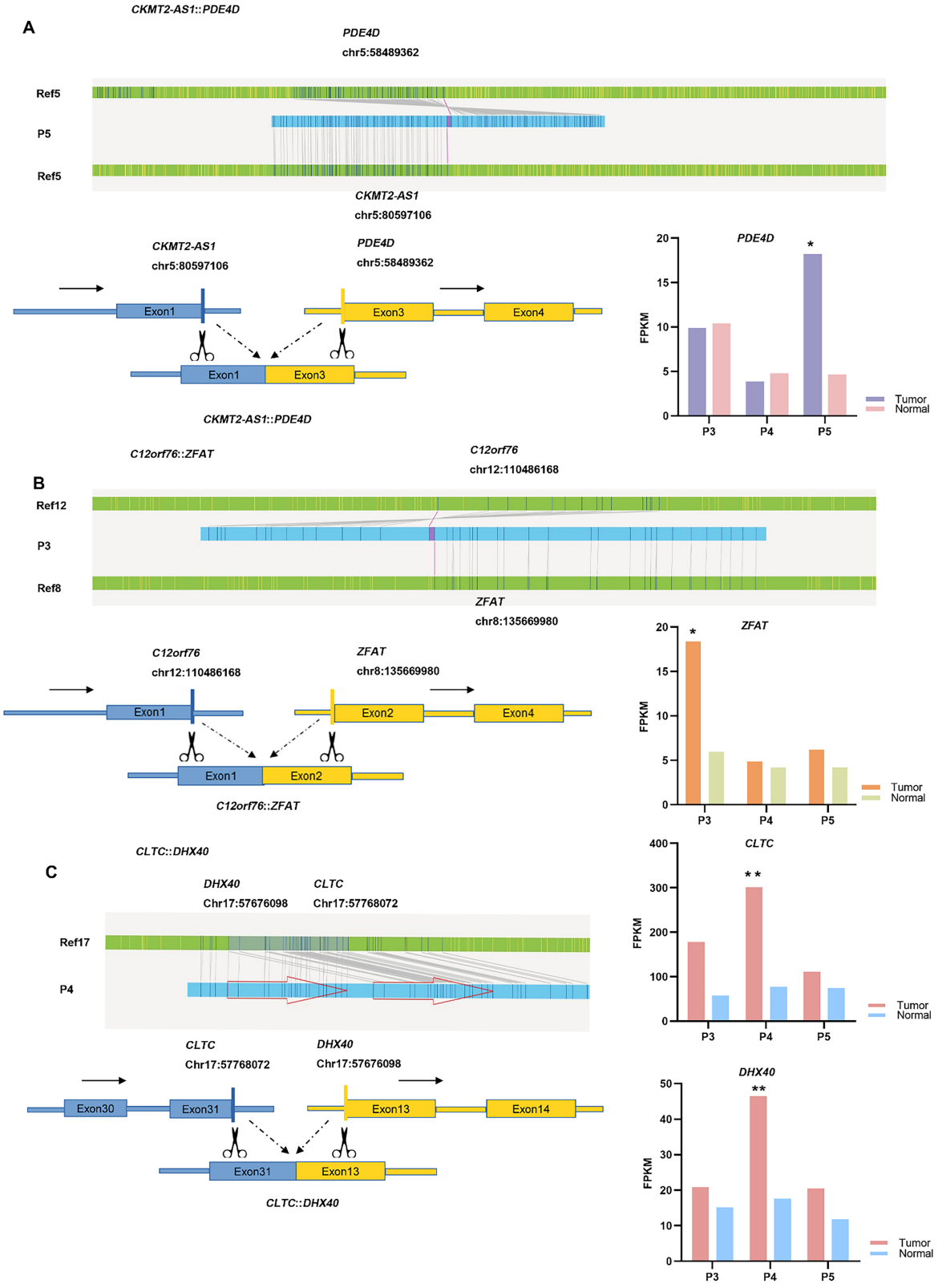

Fig. 6. Identification of novel gene fusions using OGM. (A) The top and bottom left panel show a detailed schematic of the *CKMT2-AS1*::*PDE4D* fusion. The bottom right panel shows the expression of *PDE4D* in tumor tissue compared with normal adjacent tissue. (B) The top and bottom left panel show a detailed schematic of the *C12orf76*::*ZFAT* fusion. The bottom right panel shows the expression of *ZFAT* in tumor tissue compared to normal adjacent tissue. (C) The left panel shows a detailed schematic of the *CLTC*::*DHX40* fusion. The right panel shows the expression of *CLTC* and *DHX40* in tumor tissue compared to normal adjacent tissue. Note: $*p < 0.05$, $**p < 0.01$.

addition, the mutated genes were mainly enriched in signaling pathways involved in tumorigenesis, tumor progression and resistance to therapies, including signaling pathways for EGFRi resistance, platinum drug resistance and endocrine resistance.

SV^{high} HBOC-related breast cancers also showed more CNV, of which gains in Chr 1q, 3q, 8q ,17q and losses in ChrX have previously been associated with reduced survival [35–38]. If >80% of a chromosome by length has either lower or higher CNV than baseline, a whole-chromosome aneuploidy event would be identified. This was the case for sample P3, which was not only called as having multiple CNVs, but also showed aneuploidy loss in chromosome X. Comprehensive assessment of multiple types of genomic alterations might expand our current understanding of precision medicine in breast cancer. Chromothripsis is a type of genome instability characterized by one or several chromosomes being affected by tens to hundreds of clustered DNA rearrangements [27,39–41]. In the current study, chromothripsis was found in 60% (3/5) of HBOC-related breast cancer samples in which multiple chromosomes were affected. It was reported earlier that chromothripsis events are pervasive in cancer, with a frequency of >60% in a cohort of metastatic breast cancers and 25% in a cohort comprised of predominantly luminal breast cancers [39]. In our study, chromothripsis was mainly associated with deleterious *BRCA1/2* germline mutations. This is consistent with a previous report that pathogenic germline variants of essential checkpoint or DNA repair factors may facilitate chromothripsis [27]. Chromothripsis is also thought to promote or even cause tumor development by simultaneously inactivating tumor-suppressor genes, amplifying oncogenes, and forming oncogenic fusions [39,42–45]. In the present study, more chromothripsis events were identified in SV^{high} HBOC-related breast cancers. We observed that the tumor suppressor gene *PTEN* was lost and down-regulated, while the Golgi-associated protein gene *GOLPH3* was amplified and up-regulated in sample P5, which were associated with platinum drug resistance [46]. Patient P5 showed the most chromothriptic chromosomes amongst the SV^{high} HBOC-related breast cancers.

Eight novel gene fusions were identified in this study and validated by RNA-Seq. We accessed the RNA-Seq reads, confirmed that the fusions were all in-frame, and identified that the breakpoints were located at the edges of exons without leading to premature termination codon (PTC), premature stop, and nonsense mediated decay. The IGV screenshots of the 8 fusions are included in the supplementary figures (**Supplementary Figs. 2–9**) to show the fusions identified by RNA-Seq. Two of these fusions were intrachr_trans, which can be called as either true translocation events or duplications/deletion events with Bionano software. Intra-chromosomal fusion breakpoints typically involve regions located at least 5 Mbp away from each other on the same chromosome. Meanwhile, duplication >5 Mb, deletion >5 Mb, and intra-chromosomal fusions with a reference distance between fusion points of <1 Mb are also called as intra-chromosomal fusion events. The two intra-chromosomal fusion events found in P5 in our study were confirmed not to be caused by duplication or deletion events, but rather by an extra copy of a segment fused to another segment.

The newly identified gene fusion *CKMT2-AS1*::*PDE4D* in SV^{high} HBOC-related breast cancer was formed by fusion of the lncRNA *CKMT2-AS1* with functional gene *PDE4D*. *CKMT2-AS1* is an autophagy-related lncRNA previously reported to be a prognostic biomarker in papillary renal cell carcinoma [47]. Transcriptome sequencing revealed that *CKMT2-AS1*::*PDE4D* was associated with increased *PDE4D* expression in tumor tissues. This may predict worse survival in tamoxifen-treated breast cancer patients, since *PDE4D* is known to play a pivotal role in acquired tamoxifen resistance via the blocking of cAMP/ER stress/p38-JNK signaling and apoptosis [48].

Using the OGM technique, we identified several gene fusions that interfere with gene transcription and expression. Some were previously reported, including *RBM38*::*RAE1* [49], *NAV1*::*PKP1* [50] and *EIF3L*::*TRIOBP* [51]. A novel gene fusion in SV^{high} HBOC-related breast cancer found in the present study was *C12orf76*::*ZFAT*. This fusion resulted from interchr_fusion of the two functional genes *C12orf76* and *ZFAT*. Little is known about the clinical significance and biological function of *C12orf76*. *ZFAT* was originally identified as a susceptibility gene for autoimmune thyroid disease [52], regulates apoptosis in human T-cell acute lymphocytic leukaemia (T-ALL) [53,54], and participates in the development of ovarian cancer [55]. We found that *C12orf76*::*ZFAT* fusion correlated with increased expression of *ZFAT*, which might play a critical role in breast cancer.

The OGM technique also identified the novel gene fusion *CLTC*::*DHX40* in SV^{low} HBOC-related breast cancer. This was caused by duplication of *CLTC* and *DHX40*, with both genes also showing significant up-regulation in tumor tissues. *CLTC* encodes for a major subunit of clathrin, a multimeric protein on cytoplasmic organelles, as well as being a recurrent fusion partner for the ALK tyrosine kinase gene in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor [56,57]. *DHX40* encodes for a member of the DExH/D box family of ATP-dependent RNA helicases that play an essential role in RNA metabolism [58]. Hyper/hypomethylated *DHX40* was identified in platinum-resistant ovarian cancers [59], indicating that *CLTC*::*DHX40* could also be involved in breast cancer development and drug resistance. Although multiple fusions were identified here by integrating OGM data with

RNA-seq data, these should be confirmed in larger tumor cohorts and their functional roles clarified by further investigations.

This study has several limitations. Firstly, the relatively small number of cases meant there was low statistical power to assess the associations between SVs and other clinical or genomic features. Secondly, the observed chromothripsis and novel gene fusion events should be validated in larger cohorts and their potential biological functions investigated in HBOC-related breast cancers.

5. Conclusions

In conclusion, OGM is a promising tool for the detection of SVs and CNVs in HBOC-related breast cancer. This method can efficiently characterize and quantify chromothripsis events and novel gene fusions in cancer tissues. SV^{high} HBOC-related breast cancers were associated with unfavorable clinicopathological features, with these genetic alterations having potential predictive and therapeutic significance in the clinic.

Availability of Data and Materials

The data presented in this study are available on request from the corresponding author.

Author Contributions

YNC performed the experiment and wrote the manuscript, LD performed the experiment, LH tested the software, DCB and YZ interpretated the data and reviewed the article, JTL and XJG collected the samples and the clinal-pathological information, HLX performed bioinformatics analysis and wrote the manuscript, JPY conceived and designed this study. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

Informed consent was obtained from all subjects involved in the study. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Tianjin Medical University (No. Ek2018050).

Acknowledgment

Not applicable.

Funding

This work was supported by the National Science and Technology Support Program of China (Grant No. 2018ZX09201015), National Natural Science Foundation of China (Grant No. 82072588, 82002601, 81872143, 81702280), and Projects of Science and Technology of Tianjin (Grant No. 18JCQNJC82700), Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-009A).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2901002.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021; 71: 209–249.

[2] Curtis C, Shah SP, Chin S, Turashvili G, Rueda OM, Dunning MJ, *et al*. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012; 486: 346–352.

[3] Biancolella M, Testa B, Baghernajad Salehi L, D'Apice MR, Novelli G. Genetics and Genomics of Breast Cancer: update and translational perspectives. Seminars in Cancer Biology. 2021; 72: 27–35.

[4] Li Y, Roberts ND, Wala JA, Shapira O, Schumacher SE, Kumar K, *et al*. Author Correction: Patterns of somatic structural variation in human cancer genomes. Nature. 2023; 614: E38.

[5] Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, *et al*. An integrated map of structural variation in 2,504 human genomes. Nature. 2015; 526: 75–81.

[6] Hu T, Li J, Long M, Wu J, Zhang Z, Xie F, *et al*. Detection of Structural Variations and Fusion Genes in Breast Cancer Samples Using Third-Generation Sequencing. Frontiers in Cell and Developmental Biology. 2022; 10: 854640.

[7] Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nature Reviews Cancer. 2007; 7: 233–245.

[8] Nattestad M, Goodwin S, Ng K, Baslan T, Sedlazeck FJ, Rescheneder P, *et al*. Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. Genome Research. 2018; 28: 1126–1135.

[9] Bayard Q, Cordier P, Peneau C, Imbeaud S, Hirsch TZ, Renault V, *et al*. Structure, Dynamics, and Impact of Replication Stress-Induced Structural Variants in Hepatocellular Carcinoma. Cancer research. 2022; 82: 1470–1481.

[10] Hanahan D, Weinberg R. Hallmarks of Cancer: the next Generation. Cell. 2011; 144: 646–674.

[11] Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, *et al*. Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. Nature. 2007; 448: 561–566.

[12] Wang X, Luan Y, Yue F. EagleC: a deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. Science Advances. 2022; 8: eabn9215

[13] Hasty P, Montagna C. Chromosomal rearrangements in cancer. Molecular & Cellular Oncology. 2014; 1: e29904.

[14] Jabbour E, Kantarjian H. Chronic myeloid leukemia: 2022 up-

date on diagnosis, therapy, and monitoring. American Journal of Hematology. 2022; 97: 1236–1256.
[15] Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, *et al*. Integrative detection and analysis of structural variation in cancer genomes. Nature Genetics. 2018; 50: 1388–1398.
[16] Aganezov S, Goodwin S, Sherman RM, Sedlazeck FJ, Arun G, Bhatia S, *et al*. Comprehensive analysis of structural variants in breast cancer genomes using single-molecule sequencing. Genome Research. 2020; 30: 1258–1273.
[17] Yoshida R. Hereditary breast and ovarian cancer (HBOC): review of its molecular characteristics, screening, treatment, and prognosis. Breast Cancer. 2021; 28: 1167–1180.
[18] Yamauchi H, Takei J. Management of hereditary breast and ovarian cancer. International Journal of Clinical Oncology. 2018; 23: 45–51.
[19] Hodgson A, Turashvili G. Pathology of Hereditary Breast and Ovarian Cancer. Frontiers in Oncology. 2020; 10: 531790.
[20] Lal A, Ramazzotti D, Weng Z, Liu K, Ford JM, Sidow A. Comprehensive genomic characterization of breast tumors with BRCA1 and BRCA2 mutations. BMC Medical Genomics. 2019; 12: 84.
[21] Gao H, Xu H, Wang C, Cui L, Huang X, Li W, *et al*. Optical Genome Mapping for Comprehensive Assessment of Chromosomal Aberrations and Discovery of New Fusion Genes in Pediatric B-Acute Lymphoblastic Leukemia. Cancers (Basel). 2022; 15: 35.
[22] Lei Y, Meng Y, Guo X, Ning K, Bian Y, Li L, *et al*. Overview of structural variation calling: Simulation, identification, and visualization. Computers in Biology and Medicine. 2022; 145: 105534.
[23] Sobczak K, Krzyzosiak WJ. Structural Determinants of BRCA1 Translational Regulation. Journal of Biological Chemistry. 2002; 277: 17349–17358.
[24] Xia L, Wang Z, Wu X, Zeng T, Luo W, Hu X, *et al*. Multiplatform discovery and regulatory function analysis of structural variations in non-small cell lung carcinoma. Cell Reports. 2021; 36: 109660.
[25] Dong L, Zhang H, Zhang H, Ye Y, Cheng Y, Li L, *et al*. The mutation landscape of multiple cancer predisposition genes in Chinese familial/hereditary breast cancer families. Cancer Biology & Medicine. 2022; 19: 850–870.
[26] Dong L, Wu N, Wang S, Cheng Y, Han L, Zhao J, *et al*. Detection of novel germline mutations in six breast cancer predisposition genes by targeted next-generation sequencing. Human Mutation. 2018; 39: 1442–1455.
[27] Voronina N, Wong JKL, Hubschmann D, Hlevnjak M, Uhrig S, Heilig CE, *et al.* The landscape of chromothripsis across adult cancer types. Nature Communications. 2020; 11: 2320.
[28] Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, *et al*. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. Nucleic Acids Research. 2021; 49: W317–W325.
[29] Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, *et al*. Ki67 Index, her2 Status, and Prognosis of Patients with Luminal B Breast Cancer. JNCI: Journal of the National Cancer Institute. 2009; 101: 736–750.
[30] Yang H, Garcia-Manero G, Sasaki K, Montalban-Bravo G, Tang Z, Wei Y, *et al*. High-resolution structural variant profiling of myelodysplastic syndromes by optical genome mapping uncovers cryptic aberrations of prognostic and therapeutic significance. Leukemia. 2022; 36: 2306–2316.
[31] Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, *et al*. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clinical Cancer Research. 2016; 22: 3764–3773.
[32] Pilie PG, Gay CM, Byers LA, O'Connor MJ, Yap TA. PARP Inhibitors: Extending Benefit Beyond BRCA-Mutant Cancers. Clinical Cancer Research : an official journal of the American Association for Cancer Research. 2019; 25: 3759–3771.
[33] Vitale I, Shema E, Loi S, Galluzzi L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. Nature Medicine. 2021; 27: 212–224.
[34] Cosenza MR, Rodriguez-Martin B, Korbel JO. Structural Variation in Cancer: Role, Prevalence, and Mechanisms. Annual Review of Genomics and Human Genetics. 2022; 23: 123–152.
[35] Pang JB, Savas P, Fellowes AP, Mir Arnau G, Kader T, Vedururu R, *et al*. Breast ductal carcinoma *in situ* carry mutational driver events representative of invasive breast cancer. Modern Pathology. 2017; 30: 952–963.
[36] Iddawela M, Rueda O, Eremin J, Eremin O, Cowley J, Earl HM, *et al*. Integrative analysis of copy number and gene expression in breast cancer using formalin-fixed paraffin-embedded core biopsy tissue: a feasibility study. BMC Genomics. 2017; 18: 526.
[37] Chin S, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M, *et al*. Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. Oncogene. 2007; 26: 1959–1970.
[38] Andre F, Job B, Dessen P, Tordai A, Michiels S, Liedtke C, *et al*. Molecular Characterization of Breast Cancer with High-Resolution Oligonucleotide Comparative Genomic Hybridization Array. Clinical Cancer Research. 2009; 15: 441–451.
[39] Bolkestein M, Wong JKL, Thewes V, Körber V, Hlevnjak M, Elgaafary S, *et al*. Chromothripsis in Human Breast Cancer. Cancer Research. 2020; 80: 4918–4931.
[40] Cortes-Ciriano I, Lee JJ, Xi R, Jain D, Jung YL, Yang L, *et al*. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Nature Genetics. 2020; 52: 331–341.
[41] Korbel J, Campbell P. Criteria for Inference of Chromothripsis in Cancer Genomes. Cell. 2013; 152: 1226–1236.
[42] Rausch T, Jones DW, Zapatka M, Stütz A, Zichner T, Weischenfeldt J, *et al*. Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. Cell. 2012; 148: 59–71.
[43] Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, *et al*. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. Cell. 2011; 144: 27–40.
[44] Rode A, Maass KK, Willmund KV, Lichter P, Ernst A. Chromothripsis in cancer cells: an update. International Journal of Cancer. 2016; 138: 2322–2333.
[45] Li Y, Schwab C, Ryan S, Papaemmanuil E, Robinson HM, Jacobs P, *et al*. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. Nature. 2014; 508: 98–102.
[46] Huang D, Savage SR, Calinawan AP, Lin C, Zhang B, Wang P, *et al*. A highly annotated database of genes associated with platinum resistance in cancer. Oncogene. 2021; 40: 6395–6405.
[47] Pang Y, Wang Y, Zhou X, Ni Z, Chen W, Liu Y, *et al*. Cuproptosis-Related LncRNA-Based Prediction of the Prognosis and Immunotherapy Response in Papillary Renal Cell Carcinoma. International Journal of Molecular Sciences. 2023; 24: 1464.
[48] Mishra RR, Belder N, Ansari SA, Kayhan M, Bal H, Raza U, *et al*. Reactivation of cAMP Pathway by PDE4D Inhibition Represents a Novel Druggable Axis for Overcoming Tamoxifen Resistance in ER-positive Breast Cancer. Clinical Cancer Research. 2018; 24: 1987–2001.
[49] Hu X, Wang Q, Tang M, Barthel F, Amin S, Yoshihara K, *et al*. TumorFusions: an integrative resource for cancer-associated

transcript fusions. Nucleic Acids Research. 2018; 46: D1144–D1149.

[50] Ghandi M, Huang FW, Jane-Valbuena J, Kryukov GV, Lo CC, McDonald ER, 3rd, *et al*. Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature. 2019; 569: 503–508.

[51] Dehghannasiri R, Freeman DE, Jordanski M, Hsieh GL, Damljanovic A, Lehnert E, *et al*. Improved detection of gene fusions by applying statistical methods reveals oncogenic RNA cancer drivers. Proceedings of the National Academy of Sciences. 2019; 116: 15524–15533.

[52] Shirasawa S, Harada H, Furugaki K, Akamizu T, Ishikawa N, Ito K, *et al*. SNPs in the promoter of a B cell-specific antisense transcript, SAS-ZFAT, determine susceptibility to autoimmune thyroid disease. Human Molecular Genetics. 2004; 13: 2221–2231.

[53] Tsunoda T, Shirasawa S. Roles of ZFAT in haematopoiesis, angiogenesis and cancer development. Anticancer Research. 2013; 33: 2833–2837.

[54] Fujimoto T, Doi K, Koyanagi M, Tsunoda T, Takashima Y, Yoshida Y, *et al*. ZFAT is an antiapoptotic molecule and critical for cell survival in MOLT-4 cells. FEBS Letters. 2009; 583: 568–572.

[55] Ramakrishna M, Williams LH, Boyle SE, Bearfoot JL, Sridhar A, Speed TP, *et al*. Identification of candidate growth promoting genes in ovarian cancer through integrated copy number and expression analysis. PLoS ONE. 2010; 5: e9983.

[56] Argani P, Lui MY, Couturier J, Bouvier R, Fournet J, Ladanyi M. A novel CLTC-TFE3 gene fusion in pediatric renal adenocarcinoma with t(X;17)(p11.2;q23). Oncogene. 2003; 22: 5374–5378.

[57] Georgantzoglou N, Green D, Winnick KN, Sumegi J, Charville GW, Bridge JA, *et al*. Molecular investigation of ALK-rearranged epithelioid fibrous histiocytomas identifies CLTC as a novel fusion partner and evidence of fusion-independent transcription activation. Genes, Chromosomes & Cancer. 2022; 61: 471–480.

[58] Paine I, Posey JE, Grochowski CM, Jhangiani SN, Rosenheck S, Kleyner R, *et al*. Paralog Studies Augment Gene Discovery: DDX and DHX Genes. American Journal of Human Genetics. 2019; 105: 302–316.

[59] Hua T, Kang S, Li XF, Tian YJ, Li Y. DNA methylome profiling identifies novel methylated genes in epithelial ovarian cancer patients with platinum resistance. The Journal of Obstetrics and Gynaecology Research. 2021; 47: 1031–1039.