Breast Cancer Research and Treatment (2019) 177:145–153
https://doi.org/10.1007/s10549-019-05286-w

EPIDEMIOLOGY

Assessment of *PARP4* as a candidate breast cancer susceptibility gene

Aldo Prawira[1] · Prabhakaran Munusamy[1] · Jimin Yuan[1] · Claire Hian Tzer Chan[1] · Geok Ling Koh[1] · Timothy Wai Ho Shuen[2] · Jiancheng Hu[1,3] · Yoon Sim Yap[2] · Min Han Tan[2,4,5] · Peter Ang[2,6] · Ann Siew Gek Lee[1,7,8]

Received: 29 October 2018 / Accepted: 17 May 2019 / Published online: 22 May 2019

Abstract

Purpose *PARP4* has been proposed as a candidate breast cancer susceptibility gene. However, its function and involvement in breast carcinogenesis is unclear. We sought to determine the variant frequency of *PARP4* in *BRCA*-negative women referred for genetic testing from Singapore and to perform functional analyses of *PARP4*.

Methods Next-generation sequencing of *PARP4* was conducted for 198 *BRCA*-negative cases from Singapore. Three independent case–control association analyses of *PARP4* were performed for (1) our Singaporean cohort, (2) three dbGaP datasets, and (3) cases from TCGA, with controls from the Exome Aggregation Consortium (ExAC). *PARP4* knockout cells were generated utilizing the CRISPR-Cas9 approach in MDA-MB-231 (breast cancer) and MCF10A (normal breast) cell lines, and colony formation, cell proliferation, and migration assays carried out.

Results Candidate variants in *PARP4* were identified in 5.5% (11/198) of our Singapore cohort. Case–control association studies for our cases and the dbGaP datasets showed no significant association. However, a significant association was observed for *PARP4* variants when comparing 988 breast cancer cases from the TCGA provisional data and 53,105 controls from ExAC (ALL) (OR 0.249, 95% CI 0.139–0.414, $P = 2.86 \times 10^{-11}$). *PARP4* knockout did not affect the clonogenicity, proliferation rate, and migration of normal breast cells, but appeared to decrease the proliferation rate and clonogenicity of breast cancer cells.

Conclusions Taken together, our results do not support that *PARP4* functions as a cancer susceptibility gene. This study highlights the importance of performing functional analyses for candidate cancer predisposition genes.

Keywords *PARP4* · Breast cancer · Germline variants · Cancer predisposition

Introduction

Germline mutations in *BRCA1* and *BRCA2* are associated with hereditary breast and/or ovarian cancer, with women harboring pathogenic variants having an increased cumulative risk of developing breast and/or ovarian cancer

Aldo Prawira and Prabhakaran Munusamy have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-019-05286-w) contains supplementary material, which is available to authorized users.

✉ Ann Siew Gek Lee
dmslsg@nccs.com.sg

1 Division of Cellular and Molecular Research, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore, Singapore

2 Department of Medical Oncology, National Cancer Centre, Singapore, Singapore

3 Cancer and Stem Cell Biology Programme, Duke-NUS Medical School, Singapore, Singapore

4 Institute of Bioengineering and Nanotechnology, Singapore, Singapore

5 Lucence Diagnostics Pte Ltd, Singapore, Singapore

6 Oncocare Cancer Centre, Gleneagles Medical Centre, Singapore, Singapore

7 Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

8 Office of Clinical and Academic Faculty Affairs, Duke-NUS Medical School, Singapore, Singapore

of between 55 and 70% [1]. Nevertheless, mutations in *BRCA1*/2 are estimated to occur in less than 25% of hereditary cases [2], implying that there could exist other susceptibility genes that contribute to the development of these cancers.

Rapid advancements in next-generation sequencing have allowed for the identification of a myriad of variants in predisposition genes such as *TP53*, *PTEN*, *PALB2*, and *CDH1* [3]. Recently, *PARP4* (Poly-ADP-ribose Polymerase 4) has been described as a novel candidate gene by several sequencing studies on breast and thyroid cancer [4], pancreatic cancer [5], and hepatocellular carcinoma [6]. However, these studies involved small cohorts of patients, ranging from 3 to 24 patients and lacked detailed functional studies.

The *PARP4* gene is located at chromosome 13q11 and comprises of 34 exons, encoding 1724 amino acids [7]. Also known as vPARP, the PARP4 protein is one of the largest members of the PARP family [8] and exhibits poly-ADP ribosyltransferase activity that catalyses the post-translational modification of target proteins through the addition of ADP-ribose [9]. PARP4 has been predicted to be involved in the base excision repair pathway due to the presence of the BRCT domain commonly found in other DNA repair pathway proteins such as PARP1 and XRCC1, although a previous study has shown that PARP4 did not localize to the nucleus following UV treatment [10]. PARP4 associates with the Major Vault Protein (MVP), Telomerase-associated Protein (TEP1), and a small untranslated RNA (vRNA) to form a cytoplasmic ribonucleoprotein complex with a hollow barrel-like structure known as 'vaults' [11]. The precise functions of vaults are largely not understood, although the upregulation of these complexes has been associated with multidrug resistance in various cancers [12, 13].

In this study, we performed targeted next-generation sequencing of the *PARP4* gene on 198 *BRCA1*/2 mutation-negative patients referred for risk assessment by genetic testing. Furthermore, case–control association analyses using publicly available databases and in vitro functional characterization using CRISPR-Cas9 gene editing were also performed to investigate the role of *PARP4* in breast cancer and normal breast cells.

Materials and methods

Patient characteristics and sample collection

Cases were referred for genetic counselling at the National Cancer Centre Singapore, based on previously described criteria [3]: (1) an existing family history of breast cancer in first- and/or second-degree relatives; (2) personal history of breast and/or ovarian cancer or bilateral breast cancer; (3) early onset breast cancer ($\leq$40 years old) [3]. Demographic and clinicopathological characteristics of these cases (n = 198) are shown in Table 1. Written informed consent was obtained from all participants, and the study was approved by the SingHealth Centralised Institutional Review Board. Peripheral blood samples were collected in EDTA tubes and genomic DNA subsequently extracted using an optimized in-house protocol [3].

Next-generation sequencing

Of the 198 unrelated *BRCA1/2*-mutation-negative cases, 166 cases were subjected to targeted sequencing for the *PARP4* gene using the KAPA HyperPrep Kit (Roche, Basel, Switzerland), according to the manufacturer's protocol. Briefly, 100 ng of genomic DNA was fragmented using a Covaris S220 ultrasonicator and subsequently ligated to the TruSeq Dual Index Duplex Mixed Adapter (IDT, Coralville, IA, USA). Following DNA library preparation, amplicons were electrophoresed on the Agilent Bioanalyzer 2100 and the DNA1000 kit (Agilent, Santa Clara, CA, USA) was used to determine fragment size and quantity. DNA libraries were then target-captured for the *PARP4* gene using the xGen Lockdown Panels kit (IDT) by overnight hybridization at 65 °C. Paired-end (2 × 150 bp) sequencing was then performed using the Illumina MiSeq platform. Additionally, DNA from 32 *BRCA1/2*-negative cases were sent to a sequencing provider for whole exome sequencing (Personalis, Inc, Menlo Park, CA, USA), while DNA from 46 Singaporean controls were whole exome sequenced by another sequencing provider (Novogene Co., Ltd).

Bioinformatic data analysis

All samples that were either targeted sequenced or whole exome sequenced were analyzed using the same pipeline, as follows: Raw reads were aligned to the human reference genome (hg19 version) using the Burrows-Wheeler alignment (BWA) tool. Aligned reads were sorted and processed for PCR duplicate removal using the SAMtool and Picard tool (v1.74). Insertion or deletion (indel) realignment and base score recalibration were subsequently carried out using the functionalities of the GATK toolkit (v3.4-46). Finally, variant discovery of SNPs and indels were performed with the GATK HaplotypeCaller module (v3.4-46).

Variants present in the *PARP4* chromosomal region (chr13: 24995069–25086948) were extracted from the VCF of our whole exome sequenced samples and public databases (dbGaP, TCGA, ExAC) for further analysis using the tabix tool (version: 1.3.2).

To identify the impact of variants on gene function and to prioritize the variants, functional annotation of the variants was performed using the ANNOVAR pipeline (http://annovar.openbioinformatics.org/). Exonic and intronic variants

Table 1 Characteristics of patients involved in this study

Characteristics	Patients ($n = 198$)	
	No. of patients	Percentage
Sex		
Female	196	99.0
Male	2	1.0
Race/Ethnicity		
Chinese	167	84.3
Malay	15	7.6
Indian	5	2.5
Others	11	5.6
Personal history of breast and/or ovarian cancer		
Unilateral breast cancer	164	82.8
Bilateral breast cancer	15	7.6
Ovarian cancer	10	5.0
No personal history of breast and/or ovarian cancer, but with a family history[a]	9	4.6
Patients with a family history of breast and/or ovarian cancer	82	41.4
Patients aged ≤ 40 years	131	66.2
Age at first diagnosis (years)		
Mean	39	
Median	37	
Range (unknown for 1 patient)	19–67	

[a]Details of the inclusion criteria are in the methods section and have been previously published [3]

within ± 50 bp of the exon boundaries were retained, and synonymous variants were discarded. Subsequently, variants with minor allele frequency (MAF) of ≤ 0.01 (1%) available from the ExAC East Asian population databases were selected. The variants were then manually visualized using the Integrative Genome Viewer (IGV, v2.4.6). Five functional prediction algorithms (SIFT [14], PolyPhen-2 [15], MutationTaster [16], Mutation Assessor [17], and CADD [18]) were utilized to assess the deleteriousness of each identified variant. Variants were considered as candidate variants if they were classified as deleterious by at least one tool.

Cell culture

MDA-MB-231 (ATCC HTB-26, hereby referred to as MB-231) breast epithelial carcinoma and MCF10A (ATCC CRL-10317) normal breast epithelial cells were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). MB-231 cells were cultured in DMEM medium supplemented with 10% fetal bovine serum and 1 mM sodium pyruvate. MCF10A cells were grown in DMEM/F12 medium supplemented with 5% horse serum, 20 ng/ml epidermal growth factor (EGF), 0.5 µg/ml hydrocortisone, 100 ng/ml cholera toxin, and 10 µg/ml human recombinant insulin. All cells were cultured according to standard aseptic technique and maintained in a 37 °C humidified incubator with 5% CO_2.

CRISPR knockout

LentiCRISPRv2 (Addgene #52961, Cambridge, MA, USA) plasmid was initially digested with *BsmBI* and subsequently purified following agarose gel electrophoresis using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Oligonucleotides encoding single guide RNA (sgRNA) targeting exon 2 of the *PARP4* gene (sgPARP4; sequence: 5′-GAACTG CTCGCCTTCTGATC-3′) was subsequently cloned into the digested plasmid. For control vector, lentiCRISPRv2 with sgRNA targeting enhanced green fluorescence protein (EGFP) (sgEGFP; sequence: 5′-GGGCGAGGAGCT GTTCACCG-3′) was constructed. All plasmids were then Sanger sequenced to confirm correct insertion prior to lentiviral generation.

To produce lentivirus, a third-generation lentiviral packaging system was used by co-transfecting PARP4-sgRNA- or EGFP-sgRNA-containing plasmid with pMD2.G (Addgene #12259), pRSV-Rev (Addgene #12253), and pMDLg/pRRE (Addgene #12251) into subconfluent lentiX-293T cells (Clontech, Mountain View, CA, USA) using TransIT-2020 transfection reagent (Mirus Bio, Madison, WI, USA). Lentivirus was collected after 72 h of incubation and used to transduce MCF10A or MB-231 in the presence of 10 µg/ml protamine sulfate overnight. Cells were then subjected to selection pressure with 4 µg/ml puromycin (Sigma-Aldrich,

St. Louis, MO, USA) for 2 weeks and maintained in normal growth medium thereafter. Clonal selection was performed to obtain monoclonal CRISPR knockout cells, and three clones were randomly selected for in vitro assays to account for possible off-target effects and/or random lentiviral insertional mutagenesis.

Sanger sequencing

Sanger sequencing was performed by subjecting 5 ng of genomic DNA to polymerase chain reaction (PCR) amplification using HotStarTaq DNA polymerase (Qiagen). PCR products were then analyzed on agarose gels and purified using Exo/SAP (Thermo Fisher, Waltham, MA, USA). Sanger sequencing was carried out using the BigDye Terminator v3.1 sequencing kit (Applied Biosystems, Foster City, CA, USA) on a 3130xl Genetic Analyser sequencer (Applied Biosystems). Electropherograms were analyzed using SeqMan Pro v.12 (DNASTAR, Madison, WI, USA) by aligning to a reference sequence (NC_000013.11). Primers used for Sanger sequencing are as follows: F: 5′-GTGCATTCGTAG CCATATCC-3′; R: 5′-CCTCCCCCTAGTTAAAAGCAC-3′.

Western blotting

Whole protein was extracted from CRISPR-modified cell lines using RIPA buffer with EDTA and protease inhibitor (Thermo Fisher). Equal amounts of total protein were electrophoresed on a 4–20% SDS-PAGE gel (Bio-Rad, Hercules, CA, USA) and subsequently transferred onto a polyvinylidene difluoride membrane. After blocking with 5% skim milk in TBST for 60 min, membranes were then incubated with antibodies against *PARP4* (1:5000, Abcam #ab133745, Cambridge, UK), or α-tubulin (1:5000, CST #3873, Danvers, MA, USA) in 5% BSA in TBST at 4 °C overnight. Membranes were then washed 3 times with TBST and incubated with horseradish peroxidase-conjugated secondary antibody against mouse (1:1000, Thermo Fisher #PA1-28748) or rabbit (1:1000, CST #7074) for 60 min at room temperature. Immunoblots were subsequently developed with the enhanced chemiluminescence system (Thermo Fisher) according to the manufacturer's protocol.

Colony formation, cell proliferation, and cell migration assays

For the colony formation assay, 500 cells were seeded and incubated for 12 (MB-231) or 8 (MCF10A) days. Cells were then fixed and stained with 25% methanol and 0.5% crystal violet solution for 1 h. Colonies with at least 50 cells were subsequently counted.

To assess cell proliferation, *PARP4*-null or control cells were seeded in five replicates, at an initial density of 2000 cells (MB-231) or 500 cells (MCF10A) in a 96-well plate and allowed to adhere overnight. Cell proliferation was measured using the CellTiter 96 AQueous One Solution Cell Proliferation (MTS) assay (Promega, Madison, WI, USA) according to the manufacturer's protocol. The absorbance of formazan product was measured every 24 h for 4 days at 490 nm.

Cell migration was performed using the Boyden chamber assay method. Firstly, 5×10^4 *PARP4*-null or control cells were seeded in Transwell Inserts with 8 µm pore (Corning Inc, Corning, NY, USA) containing 300 µl of serum-free media. Cells were then allowed to settle for 15 min before placing the insert into 24 well plates containing 500 µl of full growth medium. Cells were incubated for 18 h at 37 °C and subsequently fixed and stained with 25% methanol and 0.5% crystal violet solution for 30 min. Insert membranes were then excised, mounted onto a slide, and migrated cells were visualized using ×10 magnification and quantified manually using the QuPath image processing software (v. 0.1.2) [19].

Statistical analyses

Comparison of the age of diagnosis was performed using the Mann–Whitney *U* test. Clinico-pathological associations were investigated in *PARP4* carriers and non-carriers using the two-tailed Fisher's exact test. Association analyses between breast cancer cases (the Singapore cohort; three dbGaP datasets; cases from TCGA) and control datasets (from ExAC and from our Singaporean population) were calculated using the two-tailed Fisher's exact test. All in vitro experiments were statistically analyzed using the Student's *t* test.

Results

Clinicopathological characteristics of our study population

This study involved 198 *BRCA1/2*-negative cases referred for risk assessment of hereditary breast and/or ovarian cancer. The majority of the study population are females (99%), where a large proportion were of Chinese ancestry (84%) (Table 1). The age of first cancer diagnosis ranged from 19 to 67 years old, with a mean and median age of 39 and 37 years, respectively. Eighty-two patients (41.4%) presented with a family history of breast and/or ovarian cancer, of whom 28 had early-onset breast and/or ovarian cancer (≤ 40 years). Overall, 131 (66.2%) of our cases were aged ≤ 40 years.

Candidate variants identified in *PARP4*

Eight rare missense variants were identified in 11 subjects (~5.5%), all of whom had breast cancer (Table 2). The age range of these patients is 22–55 years old, with a mean age of 38.9 years. All variants were predicted to be deleterious by at least one or more in silico tool (Supplementary Table 1). Importantly, the C-scaled CADD score of these variants was > 10 (range 10.63–31) which indicates high deleteriousness and that they are in the top 10% deleterious variants of the human genome. Furthermore, the estimated conservation score using SiPhy tool for all the variants was found to be positive (range 5.751–15.593) implying the sites in which the variants were found to be highly conserved.

The variant p.Q20R is located in exon 2 and is within the *BRCA1* N-terminus (BRCT) domain (Supplementary Fig. 1). Two other variants, p.G681E and p.A703T identified in exons 16 and 17, respectively, are located in the vault protein inter-alpha trypsin (VIT) domain. Additionally, variant p.A750T was found to be located in close proximity with the VIT domain, whilst variants p.L868R and p.P1049L were situated near to the Von-Willebrand Factor Type A (VWA) domain. Two other variants, p.R1332C and p.Q1634L, were found close to the C-terminus of the PARP4 protein, where the MVP docking site is located [20]. These suggest that the changes in amino acids likely impact structural integrity, hence impairing the function of the *PARP4* protein and its interaction with other proteins.

Table 2 Clinical features of *PARP4* carriers

Patient ID	Ethnicity	Age at diagnosis (years)	Breast cancer histology	Family history	ER	PR	HER2	Nucleotide change	Amino acid change
YP12	Chinese	40	IDC	Paternal aunt, breast cancer	+	+	−	c.4901A>T	p.Q1634L
YP34	Chinese	35	IDC + DCIS	NA.; early-onset breast cancer	+	+	−	c.4901A>T	p.Q1634L
130	Chinese	48	IDC + ILC + DCIS + LCIS	Unknown family history; patient has bilateral breast cancer	+	+	−	c.3994C>T	p.R1332C
YP10	Chinese	33	IDC + DCIS	NA.; early-onset breast cancer	+	−	+	c.3146C>T	p.P1049L
170	Others	38	IDC + DCIS	NA.; early-onset breast cancer	−	−	−	c.2248G>A	p.A750T
37	Indian	37	IDC	NA.; early-onset breast cancer	U	U	U	c.2248G>A	p.A750T
66	Chinese	55	IDC	Daughter Ovarian Cancer (25 years), died at 28	+	+	+	c.2107G>A	p.A703T
YP48	Chinese	35	IDC + DCIS	NA.; early-onset breast cancer	+	+	+	c.2107G>A	p.A703T
161	Chinese	22	IDC + DCIS	NA.; early-onset breast cancer	+	+	−	c.2042G>A	p.G681E
155	Indian	53	Breast mucinous carcinoma + DCIS (Male Br Ca)	No family history	+	+	−	c.59A>G	p.Q20R
53	Malay	32	IDC	Mother, bilateral breast cancer (42 years) and ovarian cancer (52 years); Maternal Grandmother, breast cancer (35 years), died at 93	−	−	−	c.2603T>G	p.L868R

IDC invasive ductal carcinoma, *DCIS* ductal carcinoma in situ, *ILC* invasive lobular carcinoma, *LCIS* lobular carcinoma in situ, *NA* not applicable, *ER* oestrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor 2, *U* unknown

Lack of association of *PARP4* variants with clinicopathological parameters

No difference was observed in the median age of diagnosis between *PARP4* carriers and non-carriers (37 years vs. 37 years, $p=0.972$). Further comparison of clinicopathological features showed that there was no significant difference between oestrogen receptor (ER) positive- (88.9% vs. 67.1%, $p=0.275$), progesterone receptor (PR) positive- (77.8% vs. 56.1%, $p=0.303$), and human epidermal growth factor receptor 2 (HER2) positive- (22.2% vs. 35.1%, $p=0.719$) statuses between *PARP4* carriers and non-carriers. In addition, there was also no significant increase in the frequency of grade 3 breast cancer in *PARP4*-positive versus -negative patients (57.1% vs. 55.3%, $p>0.999$). These results indicate the lack of association of *PARP4* variants with clinicopathological features of breast cancer.

Case–control association analyses

Case–control association analyses of *PARP4* rare variants (MAF < 1%) in breast cancer cases from our cohort versus Singaporean controls (OR 0.619, 95% CI 0.172–2.797, $p=0.493$) and East Asian controls from the ExAC database (OR 1.443, 95% CI 0.693–2.716, $p=0.260$) did not show any significant difference in variant frequency (Table 3). Similarly, *PARP4* variant frequency for three germline DNA breast cancer datasets obtained from the dbGaP database (Accession numbers phs001050.v1.p1, phs000601.v1.p1, and phs000822.v1.p1) compared with ExAC controls from all populations (ALL) did not show any significant difference in cases (4.04% (24/594)) versus controls (5.81% (3088/53,105); OR 0.682, 95% CI 0.432–1.027, $p=0.076$) (Table 3). Interestingly, a significant association was observed for *PARP4* variants when comparing 988 breast cancer cases from the TCGA provisional data and 53,105 healthy controls from ExAC (ALL) with controls having a higher proportion of *PARP4* variant carriers than cases (1.52% (15/988) in cases versus 5.81% (3088/53,105) in controls, OR 0.249, 95% CI 0.139–0.414, $p=2.86\times10^{-11}$) (Table 3).

CRISPR-Cas9 knockout of *PARP4* in MB-231 and MCF10A cells

PARP4-deficient cell lines were generated to further decipher its role in normal and breast cancer cells. Six breast cancer cell lines (MCF7, SKBR3, Hs578T, T47D, MB-231, and MB-468) and 2 non-tumorigenic breast cell lines (MCF10A and MCF12A) were screened for *PARP4* protein expression (Fig. 1a). MB-231 and MCF10A were subsequently chosen based on their higher *PARP4* expression level relative to other the other cell lines and subjected to CRISPR-Cas9 gene editing using lentivirus containing sgPARP4 (Fig. 1b). The expression of *PARP4* in parental cells, control cells transduced with sgEGFP, and three randomly selected clones are shown in Fig. 1c, where *PARP4* protein expression was abolished in knockout cells. Furthermore, sequencing analyses were performed, indicating the presence of indels across all knockout cells, but not for control cells (Fig. 1d).

PARP4-deficient breast cancer cells show decreased colony formation and proliferation

In MB-231 cells, all three knockout clones showed a significantly decreased colony number relative to control cells ranging from a decrease of 14.87–32.18% (Fig. 2a). Similarly, a significant reduction in proliferation was observed for all 3 MB-231 knockout clones compared with control cells (Fig. 2b). One *PARP4*-null MB-231 clone showed an increase in migratory capability; however, this could be due to off-target effect as this phenotype was not observed in other clones (Fig. 2c).

In contrast, only one MCF10A clone showed a significant reduction in colony number. However, this trend was not observed in other clones suggesting that overall, *PARP4* may not affect MCF10A clonogenicity (Fig. 2d). No significant

Table 3 Case-control association analyses for *PARP4* carriers

	Breast cancer cases				Healthy controls			Odds ratio (95% CI)	*p* value
	PARP4 carriers	Wild type	%		*PARP4* carriers	Wild type	%		
Our patient cohort	11	187	5.5	Singaporean controls	4	42	8.7	0.619 (0.172–2.797)	0.493
Our patient cohort[a]	11	187	5.5	ExAC (East Asian)	154	3779	3.9	1.443 (0.693–2.716)	0.260
dbGaP cohort[b]	24	570	4.04	ExAC (ALL population)	3088	50,017	5.81	0.682 (0.432–1.027)	0.076
TCGA cohort[b]	15	973	1.52	ExAC (ALL population)	3088	50,017	5.81	0.249 (0.139–0.414)	2.863×10^{-11}

[a]Breast cancer cases from our patient cohort were compared against the East Asian healthy control population from the ExAC dataset

[b]Breast cancer cases from the dbGaP and TCGA cohorts were compared against the general healthy control population from the ExAC dataset

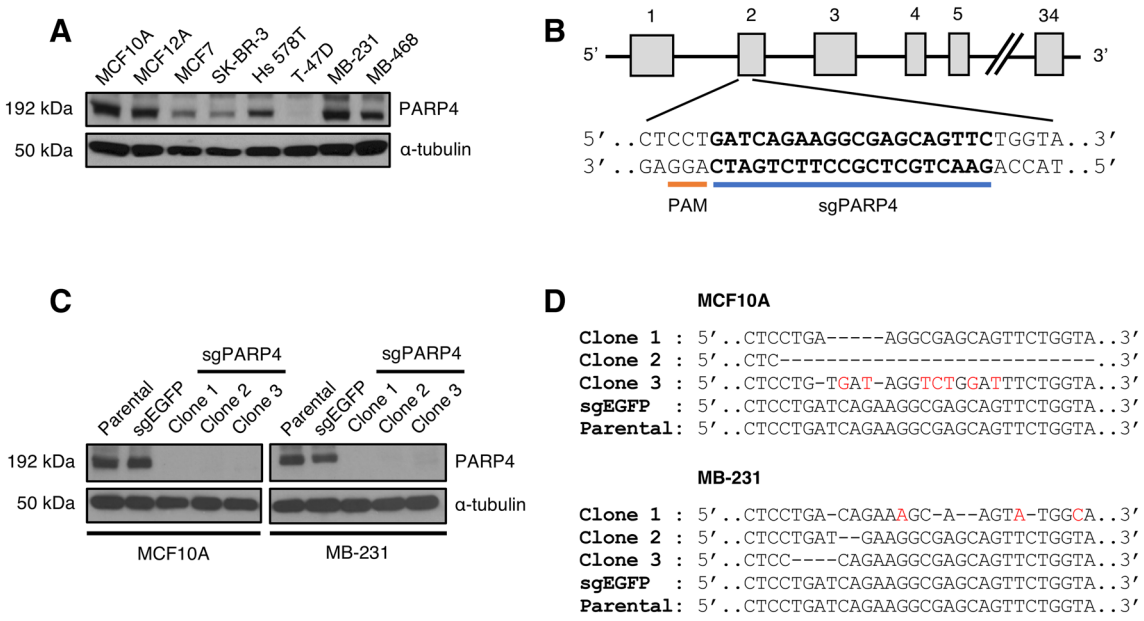

Fig. 1 Generation of PARP4 knockout MCF10A and MB-231 cell lines. **a** Western blotting of PARP4 expression in breast cancer and normal breast cells. **b** Guide RNA (gRNA) targeting exon 2 of the PARP4 gene. PAM sequence is highlighted in red. **c** Western blots showing expression of PARP4 in parental cells, mock (sgEGFP) knockout cells, and three clones of randomly selected PARP4 knockout cells. **d** Sequencing analysis of the knockout and control cells. Hyphens (-) denote deletions, whilst single base alterations are indicated in red

change in growth (Fig. 2e) and migration (Fig. 2f) was observed in *PARP4*-null MCF10A cells when compared with its control. Taken together, these indicate that *PARP4* plays a role in cell survival and growth of cancer cells (MB-231), but not in that of normal cells (MCF10A).

Discussion

Our studies on the case–control association and functional analyses of *PARP4* did not show strong evidence for *PARP4* as a candidate predisposition gene in breast cancer. Overall, *PARP4* variants were found in 5.5% of our study population. A study by Chen et al. has reported a higher frequency of *PARP4* variant (9/88; 10.2%) amongst hepatitis B-induced hepatocellular carcinoma patients [6], while another study focusing on primary co-occurring thyroid and breast cancer involving 14 patients showed that variants in *PARP4* were observed in 43%, with the majority of variant carriers being Caucasians [4].

In this present study, case–control analysis of our Singaporean cohort showed no significant association of *PARP4* with breast cancer risk. In addition, further examination utilizing germline sequencing data of 594 cases from three dbGaP datasets did not show any evidence of increased breast cancer risk. Interestingly, case–control association analyses on a TCGA dataset showed a significant association for *PARP4* variants ($p = 2.86 \times 10^{-11}$) with a relatively higher proportion of *PARP4* carriers in controls as compared to cases (Table 3), thus providing evidence against the role of *PARP4* as a breast cancer susceptibility gene. However, it should be noted that this dataset utilized somatic mutation data from tumor tissue, as access to germline datasets from TCGA is controlled. Therefore, it may not truly represent the genetic landscape of germline cases.

Our in vitro data suggest that the loss of *PARP4* did not increase proliferation in MB-231 or MCF10A cells. In fact, *PARP4*-deficient MB-231 had a decreased proliferation rate compared with its respective control, with no observable change in cell migration or cell morphology in both cell lines, consistent with a previous study which observed that knockdown of *PARP4* did not affect HeLa cell viability and morphology [21]. Furthermore, a study on colon and lung cancer showed that *parp4*$^{-/-}$ mice did not show any noticeable change in phenotype and remained healthy. However, following exposure to the carcinogen dimethylhydrazine, there was an increase in colon tumor incidence, multiplicity, and decreased tumor latency, suggesting the possible role of *PARP4* as a tumor suppressor [22]. Interestingly, the same result was not observed in a lung cancer mice model after urethane injection, raising the possibility that *PARP4*'s tumor suppressive role could be tissue-dependent [22], hence warranting similar studies on breast cancer. Another study examining the role of *parp4* in telomerase function showed

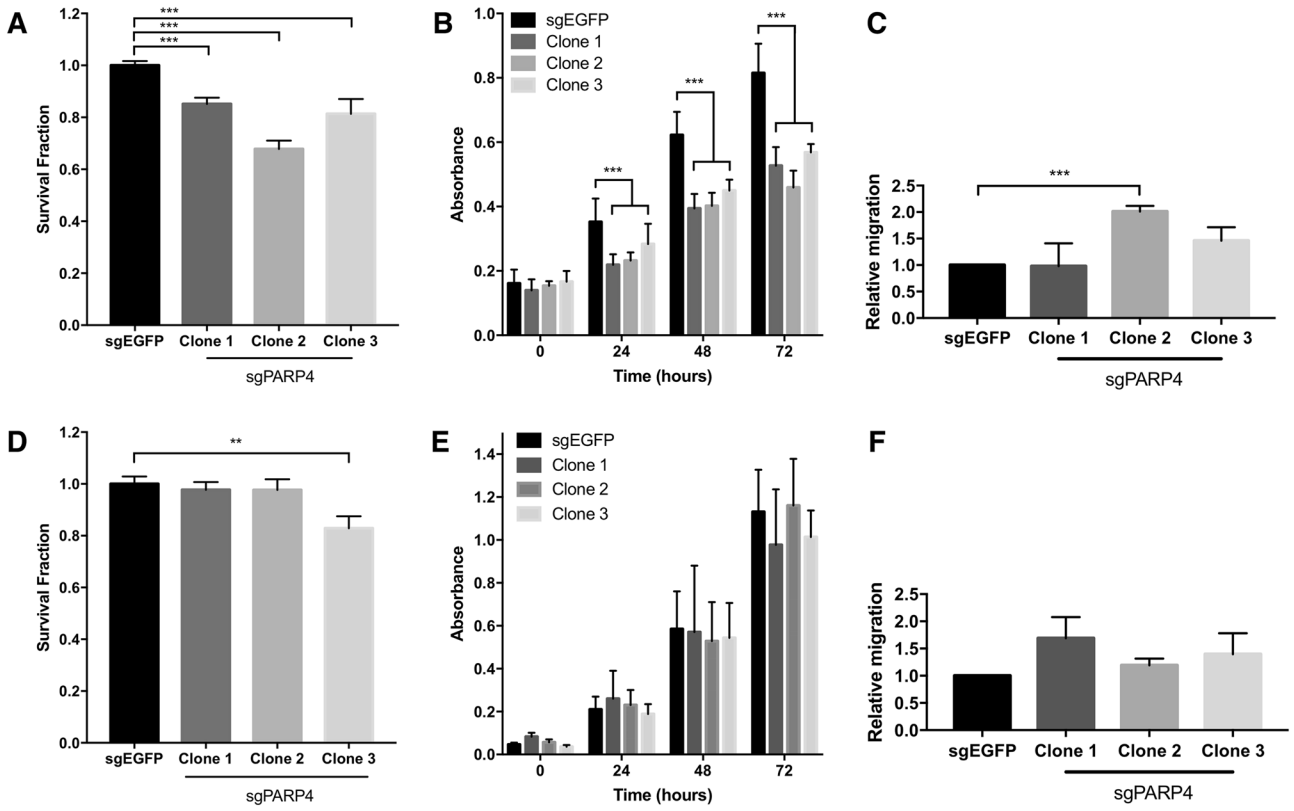

Fig. 2 Colony formation, proliferation and migration assays of PARP4 knockout versus control cells. Relative colony fraction (**a**), proliferation (**b**), and migration (**c**) of PARP4 knockout MB-231 cells compared with mock control. Relative colony fraction (**d**), proliferation (**e**), and migration (**f**) of *PARP4* knockout MCF10A cells compared with control. Data from three independent experiments were collated and subjected to Student's t test. Graphs show the mean value across all readings. Error bars in colony formation assays represent standard error of the mean (SEM). $^{**}p \leq 0.01$, $^{***}p \leq 0.001$. Error bars in proliferation assays represent standard deviation (SD)

that knockout mice did not demonstrate any significant changes in their vault assembly and telomerase activity and length, while confirming that the loss of *parp4* in mice did not alter their health and fertility [23]. In vitro knockdown of *PARP4* in one breast cancer cell line, HCC1143, suggested that *PARP4* may act as a tumor suppressor in contrast to our data [4]. We speculate that phenotypic changes observed between studies could be due to differences in pathological and molecular characteristics between cell types [24, 25]. Together, these suggest that the loss of *PARP4* by itself may not be sufficient to drive cellular transformation; however, it might be involved in breast cancer cell growth. Further studies are warranted to decipher the molecular pathways affected by *PARP*4 in breast cancer cells.

A limitation of this study is that we were unable to obtain DNA samples from family members to perform segregation analysis. Furthermore, although we have performed a functional study of *PARP4*, it is still unclear whether specific variants could cause phenotypic changes through mechanisms similar to those of oncogenes [26]. It is therefore imperative to study the effect of individual variants in *PARP4* to determine their pathogenicity.

In conclusion, results from the functional analyses performed in this study do not provide evidence for *PARP4* as a breast cancer susceptibility gene. This study highlights the importance of exercising caution when evaluating candidate cancer susceptibility genes and emphasizes the importance of performing functional analyses to verify the role of candidate breast cancer susceptibility genes.

Acknowledgements The authors are grateful to the volunteers who have participated in this study and colleagues who have referred patients. This study utilized datasets retrieved from dbGaP through the following accession numbers: phs001050.v1.p1 (BEAUTY), phs000822.v1.p1, and phs000601.v1.p1 (FAVR). The BEAUTY study was supported by Mayo Clinic Center for Individualized Medicine, Nadia's Gift Foundation, John P. Guider, The Pharmacogenomics Research Network (U10GM 61388-15), Mayo Clinic Cancer Center (CA15083-40A2), and Mayo Clinic Breast SPORE P50CA 116201-9 (Goetz, Ingle, Kalari, Suman). The phs000822.v1.p1 study was supported by the NHGRI Grant fund U54 HG003067, and the FAVR study was supported by NIH (R01CA155767), National Health and Medical Research Council, Australia (APP1025145), Cancer Council Victoria (Grant #628774), and Susan G Komen for the Cure (KG111155). We thank all the study participants, submitters, and all funding bodies of these studies.

Authors' contributions Conception and design: ASGL, PM, AP. Development of methodology: AP, JY, JCH, TWHS. Acquisition of material (including recruitment of patients): YSY, MHT, PA. Acquisition of data (conducted experiments and tests): AP, CHTC, GKL. Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): AP, PM, ASGL. Writing, review, and/or revision of the manuscript: AP, PM, ASGL. Study support and supervision: ASGL. All authors read and approved the final manuscript.

Funding This work was supported by a grant from the National Medical Research Council (NMRC) of Singapore (NMRC/CBRG/0034/2013) awarded to A.S.G. Lee and by Centre Grant NMRC support to the National Cancer Centre of Singapore.

Data availability Additional data are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest MHT has executive responsibilities at and is a shareholder of Lucence Diagnostics Pte. Ltd. No potential conflicts of interest were disclosed by the other authors.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Written informed consent was obtained from all study participants.

References

1. Kuchenbaecker KB et al (2017) Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA 317(23):2402–2416
2. Larsen MJ, Thomassen M, Gerdes A-M, Kruse TA (2014) Hereditary breast cancer: clinical, pathological and molecular characteristics. Breast Cancer (Auckl) 8:145–155
3. Wong ESY et al (2016) Inherited breast cancer predisposition in Asians: multigene panel testing outcomes from Singapore. NPJ Genom Med 1:15003
4. Ikeda Y et al (2016) Germline PARP4 mutations in patients with primary thyroid and breast cancers. Endocr Relat Cancer 23(3):171–179
5. Alimirzaie S et al (2018) Mutations in known and novel cancer susceptibility genes in young patients with pancreatic cancer. Arch Iran Med 21(6):228–233
6. Chen Y, Wang L, Xu H, Liu X, Zhao Y (2013) Exome capture sequencing reveals new insights into hepatitis B virus-induced hepatocellular carcinoma at the early stage of tumorigenesis. Oncol Rep 30(4):1906–1912
7. Still IH, Vince P, Cowell JK (1999) Identification of a novel gene (ADPRTL1) encoding a potential Poly(ADP-ribosyl)transferase protein. Genomics 62(3):533–536
8. Gibson BA, Kraus WL (2012) New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. Nat Rev Mol Cell Biol 13(7):411–424
9. Morales J et al (2014) Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases. Crit Rev Eukaryot Gene Expr 24(1):15–28
10. Kickhoefer VA et al (1999) The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase. J Cell Biol 146(5):917–928
11. van Zon A, Mossink MH, Scheper RJ, Sonneveld P, Wiemer EAC (2003) The vault complex. Cell Mol Life Sci 60(9):1828–1837
12. Izquierdo MA, Scheffer GL, Schroeijers AB, de Jong MC, Scheper RJ (1998) Vault-related resistance to anticancer drugs determined by the expression of the major vault protein LRP. Cytotechnology 27(1–3):137–148
13. Siva AC et al (2001) Up-regulation of vaults may be necessary but not sufficient for multidrug resistance. Int J Cancer 92(2):195–202
14. Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31(13):3812–3814
15. Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 76(1):7–20
16. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D (2010) Mutation Taster evaluates disease-causing potential of sequence alterations. Nature Methods 7(8):575–576
17. Reva B, Antipin Y, Sander C (2011) Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 39(17):e118
18. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46(3):310–315
19. Bankhead P et al (2017) QuPath: open source software for digital pathology image analysis. Sci Rep 7(1):16878
20. van Zon A et al (2002) Structural domains of vault proteins: a role for the coiled coil domain in vault assembly. Biochem Biophys Res Commun 291(3):535–541
21. Vyas S, Chesarone-Cataldo M, Todorova T, Huang Y-H, Chang P (2013) A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. Nat Commun 4:2240
22. Raval-Fernandes S, Kickhoefer VA, Kitchen C, Rome LH (2005) Increased susceptibility of vault poly(ADP-ribose) polymerase-deficient mice to carcinogen-induced tumorigenesis. Cancer Res 65(19):8846–8852
23. Liu Y et al (2004) Vault poly(ADP-ribose) polymerase is associated with mammalian telomerase and is dispensable for telomerase function and vault structure in vivo. Mol Cell Biol 24(12):5314–5323
24. Prat A et al (2013) Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. Breast Cancer Res Treat 142(2):237–255
25. Dai X, Cheng H, Bai Z, Li J (2017) Breast cancer cell line classification and its relevance with breast tumor subtyping. J. Cancer 8(16):3131–3141
26. Tabin CJ et al (1982) Mechanism of activation of a human oncogene. Nature 300(5888):143–149