



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Lee, JEA;Li, N;Rowley, SM;Cheasley, D;Zethoven, M;McInerny, S;Gorringe, KL;James, PA;Campbell, IG

Title:

Molecular analysis of PALB2-associated breast cancers

Date:

2018-05-01

Citation:

Lee, J. E. A., Li, N., Rowley, S. M., Cheasley, D., Zethoven, M., McInerny, S., Gorringe, K. L., James, P. A. & Campbell, I. G. (2018). Molecular analysis of PALB2-associated breast cancers. *Journal of Pathology*, 245 (1), pp.53-60. <https://doi.org/10.1002/path.5055>.

Persistent Link:

<https://hdl.handle.net/11343/283590>

Molecular analysis of *PALB2* associated breast cancers

Jue Er Amanda Lee¹, Na Li^{1,2,3}, Simone M. Rowley¹, Dane Cheasley^{1,3}, Magnus Zethoven⁴, Simone McInerny⁵, Kylie L. Gorringer^{3,6,7}, Paul A. James^{3,5,*} and Ian G. Campbell^{1,3,7,*}

¹Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia;

²Cancer Biology Medical Center, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China;

³Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Victoria, Australia;

⁴Bioinformatics Consulting Core, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia;

⁵Parkville Familial Cancer Centre, Peter MacCallum Cancer Centre and Royal Melbourne Hospital, Melbourne, Victoria, Australia;

⁶Cancer Genomics Program, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

⁷Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia;

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/path.5055](https://doi.org/10.1002/path.5055)

*These Authors contributed equally

Correspondence to: *Ian Campbell, Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, Victoria, 3000, Australia.*

E-mail: ian.campbell@petermac.org.

Conflict of interest statements

No conflicts of interest were declared.

Running title: Molecular analysis of *PALB2* associated breast cancers

Abstract

PALB2 is established as the most clinically important moderate to high penetrance breast cancer predisposition gene after *BRCA1* and *BRCA2*. Mutations in classical familial cancer predisposition genes are presumed to be recessive at the cellular level and therefore a second inactivating somatic mutation is required in the tumour tissue. However, from the limited data that exists, *PALB2* may be an example of a cancer predisposition gene that does not conform to Knudsen's "two hit" paradigm. We conducted genome-wide copy number analysis and targeted sequencing of *PALB2* and other breast cancer driver genes in 15 invasive breast cancers from individuals carrying pathogenic germline mutations in *PALB2*. The majority of cancers showed clear evidence of bi-allelic inactivation of *PALB2* (10/15) either as loss of heterozygosity involving the wild-type allele (6 tumours) or somatic point mutations (4 tumours). All *PALB2* null cancers had high homologous recombination deficiency (HRD) scores consistent with a homologous recombination repair deficiency. Interestingly, all but one of the *PALB2* heterozygous cancers also had high HRD scores suggesting alternative mechanisms of *PALB2* functional loss might be operating in these cancers. Our findings demonstrate that *PALB2* does undergo bi-allelic inactivation in the majority of breast cancers from *PALB2* germline mutation carriers. This feature has implications for the discovery of new moderate to high penetrance breast cancer predisposition genes as it supports using the existence of a "second hit" and mutation signatures as important search criteria.

Keywords

PALB2; Loss of heterozygosity; tumour suppressor gene; homologous recombination deficiency

Introduction

PALB2 was first identified as a *BRCA2*-interacting protein and involved in homologous recombination-mediated repair of double-strand DNA breaks [1]. Bi-allelic mutations in *PALB2* are responsible for Fanconi anaemia complementation group N [2,3] while mono-allelic mutations are associated with increased breast cancer susceptibility [4]. *PALB2* is now established as the most clinically important breast cancer predisposition gene after *BRCA1* and *BRCA2* [5].

Mutations in classical familial cancer predisposition genes are presumed to be recessive at the cellular level and therefore require a second somatic inactivating mutation in the tumour according to the model first proposed by Knudson from his observations of familial retinoblastoma [6]. In the search for new moderate to high penetrance familial breast cancer genes, the absence of a “second hit” in a candidate gene has generally been considered sufficient to nullify its candidacy. However, since its discovery, *PALB2* has been cited as an example of a predisposition gene that may not require bi-allelic inactivation in tumours [7-9]. Data from seven studies that have collectively analysed 20 breast tumours from germline carriers of *PALB2* loss of function (LoF) mutations [7-13] provided mixed results but overall over 50% retained both the wild-type and mutant *PALB2* alleles (Table 1) although for one study a second nonsense somatic mutation was

detected [13]. Since only one of the seven studies sequenced the entire coding region of the *PALB2* gene in the tumours, the presence of second somatic point mutations cannot be excluded and it is consequently premature to conclude that *PALB2* does not undergo bi-allelic inactivation. Resolving this issue has important implications for the strategies that should be adopted to identify the genes responsible for the ~50% of breast cancer families that are not explained by known predisposition genes.

In order to determine whether *PALB2* requires bi-allelic inactivation and if this is associated with a characteristic genomic scar signature, we analysed DNA extracted from 15 breast cancer tumours from women with a *PALB2* germline loss of function (LoF) mutation.

Materials and methods

Breast tumour cohort

PALB2 LoF mutation carrier families were identified from the Variants in Practice (ViP) Study through the Victorian and Tasmanian Familial Cancer Clinics, Australia. The family and/or personal history of all the cases was assessed by a specialist Familial Cancer Clinic and determined to be sufficiently strong to be eligible for clinical genetic testing by local criteria. All cases tested negative for pathogenic mutations in *BRCA1*, *BRCA2* and other hereditary breast and ovarian cancer genes as described previously [14]. Fifteen cases where an actionable germline LoF *PALB2* was identified also had formalin fixed paraffin embedded (FFPE) breast cancer biopsies available. Carrier subject details and tumour characteristics obtained from pathology reports are listed in Table 2 and Supplementary Table S1. All participants provided informed consent for genetic analysis of their germline and tumour DNA. This study was approved by the Human Research Ethics Committees at each participating ViP study recruitment centre and the Peter MacCallum Cancer Centre (Approval # 09/29).

Tumour DNA extraction

For each FFPE tumour, a representative haematoxylin and eosin (H&E) stained section was prepared and used as a template to needle microdissect cancer cells from H&E stained 10 µm sections. DNA was extracted using the QIAamp DNA FFPE Tissue Kit (Qiagen, Valencia, CA, USA) and quantified using Qubit dsDNA high-sensitivity Assay kit (ThermoFisher Scientific, Australia Pty Ltd).

Targeted sequencing library preparation, enrichment and sequencing

Targeted sequencing of tumour DNA was performed using an Agilent SureSelect XT Custom Panel which targeted 487 genes (total targeted region of 1.337 Mb) including *BRCA1*, *BRCA2*, *PTEN*, *ATM*, *ATR*, *BRIP1* and *RAD51C* and the entire coding region and introns of *PALB2* spanning coordinates chr16:23614482-23657678 (*Homo sapiens*, hg19, GRCh37, February 2009). Library preparation was performed from an input of 300 ng of tumour DNA using the KAPA Hyper system (Agilent, Santa Clara, USA). Sequencing of target-enriched DNA libraries were performed using the Illumina Next Seq 500 generating 75 bp paired-end sequence reads.

Variant calling, somatic mutation detection and genome-wide copy number analysis

Paired-end sequence reads were aligned to the g1k v37 hg19 reference genome using BWA [15]. Optical duplicate reads were removed using Picard (<http://broadinstitute.github.io/picard/>), then local realignment around indels and base quality score recalibration were performed using the Genome Analysis Tool Kit (GATK) in accordance with their recommended best practice workflow [16]. SNP and indel variants were called using GATK UnifiedGenotyper, Platypus [17] and VarScan 2 [18]. Called variants were additionally annotated using the Ensembl Variant Effect Predictor [19]. Somatic mutations in the tumour sequencing data were identified by removing previously available germline variant data for hereditary breast and ovarian cancer panel genes, including *PALB2* [14] and where this information was not available, by applying the following filters: canonical transcript; bidirectional read; quality ≥ 100 ; variants identified by

at least two variant callers; minor allele frequency (MAF) present at ≤ 0.0001 in ExAc (Version 0.3.1, excluding TCGA data, released March 14 2016) [20], GnomAD (Version 2.0, released 27 February 2017) [20], EVS (Version ESP6500SI-V2-SSA137) [21] and 1000 Genomes [22] databases.

Off-target sequencing reads were used to generate genome-wide copy number (CN) data using copywriteR [23] utilising a normal lymphocyte DNA control (NA12878, Coriell Institute) run in the same sequencing batch for the normalisation baseline. Data was then imported to NEXUS Copy Number™ (software v8.0 with build version 9169, BioDiscovery Inc), segmented using a FASST2 segmentation algorithm and visualized.

Analysis of Loss-of-Heterozygosity (LOH) across chromosome 16

In addition to copy number data, allelic information across the *PALB2* locus on chromosome 16 was used to substantiate LOH. High confidence germline heterozygous SNPs across chromosome 16 were identified from germline DNA data generated previously [14] using the following filters: canonical transcript; called by two or more callers; bidirectional read; quality ≥ 50 ; present in one or more population databases ExAc, GnomAD, EVS and 1000 Genomes. The alternate allele frequency of these SNPs in the tumour tissue were then used to infer allelic status.

Generation of homologous recombination deficiency (HRD) scores and mutational signatures

Using the genome-wide CN data, a HRD score [24] was calculated for each tumour by summing the individual scores for Telomeric allelic imbalances (NtAI) [25], Large-scale

State transitions (LSTs) [26] and Homologous recombination deficiency-Loss-of-heterozygosity (HRD-LOH) [27] following methods previously described [25-27].

Mutational signatures were defined by measuring the mutational context of somatic single nucleotide variation (SNVs) within the target regions. Somatic mutations located in the capture target region from tumour sequencing were analysed and enriched for likely somatic mutations using the following filters: called by two or more callers; alternate variant proportion more than 10%; $MAF \leq 0.0001$ in ExAC, GnomAD, EVS, and 1000 Genomes in overall population and European population; variant recurrence is no more than 2 tumours, except for genes known to be frequently mutated genes in breast cancer (*PIK3CA*, *TOX3*, *TP53*, *PTEN*, *SPEN*, *GATA3*, *KRAS* and *BRAF*) [28]. For each of the 15 tumours with germline *PALB2* mutation and 35 sporadic breast tumours (negative for germline *BRCA1*, *BRCA2* and *PALB2* LoF mutations), the somatic SNVs were categorized into six basic base substitutions (C>A, C>G, C>T, T>A, T>C and T>G) and subcategorized into 96 subcategories according to the trinucleotide context. To overcome the limitation of the small number of somatic mutations in each sample detected using the targeted sequencing panel (average 15 per sample), we pooled the mutations identified in each tumour set and analysed them as a single group. The mutational signatures were calculated and plotted using deconstructSigs [29] package in R v3.3.2 [30].

Promoter methylation analysis of *PALB2* and *BRCA1*

DNA from tumour (200ng-500ng) and Female Genomic Reference DNA control (Promega, Wisconsin, USA) (500ng) were bisulfite treated using the EpiTect Fast DNA Bisulfite Kit

(Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. A 135bp region of the CpG island located in the *PALB2* promoter, (chr16: 23,652,428 to 23,652,562; GRCh37/hg19) was PCR amplified as described by [31] using the primers 5'-TTTTCGGTTTAGGGTTAATTGGGTT-3' (forward) and 5'-CACCTTTTCCTTCTCCTCACAACCTAAA-3' (reverse). A 232bp region of the *BRCA1* promoter region was amplified as described in [32] using the primers 5'-GGTAGATTGGGTGGTTAATTTAGA-3' (forward) and 5'-CTAAAAAACCCACAACTATCC-3'. PCR products were then assessed for the extent of methylation by Sanger sequencing.

Results

***PALB2* germline mutation families and tumour characteristics.**

Families with germline LoF mutations in *PALB2* were identified by targeted gene panel sequencing of participants of the ViP study as reported previously [14]. Fifteen *PALB2* LoF mutations carriers from 15 different families had archival FFPE breast cancer blocks available for analysis. Tumour and patient details are summarized in Table S1. Seven cases were ER-positive, including 4 that were also HER2 positive and 8 patients were “triple negative” (ER/PR/HER2 negative). Hormone receptor statuses were confirmed by CN data. All 15 cases were Grade 3 and invasive ductal cancers.

Identification of somatic *PALB2* mutations in tumours

All exons and introns of *PALB2* were sequenced to a mean read depth of 382 across the entire gene for all samples and 99% of the coding region bases were covered by >20 reads. The respective germline LoF *PALB2* mutations were identified in all 15 cases and the variant allele read proportion was used as a measure of allelic status together with the allelic status of other germline SNPs detected across chromosome 16 and copy number data derived using off-target reads (Table 2; Figure S1). Four tumours harboured somatic point mutations in *PALB2* (three truncating mutations and one missense mutation) with each retaining the germline LoF mutation (cases 0306571, 0279859, 0393382 and 0410313). The p.Gly1068Arg somatic missense mutation was predicted by *in-silico* tools Condel [33], PolyPhen2 [34] and SIFT [35] to be deleterious and the CADD score [36] was high at 15.7. In addition, the p.Gly1068Arg mutation is located within the functionally important WD40 domain of *PALB2* and previously reported biochemical assays of *PALB2*

Author Manuscript

harbouring WD40 domain missense variants p.Leu939Trp, p.Thr1030Ile and p.Leu1143Pro have shown altered binding to RAD51C, RAD51 and BRCA2 [37]. A further 6 tumours showed strong evidence of LOH across the *PALB2* locus with all 6 retaining the *PALB2* LoF allele (cases 0149217, 0241796, 0319619, 0388519, 0401044 and 0112255) as evidenced by the high LoF allele frequency (Table 2). The LoF allele frequency for case 0388519 was only 0.65 but this is consistent with the ~30% tumour purity estimated for this case. Five tumours showed neither LOH nor somatic mutations at the locus (cases 0233453, 0252381, 0057225, 0317466 and 0324737). For case 0317466, the variant allele frequency was 0.63, but in the context of an estimated tumour purity of 80%, this frequency is more consistent with allelic imbalance rather than LOH. The presence of allelic imbalance is supported by SNP allele frequency data across the *PALB2* locus and chromosome 16, which show variant allele frequencies clustering around 0.3 and 0.6 consistent with retention of the wild-type allele and duplication of the chromosome harbouring the mutant allele in this case. Overall, 67% (10/15) of the tumours from germline *PALB2* LoF mutations carriers are concluded to be null for *PALB2*.

Homologous recombination deficiency in *PALB2* tumours

We scored the extent of large-scale genomic alterations consistent with homologous recombination deficiency to assess for correlation with *PALB2* status. Genome-wide copy number data were generated from off-target sequencing reads from the custom SureSelect targeted sequencing assay and a combined HRD score for genomic scar was generated by quantitating three separate parameters; NtAI, LST and HRD-LOH. In addition to the 15 *PALB2* tumours, we also scored a cohort of 35 sporadic breast cancers

that were assayed on the same SureSelect panel as the *PALB2* tumours. Compared to sporadic breast cancers, the *PALB2* tumours had significantly higher HRD scores regardless of *PALB2* allelic status (Figure 1). Interestingly, there was no significant difference in the HRD scores between *PALB2* heterozygotes and null tumours (Table 2; Figure 1) although one *PALB2* heterozygous tumour (case 0324737) has a HRD score of 1 with a near-normal chromosome complement. None of the *PALB2* tumours harboured additional germline or somatic mutations in hereditary breast cancer genes *BRCA1*, *BRCA2*, *PTEN*, *ATM*, *ATR*, *BRIP1*, *RAD51* or any of the Fanconi anaemia genes that could explain the high HRD scores in the *PALB2* heterozygous cancers, although somatic *TP53* mutations were detected in 3/5 of the *PALB2* heterozygous tumours compared to 3/10 *PALB2* null tumours (Table 2).

Mutation signatures were also assessed based on 110 SNVs identified in *PALB2* germline mutation tumours (43 SNVs among the HET group and 67 among the NULL group) and 225 SNVs among the 35 sporadic breast tumours. The *PALB2* tumours, but not the sporadic tumours, showed a prominent somatic mutation signature 3 (36.5% versus 0%) consistent with defective homologous recombination [13,28,38] (Figure S2). Because of the small number of mutations it was not possible to derive reliable signatures for the *PALB2* heterozygous and *PALB2* null tumours separately. High HRD scores and somatic mutational signature 3 are both hallmark features of tumours with *BRCA1* or *BRCA2* mutations and HRD [26]. Methylation of the *PALB2* and *BRCA1* promoter regions was assessed for 4 of 5 *PALB2* heterozygous and 2 of 10 *PALB2* null tumours where sufficient quantity and quality of DNA was available. Neither the heterozygous nor null *PALB2*

tumours showed any evidence of promoter hypermethylation in *PALB2* or *BRCA1* (Figure S3A, S3B).

The genomic loci commonly targeted by CN change were similar for the *PALB2* tumours compared to the 35 sporadic cancers analysed with the same sequencing panel (Figure S4) with the exception of the presence of CN loss of chromosome 1p among the *PALB2* tumours. Loss of 1p is not commonly reported in other large cohorts of sporadic breast cancer (<6% of cases affected) [39]. The 1p loss was more frequently seen among the *PALB2* null tumours, although the number of cases is small.

Discussion

The genetic causes for a large proportion of hereditary breast cancer families remain unresolved (BRCAx families). Despite the implementation of targeted and whole exome sequencing strategies, efforts to identify new genes have been slow and hampered by the apparent high genetic heterogeneity of the majority of BRCAx families [40]. While many candidate genes have been proposed, the rarity of families carrying mutations means that many types of data will be required to corroborate a role for a particular gene in breast cancer predisposition. Based on archetypical familial cancer gene *RB1*, the expectation for a recessive tumour suppressor gene is that tumours from individuals carrying germline pathogenic mutations will acquire a second somatic mutation in the wild-type allele, thus rendering the cell null for that gene. On this basis, sequencing of both germline and tumour DNA from individuals from BRCAx families is gaining favour in the search for new

breast cancer genes, with the absence of a second hit disqualifying that gene from further investigating. However, while the “two hit” model clearly applies for genes such as *BRCA1* [41,42] and reviewed in [43], there have been doubts as to whether this is required for other hereditary breast cancer genes. It is certainly true that a second somatic genetic or epigenetic mutation may not be required if the germline mutation is acting as a dominant negative, such as for the *ATM* missense variant p. Val2424Gly and some *TP53* missense mutations, but in these cases the tumour cell is effectively rendered null for each gene and are not genuine examples of a haplo-insufficient mode of cancer predisposition.

The role of *PALB2* as a moderate to high penetrance breast cancer predisposition gene responsible for ~1% to 2.4% of hereditary breast cancer families is now well established [5,14]. Based on limited analyses, *PALB2* had emerged as an example of a hereditary breast cancer gene that may not require a second hit [9] with only 50% showing LOH of the wild-type allele (Table 1). While one study reported that 7/7 germline *PALB2* mutant tumours showed loss of the wild-type allele [11], the majority of the studies reported infrequent allelic loss. The strongest evidence to date of a bi-allelic inactivation requirement for *PALB2* is from analysis of a single *PALB2* tumour where exome sequencing identified a somatic nonsense mutation, presumably in the wild-type allele. This tumour also showed a very high LST score and a mutation signature 3, both consistent with loss of homologous DNA repair capacity [13]. However, none of the other studies performed full exon sequencing of *PALB2*.

Our data suggest that the majority of *PALB2* tumours do show bi-allelic inactivation and therefore conform to Knudsen’s two hit model. In 6/6 cases where there was LOH at the

PALB2 locus, all showed clear loss of the wild-type allele. While it is possible this was a chance occurrence, this seems unlikely in the context of the high rate of somatic point mutations identified in *PALB2* (4 of 15 tumours, 27%) compared to the somatic mutation rate of 0.9% reported in breast cancer in general (COSMIC database, Version 82, released 3 August 2017 [44]). Consistent with this, the *PALB2* null tumours all had high HRD scores (average 73.9) similar to scores reported for *BRCA1* and *BRCA2* breast tumours (average 63.1) [45]. *PALB2* tumours also had significantly higher HRD scores than observed in our own cohort of sporadic breast cancers (average 26.2) and also those reported in the literature (average 28) [24]. Interestingly, with one exception, those tumours that retained the wild-type *PALB2* allele had HRD scores similar to the *PALB2* null tumours. None of the *PALB2* heterozygous tumours showed evidence of *PALB2* promoter hypermethylation, which is consistent with other studies showing hypermethylation at the *PALB2* promoter is infrequent in breast cancer [31,46]. In addition, none of these tumours had detectable *BRCA1* promoter methylation, or germline or somatic mutations in hereditary breast cancer genes that could explain such an HR defect, which raises the possibility that the functional copy of *PALB2* in these cancers is abrogated via other epigenetic or posttranslational mechanisms.

Taken together, our findings show that *PALB2* does generally conform to the two hit model and this is associated with a genomic scar consistent with loss of homologous DNA repair capacity. This has implications for the discovery of new breast cancer predisposition genes as it validates an emphasis on the presence of a “second hit” and a genomic scar signature consistent with loss of that gene.

Acknowledgements

We thank Norah Grewal, Rebecca Driessen and the ViP study investigators at the Victorian and Tasmanian Familial Cancer Centres: Alison Trainer, Gillian Mitchell, Marion Harris, Geoff Lindeman, Thomas John, Yoland Antill, Ingrid Winship and Mary-Anne Young who provided germline and histological data. kConFab investigators, Heather Thorne, Lynda Williams and Genna Glavich for calling in the tumour blocks. David Byrne for scoring the tumour purity, and the investigators in the Peter MacCallum Cancer Centre Genomics Core Facility Gisela Mir Arnau and Tim Semple for sequencing our tumours, the Bioinformatics & Cancer Genomics Laboratory for running our data through their bioinformatics pipelines. This work was supported by the National Breast Cancer Foundation, the Victorian Cancer Agency, and the National Health and Medical Research Council of Australia.

Statement of author contributions

IGC and PAJ conceived and designed the study. JEAL carried out experiments, acquired, analysed and interpreted the data. NL and SMR carried out experiments, acquired and analysed data. DC and KLG analysed and interpreted the data. SM and PAJ provided data and samples. MZ ran our data through the bioinformatics pipeline. JEAL, IGC and PAJ were involved in drafting the manuscript and all authors read and provided critical feedback on the manuscript.

References

1. Xia B, Sheng Q, Nakanishi K, *et al.* Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* 2006; **22**: 719-729.
2. Xia B, Dorsman JC, Ameziane N, *et al.* Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet* 2007; **39**: 159-161.
3. Reid S, Schindler D, Hanenberg H, *et al.* Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet* 2007; **39**: 162-164.
4. Rahman N, Seal S, Thompson D, *et al.* PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 2007; **39**: 165-167.
5. Antoniou AC, Casadei S, Heikkinen T, *et al.* Breast-cancer risk in families with mutations in PALB2. *N Engl J Med* 2014; **371**: 497-506.
6. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 1971; **68**: 820-823.
7. Tischkowitz M, Xia B, Sabbaghian N, *et al.* Analysis of PALB2/FANCN-associated breast cancer families. *Proc Natl Acad Sci U S A* 2007; **104**: 6788-6793.
8. Balia C, Sensi E, Lombardi G, *et al.* PALB2: a novel inactivating mutation in a Italian breast cancer family. *Fam Cancer* 2010; **9**: 531-536.

9. Hartley T, Cavallone L, Sabbaghian N, *et al.* Mutation analysis of PALB2 in BRCA1 and BRCA2-negative breast and/or ovarian cancer families from Eastern Ontario, Canada. *Hered Cancer Clin Pract* 2014; **12**: 19.
10. Garcia MJ, Fernandez V, Osorio A, *et al.* Analysis of FANCB and FANCN/PALB2 fanconi anemia genes in BRCA1/2-negative Spanish breast cancer families. *Breast Cancer Res Treat* 2009; **113**: 545-551.
11. Casadei S, Norquist BM, Walsh T, *et al.* Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer. *Cancer Res* 2011; **71**: 2222-2229.
12. Pakkanen S, Wahlfors T, Siltanen S, *et al.* PALB2 variants in hereditary and unselected Finnish prostate cancer cases. *J Negat Results Biomed* 2009; **8**: 12.
13. Foo TK, Tischkowitz M, Simhadri S, *et al.* Compromised BRCA1-PALB2 interaction is associated with breast cancer risk. *Oncogene* 2017; **36**: 4161-4170.
14. Thompson ER, Rowley SM, Li N, *et al.* Panel testing for familial breast cancer: Calibrating the tension between research and clinical care. *J Clin Oncol* 2016; **34**: 1455-1459.
15. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013; arXiv:1303.3997v1 [q-bio.GN]. Available at <https://arxiv.org/archive/q-bio>.
16. Van der Auwera GA, Carneiro MO, Hartl C, *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013; **43**: 11 10 11-33.

17. Rimmer A, Phan H, Mathieson I, *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* 2014; **46**: 912-918.
18. Koboldt DC, Zhang Q, Larson DE, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; **22**: 568-576.
19. McLaren W, Gil L, Hunt SE, *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 2016; **17**: 122.
20. Lek M, Karczewski KJ, Minikel EV, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536**: 285-291.
21. Exome Variant Server. <http://evs.gs.washington.edu/EVS/>. Last accessed September 2017.
22. Genomes Project C, Auton A, Brooks LD, *et al.* A global reference for human genetic variation. *Nature* 2015; **526**: 68-74.
23. Kuilman T, Velds A, Kemper K, *et al.* CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol* 2015; **16**: 49.
24. Marquard AM, Eklund AC, Joshi T, *et al.* Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark Res* 2015; **3**: 9.
25. Birnbak NJ, Wang ZC, Kim JY, *et al.* Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Discov* 2012; **2**: 366-375.

26. Popova T, Manie E, Rieunier G, *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res* 2012; **72**: 5454-5462.
27. Abkevich V, Timms KM, Hennessy BT, *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer* 2012; **107**: 1776-1782.
28. Nik-Zainal S, Davies H, Staaf J, *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016; **534**: 47-54.
29. Rosenthal R, McGranahan N, Herrero J, *et al.* DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol* 2016; **17**: 31.
30. Team RDC. R: A Language and Environment for Statistical Computing. *Vienna, Austria : the R Foundation for Statistical Computing* 2011.
31. Mikeska T, Alsop K, Australian Ovarian Cancer Study G, *et al.* No evidence for PALB2 methylation in high-grade serous ovarian cancer. *J Ovarian Res* 2013; **6**: 26.
32. Watanabe Y, Maeda I, Oikawa R, *et al.* Aberrant DNA methylation status of DNA repair genes in breast cancer treated with neoadjuvant chemotherapy. *Genes Cells* 2013; **18**: 1120-1130.
33. Gonzalez-Perez A, Lopez-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* 2011; **88**: 440-449.

34. Adzhubei IA, Schmidt S, Peshkin L, *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248-249.
35. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res* 2001; **11**: 863-874.
36. Kircher M, Witten DM, Jain P, *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014; **46**: 310-315.
37. Park JY, Zhang F, Andreassen PR. PALB2: the hub of a network of tumor suppressors involved in DNA damage responses. *Biochim Biophys Acta* 2014; **1846**: 263-275.
38. Polak P, Kim J, Braunstein LZ, *et al.* A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. *Nat Genet* 2017; **49**: 1476-1486.
39. Curtis C, Shah SP, Chin SF, *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012; **486**: 346-352.
40. Thompson ER, Goringe KL, Rowley SM, *et al.* Prevalence of PALB2 mutations in Australian familial breast cancer cases and controls. *Breast Cancer Res* 2015; **17**: 111.
41. Esteller M, Silva JM, Dominguez G, *et al.* Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst* 2000; **92**: 564-569.
42. Staff S, Isola J, Tanner M. Haplo-insufficiency of BRCA1 in sporadic breast cancer. *Cancer Res* 2003; **63**: 4978-4983.

43. Turner N, Tutt A, Ashworth A. Hallmarks of 'BRCAness' in sporadic cancers. *Nat Rev Cancer* 2004; **4**: 814-819.
44. Simon A, Forbes, David Beare, Harry Boutselakis, *et al.* COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* 2017; **45**: D777–D783.
45. Telli ML, Timms KM, Reid J, *et al.* Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clin Cancer Res* 2016; **22**: 3764-3773.
46. Potapova A, Hoffman AM, Godwin AK, *et al.* Promoter hypermethylation of the PALB2 susceptibility gene in inherited and sporadic breast and ovarian cancer. *Cancer Res* 2008; **68**: 998-1002.

Table 1. Summary of published tumour analyses of *PALB2* carrier breast tumours

Study	No. of Tumours Studied	Allelic Status		Somatic Point Mutation	Technique
		LOH Wildtype Allele	LOH Mutant Allele		
Tischkowitz (2007) [7]	4	0	0	Not tested	Site specific Sanger sequencing
Garcia (2009) [10]	1	1	0	Not tested	Site specific Sanger sequencing
Pakkanen (2009) [12]	2	0	0	Not tested	Site specific Sanger sequencing
Balia (2010) [8]	1	0	0	Not tested	Site specific Sanger sequencing
Casadei (2011) [11]	7	7	0	Not tested	Site specific Sanger sequencing
Hartley (2014) [9]	4	1	0	Not tested	Site specific Sanger sequencing
Foo (2017) [13]	1	0	0	1	Whole exome sequencing
Total	20	9	0	1	

LOH, loss of heterozygosity

Table 2. Overall *PALB2* carrier information and mutation summary of breast cancer tumours for this study

Sample ID	<i>PALB2</i> Status	<i>PALB2</i> Mutation Nucleotide (Protein)	<i>Tp53</i> Mutation Nucleotide (Protein)	Tumour Purity	HRD Score	Read
0233453	HET	c.522_523delAA (p.Arg175ThrfsTer9)	c.1024C>T (p.Arg342Ter)	60%	73	85
0252381	HET	c.1037_1041delAAGAA (p.Lys346ThrfsTer13)	c.659A>G (p.Tyr220Cys)	60%	66	20
0057225	HET	c.3362delG (p.Gly1121ValfsTer3)	.	45%	83	28
0317466	HET	c.3113G>A (p.Trp1038Ter)	c.695T>A (p.Ile232Asn)	80%	74	75
0324737	HET	c.3332delC (p.Pro1111LeufsTer13)	.	65%	1	70
0149217	NULL	c.2386G>T (p.Gly796Ter)	.	50%	90	55
0241796	NULL	c.3113G>A (p.Trp1038Ter)	.	85%	76	15
0319619	NULL	c.2915delT (p.Leu972ArgfsTer6)	.	70%	84	74
0388519	NULL	c.3113G>A (p.Trp1038Ter)	c.734G>T (p.Gly245Val)	30%	53	27
0401044	NULL	c.3113G>A (p.Trp1038Ter)	.	65%	58	33
0112255	NULL	c.268G>T (p.Glu90Ter)	c.742C>T (p.Arg248Trp)	75%	101	28
0306571	NULL	c.268G>T (p.Glu90Ter)	.	60%	75	27
		*c.3202G>C (p.Gly1068Arg)				20
0279859	NULL	c.3113G>A (p.Trp1038Ter)	.	80%	61	46
		*c.3193_3201+6delICTGAAATGGTAAGT (p.?)				44
0393382	NULL	c.2771delC (p.Pro924GlnfsTer11)	c.909_919+13del (p.?)	60%	91	56
		*c.876_898del (p.Gln292HisfsTer3)				33
0410313	NULL	c.3507_3508delTC (p.His1170PhefsTer19)	.	70%	50	27
		*c.2630G>A (p.Trp877Ter)				47

HET, Heterozygous tumour; NULL, tumour with bi-allelic inactivation; *indicates *PALB2* somatic mutation; HRD score, sum of the individual NtAI, LSTs and HRD-LOH scores

Figure Legend.

Figure 1. Large scale genomic alteration scores for 35 sporadic breast carcinomas, 10 *PALB2* null and 5 *PALB2* heterozygous tumours plotted as (A) HRD score, (B) NtAI, (C) LST and (D) HRD-LOH. p values are shown for each group.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods **NO**

Supplementary figure legends **YES**

Table S1. Carrier subject details and tumour characteristics obtained from pathology reports

Figure S1. Analysis of Loss-of-heterozygosity and allelic imbalance across chromosome 16, including *PALB2*

Figure S2. Mutational signatures of *PALB2* tumours versus sporadic breast tumours

Figure S3. Bisulfite sequencing of the *PALB2* tumours within the *PALB2* and *BRCA1* promoter regions

Figure S4. Genome wide copy number plots of sporadic breast cancers vs *PALB2* tumours