

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Cheasley, D;Li, N;Rowley, SM;Elder, K;Mann, GB;Loi, S;Savas, P;Goode, DL;Kader, T;Zethoven, M;Semple, T;Fox, SB;Pang, JM;Byrne, D;Devereux, L;Nickson, C;Procopio, P;Lee, G;Hughes, S;Saunders, H;Fujihara, KM;Kuykhoven, K;Connaughton, J;James, PA;Gorringe, KL;Campbell, IG

Title:

Molecular comparison of interval and screen-detected breast cancers

Date:

2019-06-01

Citation:

Cheasley, D., Li, N., Rowley, S. M., Elder, K., Mann, G. B., Loi, S., Savas, P., Goode, D. L., Kader, T., Zethoven, M., Semple, T., Fox, S. B., Pang, J. M., Byrne, D., Devereux, L., Nickson, C., Procopio, P., Lee, G., Hughes, S. ,... Campbell, I. G. (2019). Molecular comparison of interval and screen-detected breast cancers. *Journal of Pathology*, 248 (2), pp.243-252. <https://doi.org/10.1002/path.5251>.

Persistent Link:

<https://hdl.handle.net/11343/285568>

## Molecular comparison of interval and screen-detected breast cancers

Dane Cheasley<sup>1,2</sup>, Na Li<sup>1,2</sup>, Simone M Rowley<sup>1</sup>, Kenneth Elder<sup>3,4,5</sup>, G Bruce Mann<sup>3,4</sup>,  
Sherene Loi<sup>2,6</sup>, Peter Savas<sup>2,6</sup>, David L Goode<sup>2</sup>, Tanjina Kader<sup>1</sup>, Magnus Zethoven<sup>7</sup>,  
Tim Semple<sup>7</sup>, Stephen B Fox<sup>9</sup>, Jia-Min Pang<sup>9</sup>, David Byrne<sup>9</sup>, Lisa Devereux<sup>1,10</sup>,  
Carolyn Nickson<sup>11</sup>, Pietro Procopio<sup>11</sup>, Grant Lee<sup>11</sup>, Siobhan Hughes<sup>1</sup>, Hugo  
Saunders<sup>1</sup>, Kenji M Fujihara<sup>1</sup>, Keilly Kuykhoven<sup>1</sup>, Jacquie Connaughton<sup>1</sup>, Paul A  
James<sup>2,12</sup>, Kylie L Gorringer<sup>2,13</sup>, and Ian G Campbell<sup>1,2\*</sup>

<sup>1</sup>Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, Victoria,  
Australia

<sup>2</sup>Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne,  
Victoria, Australia

<sup>3</sup>Department of Surgery, University of Melbourne, Melbourne, VIC, Australia

<sup>4</sup>The Royal Melbourne and Royal Women's Hospitals, Parkville, VIC, Australia

<sup>5</sup>The Edinburgh Breast Unit, Western General Hospital, Edinburgh, UK

<sup>6</sup>Division of Clinical Medicine and Research, Peter MacCallum Cancer Centre,  
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.5251](https://doi.org/10.1002/path.5251)

<sup>7</sup>Bioinformatics Consulting Core, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

<sup>8</sup>Genomics Core, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

<sup>9</sup>Department of Pathology, Peter MacCallum Cancer Centre, and University of Melbourne, Victoria, Australia

<sup>10</sup>Lifepool, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

<sup>11</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne

<sup>12</sup>Parkville Familial Cancer Centre, Peter MacCallum Cancer Centre and Royal Melbourne Hospital, Melbourne, Victoria, Australia

<sup>13</sup>Cancer Genetics and Genomics Program, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia<sup>14</sup>Cancer Research Division, Cancer Council NSW, Sydney, NSW

**\*Correspondence to:** Ian Campbell, Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, Victoria, 3000, Australia. Tel: +61 385516009. E-mail: [ian.campbell@petermac.org](mailto:ian.campbell@petermac.org).

#### **Conflict of interest statements**

No conflicts of interest were declared.



## Abstract

Breast cancer (BC) diagnosed after a negative mammogram but prior to the next screening episode is termed an “interval breast cancer” (IBC). Understanding the molecular differences between IBC and screen-detected breast cancers (SDBC) could improve mammographic screening and management options. Therefore, we assessed both germline and somatic genomic aberrations in a prospective cohort. Utilizing the Lifepool cohort of >54,000 women attending mammographic screening programs, 930 BC cases with screening status were identified (726 SDBC and 204 IBC). Clinico-pathological and family history information were recorded. Germline and tumour DNA were collected where available and sequenced for BC predisposition and driver gene mutations. Compared to SDBC, IBCs were significantly associated with a younger age at diagnosis and tumour characteristics associated with worse prognosis. Germline DNA assessment of BC cases that developed post-enrolment (276 SDBCs and 77 IBCs) for pathogenic mutations in 12 hereditary BC predisposition genes identified 8 carriers (2.27%). The germline mutation frequency was higher in IBC versus SDBC, although not statistically significant (3.90% versus 1.81%,  $p=0.174$ ). Comparing somatic genetic features of IBC and SDBC matched for grade, histological subtype and hormone receptor revealed no significant differences, with the exception of higher homologous recombination deficiency scores in IBC, and copy number changes on chromosome Xq in TN SDBCs. Our data demonstrates that while IBCs are clinically more aggressive than SDBC, when matched for confounding clinico-pathological features

they do not represent a unique molecular class of invasive BC, but could be a consequence of timing of tumour initiation and mammographic screening.

### **Keywords**

Breast cancer, interval breast cancer, screen-detected breast cancer, mammographic screening, next generation sequencing, mammographic density, genomics, somatic mutation, genetic predisposition, copy number alteration

## Introduction

Interval breast cancers (IBC) are defined as those that arise in the interval between a confirmed negative mammogram and the next scheduled screening episode. IBC can account for ~25% of invasive breast cancer diagnoses in women participating in population-based mammographic screening [1–3] and this rate is therefore a crucial determinant of the effectiveness of screening. IBCs are typically larger in size, higher grade, frequently node-positive at diagnosis, negative for oestrogen and progesterone hormone receptors and associated with a triple-negative phenotype. IBC occur more frequently in younger women, women with high mammographic density and some studies report an association with a family history of breast cancer [3-25]. IBCs have a worse average prognosis compared to screen-detected breast cancers (SDBC) suggesting they may be biologically distinct [26-28].

It is unclear whether IBCs represent a distinct molecular subtype of breast cancer (BC) or are identical to SDBC and simply a consequence of the timing of tumour initiation and/or mammographic density. Whilst large repositories of genomic data are available for BC in general, this data is seldom linked to mammographic screening status and consequently there is a critical knowledge gap in our understanding of the biology and genomics of IBCs. If IBCs represent a distinct biological subclass of BC when compared to the SDBC matched intrinsic subtype, this information might identify novel therapeutic targets or opportunities to devise effective management strategies to reduce the interval cancer rate. Therefore, the aim of this study was to determine the biological relationship of IBC and SDBC

through germline and somatic genomic analyses of a large cohort of well annotated BC for which screening status was known.

## Materials and Methods

### Study cohort

Subjects were recruited from the prospective Lifepool study ([www.lifepool.org](http://www.lifepool.org)) and a schematic of the study cohort is presented in **Figure 1**. Epidemiological and mammographic screening data were available for all study subjects and cancer incidence was determined by linkage to the Victorian Cancer Registry and BreastScreen Victoria. SDBC's were recorded by BreastScreen and IBC's were identified by linking Lifepool participant identities with the Victorian Cancer Registry and BreastScreen records, following the standard protocol used by BreastScreen [29]. IBC's in Australia were defined as a diagnosis (invasive breast cancer or *in situ* disease) diagnosed within the 27 month interval after a negative mammogram but prior to the next scheduled screening [29]. Additionally, the prior mammogram was re-reviewed by a radiologist through BreastScreen for those cases diagnosed as an IBC in order to exclude a false-negative examination. Pathology assessment such as hormone status, Ki67 and tumour grade was performed by the local pathologist/s and the treating hospital. All research pertaining to the Lifepool cohort was approved by the Peter MacCallum Cancer Centre Human Ethics Committee under ethics approval #0966.

### Tumour/normal DNA extraction

For germline analysis, DNA was extracted from peripheral blood lymphocytes or saliva using the Chemagen DNA isolation kit coupled with a Hamilton Robotics liquid handling workstation. For formalin-fixed paraffin embedded tumours, 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  $\mu\text{m}$  thick sections. DNA was extracted using the QIAamp DNA FFPE Tissue Kit (Qiagen, Valencia, CA, USA) and quantified using the Qubit dsDNA high-sensitivity assay kit (Thermo Fisher Scientific, Australia Pty Ltd) as described previously [30].

### **Targeted sequencing**

Targeted sequencing of germline DNA was performed using a custom-designed HaloPlex Targeted Enrichment Assay (Agilent Technologies, Santa Clara, CA, USA) as described previously [31-33]. The custom panel included all exons and intron-exon boundaries of 12 hereditary breast and ovarian cancer (HBOC) genes, and a panel of low penetrance breast cancer risk single nucleotide polymorphisms (SNPs) (**supplementary material, Table S1**). Indexed, 96-sample pooled libraries were analysed using a HiSeq2500 Genome Analyzer (Illumina, San Diego, CA, USA) to generate paired-end 100 bp sequencing reads.

Targeted sequencing of tumour DNA was performed using a SureSelect XT Custom Panel (Agilent Technologies), targeting all 12 HBOC genes included in the germline panel and 27 genes commonly somatically mutated in breast cancer [34] (**supplementary material, Table S1**). Library preparation was performed on an

input of 300 ng of tumour DNA using the KAPA Hyper system (Agilent). Sequencing of target-enriched DNA libraries was performed using an Illumina NextSeq500 that generated paired-end 75 bp sequence reads.

### **Germline and somatic mutation detection**

Paired-end sequence reads were aligned to the GRCh37 human reference genome using BWA-MEM [35]. Optical and PCR 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 [36]. SNP and indel variants were called using GATK UnifiedGenotyper, Platypus [37] and VarScan 2 [38] and annotated using the Ensembl Variant Effect Predictor [39].

The germline variant calling pipeline has been described previously [40–42]. The variant data were filtered for loss-of-function mutations (defined as nonsense or frameshift or essential splice site mutations) and missense mutations that were classified as known pathogenic in ClinVar [43]. Somatic mutations in the tumour sequencing data were identified by removing previously available germline variant data for HBOC panel genes, and where this information was not available, by applying the following filters: canonical transcript; quality  $\geq 100$ ; variants identified by at least two variant callers; minimum variant allele proportion  $\geq 0.2$ ; minor allele frequency present at  $\leq 0.0001$  in ExAc (Version 0.3.1, excluding TCGA data,

released March 14 2016) [44], GnomAD (Version 2.0, released 27 February 2017) [44], EVS (Version ESP6500SI-V2-SSA137) [45] and 1000 Genomes database [46].

### **Genome-wide copy number analysis**

Off-target sequencing reads from the tumour sequencing data were used to generate genome-wide copy number (CN) data for 46 IBCs and 64 SDBC's using CopywriteR [47] utilising a normal lymphocyte DNA control run in the same sequencing batch for the normalisation baseline (NA12878, Coriell Institute). An additional 9 SDBC's were analysed for genomic CN using the Affymetrix Molecular Inversion Probe 330K array. Data were imported into NEXUS Copy Number™ (software v8.0 with build version 9169, BioDiscovery Inc., El Segundo, CA, USA), segmented using a FASST2 segmentation algorithm and visualised. Comparisons between groups were made using Nexus, applying thresholds of  $p < 0.05$  and at least 25% frequency difference. Within the Nexus software package, gains were  $\log_2$  ratio defined as  $> 0.3$  and losses at  $\leq -0.3$ . High copy number gains were called if the  $\log_2$  ratio was  $> 1.2$  and homozygous deletions  $\leq -1.2$ .

### **Polygenic risk score**

To analyse the association of the risk conferred by low penetrance breast cancer SNPs between SDBC and IBC, a polygenic risk score (PRS) was generated using data for 69 BC risk SNPs. A PRS was derived for each individual using the process reported previously [48]. The PRS generated for the IBC and SDBC was normalized to the PRS generated using the same SNPs for 4,857 cancer-free

women from Lifepool.

### **Homologous recombination deficiency scoring**

Using the genome-wide CN data, a homologous recombination deficiency (HRD) score [49] was calculated for each tumour by summing the individual scores for telomeric allelic imbalances, large-scale state transitions, and homologous recombination deficiency-loss-of-heterozygosity following methods previously described [50–52].

### **Statistical analysis**

Fisher's exact test was used to examine frequency differences in germline mutations between SDBC and IBC. A two-tailed *t*-test (1x1 table), Fisher's exact test (2x2 tables) or Chi-square test (3x2 or greater tables) was applied to compare clinic-pathological differences between SDBC and IBC. The Kolmogorov-Smirnov test was performed to compare cumulative PRS differences in cancer free controls, IBCs and SDBCs.

## Results

### Interval cancers exhibit a more aggressive phenotype

A total of 1146 Lifepool participants were identified from 2010 to 2018 as having being diagnosed with a breast malignancy where screening detection status was known (**Table 1**), and comprised 930 invasive breast carcinomas [726 (78.1%) screen-detected and 204 (21.9%) interval] and 216 *in situ* cancers [189 (87.5%) screen-detected and 27 (12.5%) interval]. Women diagnosed with an invasive IBC had a significantly lower mean age (61.1 [range 44–80]) compared with SDBC (63.1 [range 40–89]) ( $p=0.0009$ ). Invasive IBCs were significantly associated with larger tumour size ( $p<0.0001$ ), higher grade ( $p<0.0001$ ), higher proliferation index ( $p<0.0001$ ), and nodal involvement ( $p=0.0003$ ) (**Table 1**).

The intrinsic subtype distribution (according to St Gallen classification) [53] but not histological subtype differed significantly between SDBC and IBC ( $p<0.0001$ ) (**Table 1, supplementary material, Table S2**). Luminal type cancers were more frequent in SDBC and subtypes associated with worse prognosis (ER-, Her2+ and TNBC) were more frequent among IBCs (**Table 1, supplementary material, Table S3**). The number of women diagnosed with either a SDBC or IBC with a self-reported strong family history of breast/ovarian cancer (28.4% versus 26.5%) or reporting a first degree relative with breast cancer (12% versus 15.7%) was not statistically different ( $p=0.659$  and  $0.1913$ , respectively).

## Frequency of pathogenic germline mutations in *HBOC* genes

Germline DNA was available from 442 participants diagnosed with BC (353 invasive and 89 *in situ*) and these were sequenced for pathogenic mutations in 12 HBOC genes (*BRCA1*, *BRCA2*, *PALB2*, *ATM*, *CHEK2*, *BRIP1*, *RAD51C*, *RAD51D*, *CDH1*, *PTEN*, *STK11* and *TP53*). Eight pathogenic mutations were identified among the invasive BCs (2.27%; 4 *BRCA2*, 1 *PALB2*, 1 *ATM* and 2 *CHEK2*) (**Table 2**) and four among *in situ* BCs (4.5%, 3 *BRCA2* and 1 *CHEK2*) (**supplementary material, Table S4**). The mutation carrier frequency was higher among the invasive IBC cases compared with the SDBC (3.90% versus 1.81%) but was not statistically significant ( $p=0.278$ ). Details of all germline mutations identified and associated clinicopathological data are described in **supplementary material, Table S5**.

PRS scores were generated for SDBC ( $n = 268$ ) and IBC ( $n = 76$ ) using 69 breast cancer SNPs and normalized against cancer-free controls ( $n = 4857$ ). As expected, the combined BC cases (SDBC and IBC) showed a significantly higher PRS compared with cancer-free controls ( $p<0.0001$ ). In a case-only analysis, IBCs were found to be associated with lower PRS than the SDBC, which was not significantly divergent from the control cohort ( $p = 0.1057$ ); whilst the SDBC cohort PRS was significantly higher than controls ( $p<0.0001$ ) (**supplementary material, Figure S1**).

## Differences in somatic aberrations

Targeted full-exon sequencing of known breast cancer driver genes was performed on tumour DNA from all available IBC (n=46) and selected SDBC (n=64) that were matched for age, grade, intrinsic subtype and histological subtype (**Table 3**) with the exception of a numerically higher number of TN and lobular breast cancers in the IBC cohort (although not statistically significant). Collectively, there were no significant differences in the average number of somatic coding mutations per tumour in IBC compared with SDBC (8.7 [range 1–42] versus 8.5 [range 1–15],  $p = 0.1605$ ), and this was not different when comparing within the luminal ( $p=0.9797$ ) or TN ( $p=0.2473$ ) subtypes (**supplementary material, Figure S2**). One IBC carried a large number of somatic mutations (42 mutations detected), which may be attributed to the occurrence of a somatic frameshift mutation in *MLH3* (p.Pro1292Glnfs\*32) that was accompanied by loss of the wild-type allele. The most commonly mutated breast cancer driver genes included *PIK3CA*, *TP53*, *KMT2C*, *CBFB*, *GATA3*, *CDH1*, *MAP3K1* and *TBX4* (**Table 4**). There were no significant differences in the frequency of mutations in individual genes between IBC and SDBC either as a whole or when analysed according to intrinsic (TN and luminal subtype) and histological subtypes (ductal and lobular).

Analysis of genome-wide CN changes did not identify any significant differences between IBCs and SDBCs with the exception of a higher frequency of loss of chromosome Xq in the TN SDBCs (**Figure 2A–C**), although the number of cases was small. The HRD sum score was significantly higher in the IBC cohort compared to the SDBC cohort (mean score 37.5 versus 28.2,  $p=0.0211$ , **Figure 2D**).

The HRD sum score was significantly higher when analysed as luminal (mean score 34.1 versus 25.8,  $p=0.0374$ , **Figure 2E**) but not TN (mean score 59.3 versus 60.0,  $p=0.9433$ , **Figure 2F**) subtype

Additionally, the frequency of IBC and SDBC with either a clinically high ( $\geq 42$ ) or low ( $< 42$ ) HRD sum score showed no significant differences when comparing the entire cohort, luminal and TN subtypes (**Figure 2G-I**).

## Discussion

While there have been numerous studies comparing the clinicopathological characteristics of SDBC and IBC, few have explored the molecular landscape of these cancers. In this study we report the largest and most detailed germline and somatic genetic analysis of IBC and SDBC. The proportion of cases and clinical features of IBCs compared with SDBC within Lifepool is broadly consistent with data reported in other studies (**Table 5**). Specifically, IBC represented 21.9% of invasive BC diagnoses (and 12.5% with *in situ* diagnoses), and cancers had clinical features associated with more aggressive disease, such as higher grade, more node positivity and a higher rate of TN and Her2+ intrinsic subtypes diagnosis when compared with SDBC.

While we performed radiological review of the prior mammograms to exclude missed cancers, some cases remaining in our analysis may have been occult on the screening mammogram due to dense tissue, and this is more likely for younger women who on average have higher mammographic density. This may have contributed to our finding that the average age at diagnosis was lower for IBCs than SDBC.

The lack of family history of BC association was consistent with most studies (**Table 5**) and this persisted when using only the number of first degree relatives with BC. Contrary to these findings, the largest familial study comparing IBC and SDBC to date (family history of BC in mother and sister in 538 IBCs and 1263 SDBC)

found a small but statistically significant association of family history with IBC after applying an age-adjusted model [4]. The lack of consistency in the various studies could be due to differing definitions of family history, a lack of accurate family history data or might reflect statistical variance around a small effect size.

Among the entire cohort with invasive BC, 2.27% carried actionable *HBOC* mutations that would initiate an offer of referral to a specialised familial cancer centre. This number is lower, although not statistically different from the 4.6% described by Tung *et al* (2016) in women >60 years of age from a hospital-based study of BC unselected for family history [54]. The higher frequency of *HBOC* mutations reported by Tung *et al* (2016) could be partly explained as they obtained samples through a specialist cancer hospital-based cohort. This is reflective in their data, as compared to our cohort, they report a higher proportion of TNBCs (17.8% compared to 7%), and a larger proportion of patients with a strong family history of breast cancer (18.6% compared to 12.8%).

It has been documented that *BRCA1/2* carriers who are part of a screening program are highly prone to developing an IBC [55–59], and while a higher frequency of *HBOC* mutations was observed in IBC versus SDBC in our study (3.90% versus 1.81%) this was not statistically significant. However, since the peak breast cancer incidence for *BRCA1* and *BRCA2* carriers occurs in the 41–50 and 51–60 age group respectively [60], carriers of these mutations will have been depleted from the Lifepool study cohort given the median age at recruitment is 60.63.

Consistent with this we did not identify any *BRCA1* mutation carriers among either the SDBC or IBC. Furthermore, it has been noted in these studies that patients who went on to develop an IBC were younger than women detected by mammographic screening; and their breast tissue described as 'dense' [55–59]. This supports the observation that density limits the sensitivity of imaging, whereas these masking effects would be reduced in our older cohort.

It is possible that some IBC cases were identified through 'pseudo-screening', where women accessed screening services outside the BreastScreen program. This is known to occur in Australia [61,62] and would be expected to attenuate differences between our SDBC and IBC cases.

We explored whether the PRS could discriminate between SDBC and IBC, as this could potentially be used to identify sub-groups of women who would benefit from a tailored screening modality. As expected, the BC cases as a whole had significantly higher PRS than cancer-free controls. Interestingly, the IBCs had a lower PRS than the SDBC which is consistent with a previous report [25] and likely reflects the fact that the SNPs used to generate the PRS are more predictive for ER+ breast cancer [63], which were more prevalent in SDBC than IBC (92.8% versus 74%,  $p < 0.0001$ ).

Because IBCs are generally more aggressive and associated with poorer outcome than SDBC, this may reflect a fundamental difference in molecular aetiology. Whilst the clinicopathological data confirmed that IBCs are enriched for

more aggressive phenotypes, when matched for age, grade, intrinsic subtypes and histological subtype there were no significant differences in the somatic genetic alterations between the two groups. Differences in IBCs reported by others, in particular a higher frequency of *TP53* mutations (**Table 5**), are largely explained by the higher proportion of TNBC and Her2+ tumours in which *TP53* mutations are more common. The only potential difference between the IBC and SDBC was more frequent loss of chr Xq in the TN SDBC (0/8 IBC versus 5/5 SDBC), but the numbers were small.

A limitation of the somatic analysis is the slightly higher proportion of TN and lobular BCs in the IBC cohort, due to the lack of suitable matching cases in the SDBC. Additionally, the number of Her2-type IBCs analysed was small, so the question as to whether they harbour unique copy number aberrations or somatic driver mutations compared to Her2-type SBDCs remains unanswered.

HRD scores  $\geq 42$  have shown clinical utility in predicting the likelihood of response to neoadjuvant platinum-containing therapy such as PARP inhibition [64,65]. Whilst IBC cases demonstrated a significantly higher mean HRD score, when categorised into the potentially clinically relevant high ( $\geq 42$ ) or low HRD ( $< 42$ ) categories there was no significant difference. This result indicates that treatment strategies for IBC and SDBC patients with DNA damaging therapeutics are likely to be similar.

Two scenarios of IBC development have been postulated when compared to SDBC of the same intrinsic subtype (**Figure 3**). Scenario 1 infers IBCs are a consequence of the convergence of tumour initiation timing and mammographic screening timing/technology; and Scenario 2 postulates that IBCs comprise a significantly different molecular class of BC. Our data is consistent with the former scenario, given our analysis revealed no major molecular divergence in the IBC cohort when matched for clinical features that include age at diagnosis, tumour grade, histological subtype and hormone status. While this study is the most comprehensive molecular analysis of IBC to date, the overall numbers are small and this data needs to be interpreted cautiously and will require validation in independent cohorts.

## **Acknowledgments**

The authors would like to acknowledge the Bioinformatics and Molecular Genomics Core Facilities of the Peter MacCallum Cancer Centre. This work would not have been possible without financial support by the Victorian Cancer Agency (ECSG15003, D.C), the National Breast Cancer Foundation (IF-15-004 & CG-08-02, I.G.C), the NHMRC (APP1079329 SF) and Cancer Australia (1084963, I.G.C). The funders had no role in the design and execution of this study.

## **Author contributions statement**

DC, KLG and IGC contributed to the study design; DC, NL, SMR, KE, GBM, TK, TS, LD, GL, SH, HS, KMF, KK, JC, PAJ, and KLG contributed to data collection; DC, NL, SMR, KE, GBM, MZ, LD, KLG and IGC performed data analysis and interpretation; DC and KLG performed the literature search for the manuscript; DC generated the figures, tables and drafted the manuscript, which was then extensively edited by KLG and IGC. All other authors read and commented on the manuscript and approved the final version.

## References

1. Rickard MT, Taylor RJ, Fazli MA, *et al.* Interval breast cancers in an Australian mammographic screening program. *Med J Aust* 1998; **169**: 184-187.
2. Rojo F, Domingo L, Sala M, *et al.* Gene expression profiling in true interval breast cancer reveals overactivation of the mTOR signaling pathway. *Cancer Epidemiol Biomarkers Prev* 2014; **23**: 288-299.
3. Eriksson L, Czene K, Rosenberg LU, *et al.* Mammographic density and survival in interval breast cancers. *Breast Cancer Res* 2013; **15**: R48.
4. Holm J, Humphreys K, Li J, *et al.* Risk factors and tumor characteristics of interval cancers by mammographic density. *J Clin Oncol* 2015; **33**: 1030-1037.
5. Blanch J, Sala M, Ibanez J, *et al.* Impact of risk factors on different interval cancer subtypes in a population-based breast cancer screening programme. *PLoS One* 2014; **9**: e110207.
6. Payne JJ, Caines JS, Gallant J, *et al.* A review of interval breast cancers diagnosed among participants of the Nova Scotia Breast Screening Program. *Radiology* 2013; **266**: 96-103.
7. Porter PL, El-Bastawissi AY, Mandelson MT, *et al.* Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 1999; **91**: 2020-2028.
8. Mandelson MT, Oestreicher N, Porter PL, *et al.* Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000; **92**: 1081-1087.
9. Li J, Ivansson E, Klevebring D, *et al.* Molecular Differences between Screen-Detected and Interval Breast Cancers Are Largely Explained by PAM50 Subtypes. *Clin Cancer Res* 2017; **23**: 2584-2592.
10. Brekelmans CT, Peeters PH, Faber JA, *et al.* The epidemiological profile of women with an interval cancer in the DOM screening programme. *Breast Cancer Res Treat* 1994; **30**: 223-232.
11. Bellio G, Marion R, Giudici F, *et al.* Interval Breast Cancer Versus Screen-Detected Cancer: Comparison of Clinicopathologic Characteristics in a Single-Center Analysis. *Clin Breast Cancer* 2017; **17**: 564-571.
12. Collett K, Stefansson IM, Eide J, *et al.* A basal epithelial phenotype is more frequent in interval breast cancers compared with screen detected tumors. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 1108-1112.
13. Meshkat B, Prichard RS, Al-Hilli Z, *et al.* A comparison of clinical-pathological characteristics between symptomatic and interval breast cancer. *Breast* 2015; **24**: 278-282.
14. Caldarella A, Puliti D, Crocetti E, *et al.* Biological characteristics of interval cancers: a role for biomarkers in the breast cancer screening. *J Cancer Res Clin Oncol* 2013; **139**: 181-185.
15. Cowan WK, Angus B, Gray JC, *et al.* A study of interval breast cancer within the NHS breast screening programme. *J Clin Pathol* 2000; **53**: 140-146.
16. Crosier M, Scott D, Wilson RG, *et al.* Differences in Ki67 and c-erbB2 expression between screen-detected and true interval breast cancers. *Clin Cancer Res* 1999; **5**: 2682-2688.
17. Palka I, Kelemen G, Ormandi K, *et al.* Tumor characteristics in screen-detected and symptomatic breast cancers. *Pathol Oncol Res* 2008; **14**: 161-167.

18. Musolino A, Michiara M, Conti GM, *et al.* Human epidermal growth factor receptor 2 status and interval breast cancer in a population-based cancer registry study. *J Clin Oncol* 2012; **30**: 2362-2368.
19. Lowery JT, Byers T, Hokanson JE, *et al.* Complementary approaches to assessing risk factors for interval breast cancer. *Cancer Causes Control* 2011; **22**: 23-31.
20. Anttinen J, Kuopio T, Nykanen M, *et al.* Her-2/neu oncogene amplification and protein over-expression in interval and screen-detected breast cancers. *Anticancer Res* 2003; **23**: 4213-4218.
21. Domingo L, Sala M, Servitja S, *et al.* Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes Control* 2010; **21**: 1155-1164.
22. Kirsh VA, Chiarelli AM, Edwards SA, *et al.* Tumor characteristics associated with mammographic detection of breast cancer in the Ontario breast screening program. *J Natl Cancer Inst* 2011; **103**: 942-950.
23. Chen X, Zhu S, Fei X, *et al.* Surgery time interval and molecular subtype may influence Ki67 change after core needle biopsy in breast cancer patients. *BMC Cancer* 2015; **15**: 822.
24. Gilliland FD, Joste N, Stauber PM, *et al.* Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2000; **92**: 743-749.
25. Li J, Holm J, Bergh J, *et al.* Breast cancer genetic risk profile is differentially associated with interval and screen-detected breast cancers. *Ann Oncol* 2015; **26**: 517-522.
26. Goncalves R, Warner WA, Luo J, *et al.* New concepts in breast cancer genomics and genetics. *Breast Cancer Res* 2014; **16**: 460.
27. Mook S, Van 't Veer LJ, Rutgers EJ, *et al.* Independent prognostic value of screen detection in invasive breast cancer. *J Natl Cancer Inst* 2011; **103**: 585-597.
28. Shen Y, Yang Y, Inoue LY, *et al.* Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *J Natl Cancer Inst* 2005; **97**: 1195-1203.
29. Kavanagh AM, Mitchell H, Farrugia H, *et al.* Monitoring interval cancers in an Australian mammographic screening programme. *J Med Screen* 1999; **6**: 139-143.
30. Lee JEA, Li N, Rowley SM, *et al.* Molecular analysis of PALB2-associated breast cancers. *J Pathol* 2018; **245**: 53-60.
31. Li N, Rowley SM, Thompson ER, *et al.* Evaluating the breast cancer predisposition role of rare variants in genes associated with low-penetrance breast cancer risk SNPs. *Breast Cancer Res* 2018; **20**: 3.
32. Li N, Rowley SM, Goode DL, *et al.* Mutations in RECQL are not associated with breast cancer risk in an Australian population. *Nat Genet* 2018.
33. Rowley SM, Mascarenhas L, Devereux L, *et al.* Population-based genetic testing of asymptomatic women for breast and ovarian cancer susceptibility. *Genet Med* 2018.
34. Pereira B, Chin S-F, Rueda OM, *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* 2016; **7**: 11479.
35. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013.
36. 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.

37. 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.
38. 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.
39. McLaren W, Gil L, Hunt SE, *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 2016; **17**: 122.
40. 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.
41. 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.
42. Li N, Thompson ER, Rowley SM, *et al.* Reevaluation of RINT1 as a breast cancer predisposition gene. *Breast Cancer Res Treat* 2016; **159**: 385-392.
43. Rehm HL, Harrison SM, Martin CL. ClinVar Is a Critical Resource to Advance Variant Interpretation. *Oncologist* 2017; **22**: 1562.
44. Lek M, Karczewski KJ, Minikel EV, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536**: 285-291.
45. Balbo Pogliano C, Gatti M, Rüttemann P, *et al.* ASH1L histone methyltransferase regulates the handoff between damage recognition factors in global-genome nucleotide excision repair. *Nat Commun* 2017; **8**: 1333.
46. Genomes Project C, Auton A, Brooks LD, *et al.* A global reference for human genetic variation. *Nature* 2015; **526**: 68-74.
47. Kuilman T, Velds A, Kemper K, *et al.* CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol* 2015; **16**: 49.
48. Mavaddat N, Pharoah PD, Michailidou K, *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 2015; **107**.
49. 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.
50. Birkbak 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.
51. 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.
52. 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.
53. Coates AS, Winer EP, Goldhirsch A, *et al.* Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015; **26**: 1533-1546.
54. Tung N, Lin NU, Kidd J, *et al.* Frequency of Germline Mutations in 25 Cancer Susceptibility Genes in a Sequential Series of Patients With Breast Cancer. *J Clin Oncol* 2016; **34**: 1460-1468.
55. Komenaka IK, Ditkoff BA, Joseph KA, *et al.* The development of interval breast malignancies in patients with BRCA mutations. *Cancer* 2004; **100**: 2079-2083.

56. Meijers-Heijboer H, van Geel B, van Putten WL, *et al.* Breast cancer after prophylactic bilateral mastectomy in women with a BRCA1 or BRCA2 mutation. *N Engl J Med* 2001; **345**: 159-164.
57. Nelson HD, Huffman LH, Fu R, *et al.* Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2005; **143**: 362-379.
58. Robson M, Offit K. Clinical practice. Management of an inherited predisposition to breast cancer. *N Engl J Med* 2007; **357**: 154-162.
59. Scheuer L, Kauff N, Robson M, *et al.* Outcome of preventive surgery and screening for breast and ovarian cancer in BRCA mutation carriers. *J Clin Oncol* 2002; **20**: 1260-1268.
60. Landrum MJ, Lee JM, Benson M, *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* 2017.
61. Giles GG, Amos A. Evaluation of the organised mammographic screening programme in Australia. *Ann Oncol* 2003; **14**: 1209-1211.
62. BreastScreen Australia Evaluation Taskforce. BreastScreen Australia Evaluation. Evaluation final report. Canberra: Australian Government Department of Health and Ageing; 2009 Jun. Report No.: Screening Monograph No.1/2009. Available from: <http://cancerscreening.gov.au/internet/screening/publishing.nsf/Content/programme-evaluation>.  
. 2009.
63. Michailidou K, Lindstrom S, Dennis J, *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017; **551**: 92-94.
64. 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.
65. Diossy M, Reiniger L, Sztupinski Z, *et al.* Breast cancer brain metastases show increased levels of genomic aberration-based homologous recombination deficiency scores relative to their corresponding primary tumors. *Ann Oncol* 2018; **29**: 1948-1954.

SUPPLEMENTARY MATERIAL ONLINE

**Supplementary materials and methods** NO

**Supplementary figure legends** NO because the legends are embedded in the Word file

**Figure S1.** Distribution of polygenic risk score

**Figure S2.** Spectrum of somatic mutations in IBC and SDBC according to subtype

**Table S1.** Genes analysed across somatic and germline sequencing panels

**Table S2.** Breast cancer subtypes diagnosed in the Lifepool cohort.

**Table S3.** Cancer subtype classification according to hormone and human epidermal growth factor receptor 2 status

**Table S4.** Germline mutations in Lifepool women diagnosed with in situ breast cancer

**Table S5.** Clinico-pathological and sequencing information for germline pathogenic mutations identified

## Legends and Tables

Figure 1. Summary of the study cohort.

Figure 2. Copy number changes in interval and screen detected breast cancers.

Copy number aberrations are shown for 46 IBCs and 73 SDBC as the frequency of copy number changes in (A) entire cohort, (B) luminal subtype and (C) triple negative breast cancer subtype. The region highlighted with a dark black box represents a significant copy change between IBCs and SDBC (thresholds of  $p < 0.05$  and at least 25% frequency difference). Copy number profiles were used to generate a HRD-sum score for (D) entire cohort, (E) luminal subtype and (F) triple negative breast cancer subtype. The frequency of IBC and SDBC with a clinically high HRD score ( $\geq 42$ ) was also calculated for (G) entire cohort, (H) luminal subtype and (I) triple negative breast cancer subtype. Mann-Whitney tests were applied to HRD sum score calculations and Fisher's exact test was applied to HRD frequency plots.

Figure 3. Development of interval breast cancers compared to screen detected breast cancer.

Table 1. Clinicopathological features of invasive screen detected and interval breast cancers in the Lifepool dataset.

| Characteristics                                | SDBC        | IBC         | p value              |
|--|-------------|-------------|----------------------|
| <b>Behaviour</b>                               |             |             |                      |
| Invasive                                       | 726 (78.1%) | 204 (21.9%) |                      |
| <i>In situ</i>                                 | 189 (87.5%) | 27 (12.5%)  |                      |
| <b>Invasive only</b>                           |             |             |                      |
| <b>Age at diagnosis</b>                        |             |             |                      |
| Mean ± SD                                      | 63.1 ± 7.4  | 61.1 ± 7.4  | 0.0009 <sup>a</sup>  |
| Median   | 63.3        | 60.8        |                      |
| Range  | 40-89       | 44-80       |                      |
| <b>Tumour Size</b>                             |             |             |                      |
| <20 mm   | 398 (61%)   | 65 (37.1%)  | <0.0001 <sup>b</sup> |
| 20-49 mm                                       | 220 (33.7%) | 86 (49.1%)  |                      |
| ≥50  | 34 (5.2%)   | 24 (13.7%)  |                      |
| NA   | 74          | 29          |                      |
| <b>Intrinsic Subtype</b>                       |             |             |                      |
| TNBC   | 34 (5.2%)   | 31 (16.9%)  | <0.0001 <sup>b</sup> |
| ER-, Her2+                                     | 9 (1.4%)    | 14 (7.7%)   |                      |
| Luminal Her2+                                  | 45 (6.9%)   | 17 (9.3%)   |                      |
| Luminal  | 564 (86.5%) | 121 (66.1%) |                      |
| NA   | 74          | 21          |                      |
| <b>Invasive cancer histology subtype</b>       |             |             |                      |
| Ductal   | 553 (76.2%) | 160 (78.4%) | 0.0615 <sup>b</sup>  |
| Lobular  | 80 (11.%)   | 29 (14.2%)  |                      |
| Other  | 93 (12.8%)  | 15 (7.4%)   |                      |
| <b>Tumour grade</b>                            |             |             |                      |
| 1  | 193 (29%)   | 21 (11.6%)  | <0.0001 <sup>b</sup> |
| 2  | 324 (48.6%) | 77 (42.5%)  |                      |
| 3  | 149 (22.4%) | 83 (45.9%)  |                      |
| NA   | 60          | 23          |                      |
| <b>Nodal status</b>                            |             |             |                      |
| Positive                                       | 117 (24.2%) | 52 (40.6%)  | 0.0003 <sup>c</sup>  |
| Negative                                       | 367 (75.6%) | 76 (59.4%)  |                      |
| NA   | 242         | 76          |                      |
| <b>Proliferation index (Ki67)</b>              |             |             |                      |
| High (≥20%)                                    | 101 (26.2%) | 45 (49.5%)  | <0.0001 <sup>c</sup> |
| Low (<20%)                                     | 284 (73.8%) | 46 (50.5%)  |                      |
| NA   | 341         | 113         |                      |
| <b>Laterality</b>                              |             |             |                      |
| Left   | 364 (52%)   | 103 (51.2%) | 0.8729 <sup>c</sup>  |
| Right  | 336 (48%)   | 98 (48.8%)  |                      |
| NA   | 26          | 3           |                      |
| <b>1st degree relatives with breast cancer</b> |             |             |                      |
| Yes  | 206 (28.4%) | 54 (26.5%)  | 0.6590 <sup>c</sup>  |

|   |             |             |                     |
|---|-------------|-------------|---------------------|
| No  | 520 (71.6%) | 150 (73.5%) |                     |
| <b>Strong family history of breast cancer</b> |             |             |                     |
| Yes   | 87 (12%)    | 32 (15.7%)  | 0.1913 <sup>c</sup> |
| No  | 639 (88%)   | 172 (84.3%) |                     |

Calculation of percentage within the cohort is shown within parentheses. Strong family history of breast cancer was defined as either three 1st or 2nd degree relatives with breast cancer; two 1st or 2nd degree relatives with breast cancer (one diagnosed <50 years); male breast cancer any age; or ovarian cancer any age. NA, data not available; a, two-tailed  $\chi^2$ -test was applied; b, Chi-square test was applied; c, Fisher's exact test.

**Table 2. Contribution of germline mutations in Lifepool women diagnosed with an invasive breast cancer.**

| Breast cancer panel genes | Total (n=353)    | IBC (n=77)       | SDBC (n=276)     | p value      | TNBC (n=23) |
|---------------------------|------------------|------------------|------------------|--------------|-------------|
| <i>BRCA2</i> LoF          | 4 (1.13%)        | 2 (2.6%)         | 2 (0.72%)        | 0.170        | 1 (4.35%)   |
| <i>PALB2</i> LoF          | 1 (0.28%)        | 1 (1.3%)         | -                | -            | 1 (4.35%)   |
| <i>ATM</i> LoF            | 1 (0.28%)        | -                | 1 (0.36%)        | -            | -           |
| <i>CHEK2</i> c.1100delC   | 2 (0.57%)        | -                | 2 (0.72%)        | -            | -           |
|                           | <b>8 (2.27%)</b> | <b>3 (3.90%)</b> | <b>5 (1.81%)</b> | <b>0.278</b> |             |

Calculation of percentage within the cohort and screening group is shown within parentheses. Two-tailed Fisher's exact tests have been calculated for each gene comparing IBC to the SDBC. No pathogenic variants were observed in screened HBOC genes *BRCA1*, *BRIP1*, *RAD51C*, *RAD51D*, *CDH1*, *ATM* (c.7271T>G), *PTEN*, *STK11* and *TP53*. LoF, loss of function

**Table 3. Breast tumours sequenced on the Agilent Sureselect panel compared for age, intrinsic subtype, grade and histological subtype .**

| Characteristics             | SDBC (n=64)  | IBC (n=46)   | p value             |
|-----------------------------|--------------|--------------|---------------------|
| <b>Age at diagnosis</b>     |              |              |                     |
| Mean $\pm$ SD               | 61.9 $\pm$ 8 | 61.2 $\pm$ 8 | 0.8731 <sup>a</sup> |
| Median                      | 60.9         | 60.1         |                     |
| Range                       | 42-80        | 45-80        |                     |
| <b>Intrinsic subtype</b>    |              |              |                     |
| ER-, Her2+                  | 1 (1.6%)     | 0            | 0.3770 <sup>b</sup> |
| Luminal Her2+               | 4 (6.3%)     | 2 (4.3%)     |                     |
| Luminal                     | 54 (84.4%)   | 36 (78.3%)   |                     |
| TNBC                        | 5 (7.8%)     | 8 (17.4%)    |                     |
| <b>Tumour grade</b>         |              |              |                     |
| 1                           | 14 (21.9%)   | 8 (17.8%)    | 0.5606 <sup>b</sup> |
| 2                           | 32 (50%)     | 20 (44.4%)   |                     |
| 3                           | 18 (28.1%)   | 17 (37.8%)   |                     |
| NA                          | 0            | 1            |                     |
| <b>Histological subtype</b> |              |              |                     |
| Ductal                      | 56 (87.5%)   | 36 (78.3%)   | 0.2749 <sup>b</sup> |
| Lobular                     | 3 (4.7%)     | 6 (13%)      |                     |
| Other                       | 5 (7.8%)     | 4 (8.7%)     |                     |

NA, clinical data not available; a, two-tailed *t*-test was applied; b, Chi-square test was applied.

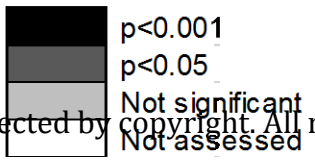


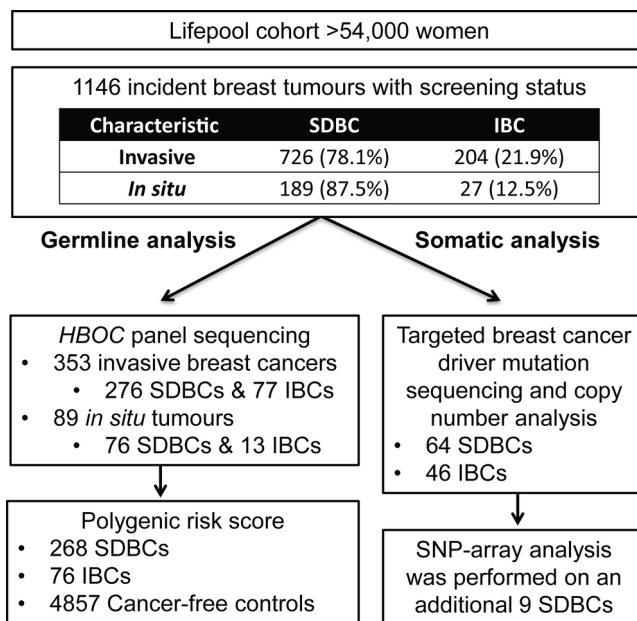
**Table 4. Breast cancer somatic driver mutation profile in interval and screen-detected breast cancers**

| Gene   | Entire cohort<br>(IBC n=46, SDBC n=64) |            |         | Luminal subtype<br>(IBC n=38, SDBC n=58) |            |         | TNBC subtype<br>(IBC n=8, SDBC n=5) |            |         | Ductal breast cancers<br>(IBC n=36, SDBC n=56) |            |         | Lobular breast cancers<br>(IBC n=6, SDBC n=3) |            |         |
|--------|--|------------|---------|--|------------|---------|-------------------------------------|------------|---------|--|------------|---------|---|------------|---------|
|        | IBC n (%)                              | SDBC n (%) | p value | IBC n (%)                                | SDBC n (%) | p value | IBC n (%)                           | SDBC n (%) | p value | IBC n (%)                                      | SDBC n (%) | p value | IBC n (%)                                     | SDBC n (%) | p value |
| PIK3CA | 21 (45.7%)                             | 24 (37.5%) | 0.393   | 18 (47.4%)                               | 23 (39.7%) | 0.457   | 3 (37.5%)                           | 1 (20%)    | 0.523   | 15 (41.7%)                                     | 20 (35.7%) | 0.568   | 5 (83.3%)                                     | 2 (66.7%)  | <0.999  |
| TP53   | 9 (19.6%)                              | 14 (21.9%) | 0.770   | 3 (7.9%)                                 | 10 (17.2%) | 0.193   | 6 (75%)                             | 4 (80%)    | 0.841   | 7 (19.4%)                                      | 13 (23.2%) | 0.670   | 1 (16.7%)                                     | 0          | <0.999  |
| KMT2C  | 7 (15.2%)                              | 13 (20.3%) | 0.496   | 6 (15.8%)                                | 13 (22.4%) | 0.428   | 1 (12.5%)                           | 0          | <0.999  | 3 (8.3%)                                       | 12 (21.4%) | 0.099   | 3 (50%)                                       | 0          | 0.464   |
| CBFB   | 7 (15.2%)                              | 10 (15.6%) | 0.954   | 8 (21.1%)                                | 10 (17.2%) | 0.642   | 0                                   | 0          | -       | 6 (16.7%)                                      | 9 (16.1%)  | 0.940   | 2 (33.3%)                                     | 0          | 0.500   |
| GATA3  | 7 (15.2%)                              | 8 (12.5%)  | 0.683   | 7 (8.4%)                                 | 8 (13.8%)  | 0.543   | 0                                   | 0          | -       | 6 (16.7%)                                      | 6 (10.7%)  | 0.411   | 0   | 0          | -       |
| CDH1   | 7 (15.2%)                              | 8 (12.5%)  | 0.683   | 6 (15.8%)                                | 7 (12.1%)  | 0.604   | 1 (12.5%)                           | 1 (20%)    | <0.999  | 1 (2.8%)                                       | 5 (8.9%)   | 0.246   | 6 (100%)                                      | 3 (100%)   | <0.999  |
| MAP3K1 | 5 (10.9%)                              | 8 (12.5%)  | 0.795   | 4 (10.5%)                                | 8 (13.8%)  | 0.638   | 1 (12.5%)                           | 0          | <0.999  | 3 (8.3%)                                       | 7 (12.5%)  | 0.533   | 1 (16.7%)                                     | 0          | <0.999  |
| TBX3   | 3 (6.5%)                               | 5 (7.8%)   | 0.798   | 3 (7.9%)                                 | 5 (8.6%)   | 0.900   | 0                                   | 0          | -       | 2 (5.6%)                                       | 4 (7.1%)   | 0.765   | 1 (16.7%)                                     | 1 (33.3%)  | <0.999  |
| NCOR1  | 6 (13%)                                | 3 (4.7%)   | 0.116   | 5 (13.2%)                                | 3 (5.2%)   | 0.168   | 1 (12.5%)                           | 0          | <0.999  | 5 (13.9%)                                      | 2 (3.6%)   | 0.070   | 1 (16.7%)                                     | 1 (33.3%)  | <0.999  |
| PTEN   | 2 (4.3%)                               | 5 (7.8%)   | 0.465   | 2 (5.3%)                                 | 5 (8.6%)   | 0.538   | 0                                   | 0          | -       | 2 (5.6%)                                       | 5 (8.9%)   | 0.554   | 0   | 0          | -       |
| ARID1A | 2 (4.3%)                               | 6 (9.4%)   | 0.319   | 2 (5.3%)                                 | 6 (10.3%)  | 0.381   | 0                                   | 0          | -       | 1 (2.8%)                                       | 4 (7.1%)   | 0.370   | 1 (16.7%)                                     | 1 (33.3%)  | <0.999  |
| AKT1   | 2 (4.3%)                               | 3 (4.7%)   | 0.933   | 2 (5.3%)                                 | 3 (5.2%)   | 0.984   | 0                                   | 0          | -       | 2 (5.6%)                                       | 3 (5.4%)   | 0.967   | 0   | 0          | -       |
| BRCA1  | 2 (4.3%)                               | 2 (3.1%)   | 0.737   | 1 (2.6%)                                 | 1 (1.7%)   | <0.999  | 1 (12.5%)                           | 1 (20%)    | <0.999  | 2 (5.6%)                                       | 2 (3.6%)   | 0.651   | 0   | 0          | -       |
| RUNX1  | 0                                      | 3 (4.7%)   | 0.138   | 0  | 3 (5.2%)   | 0.156   | 0                                   | 0          | -       | 0  | 3 (5.4%)   | 0.160   | 0   | 0          | -       |
| BRCA2  | 0                                      | 3 (4.7%)   | 0.138   | 0  | 3 (5.2%)   | 0.156   | 0                                   | 0          | -       | 0  | 3 (5.4%)   | 0.160   | 0   | 0          | -       |
| NF1    | 2 (4.3%)                               | 1 (1.6%)   | 0.379   | 1 (2.6%)                                 | 1 (1.7%)   | <0.999  | 1 (12.5%)                           | 0          | <0.999  | 2 (5.6%)                                       | 0          | 0.151   | 0   | 0          | -       |
| MAP2K4 | 2 (4.3%)                               | 1 (1.6%)   | 0.379   | 2 (5.3%)                                 | 1 (1.7%)   | 0.332   | 0                                   | 0          | -       | 2 (5.6%)                                       | 1 (1.8%)   | 0.323   | 0   | 0          | -       |
| PIK3R1 | 2 (4.3%)                               | 0          | -       | 1 (2.6%)                                 | 0          | 0.396   | 1 (12.5%)                           | 0          | <0.999  | 2 (5.6%)                                       | 0          | 0.151   | 0   | 0          | -       |
| SF3B1  | 0                                      | 1 (1.6%)   | -       | 0  | 1 (1.7%)   | <0.999  | 0                                   | 0          | -       | 0  | 1 (1.8%)   | <0.999  | 0   | 0          | -       |
| FOXP1  | 0                                      | 1 (1.6%)   | -       | 0  | 1 (1.7%)   | <0.999  | 0                                   | 0          | -       | 0  | 1 (1.8%)   | <0.999  | 0   | 0          | -       |

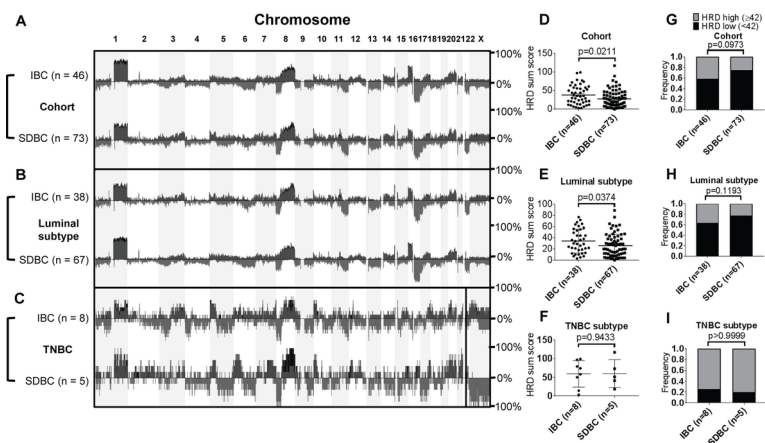
|        |          |   |   |          |   |       |   |   |   |          |   |       |   |   |   |
|--------|----------|---|---|----------|---|-------|---|---|---|----------|---|-------|---|---|---|
| MEN1   | 1 (2.2%) | 0 | - | 1 (2.6%) | 0 | 0.396 | 0 | 0 | - | 1 (2.8%) | 0 | 0.391 | 0 | 0 | - |
| CDKN2A | 1 (2.2%) | 0 | - | 1 (2.6%) | 0 | 0.396 | 0 | 0 | - | 1 (2.8%) | 0 | 0.391 | 0 | 0 | - |

| Author     | Year | Reference   | ER-   | PR-   | HER2 +ve | Ki67 high | Higher grade | Larger size | Nodal status +ve | Young age | TP53 +ve | CNAs  | Somatic mutations | Germline mutations | PRS   | Family history | # of IBCs |
|------------|------|-------------|-------|-------|----------|-----------|--------------|-------------|------------------|-----------|----------|-------|-------------------|--------------------|-------|----------------|-----------|
| Cheasley   | 2019 | This report | Black | Black | Black    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 201       |
| Holm       | 2015 | 4           | Black | Black | Black    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 1247      |
| Li         | 2015 | 25          | White | White | White    | White     | White        | White       | White            | White     | White    | White | White             | White              | White | White          | 782       |
| Blanch     | 2014 | 5           | White | White | White    | White     | White        | White       | White            | White     | White    | White | White             | White              | White | White          | 484       |
| Kirsh      | 2011 | 22          | Black | Black | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 288       |
| Eriksson   | 2013 | 3           | Black | Black | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 285       |
| Payne      | 2013 | 6           | Black | Grey  | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 233       |
| Porter     | 1999 | 7           | Black | Grey  | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 150       |
| Mandelson  | 2000 | 8           | White | White | White    | White     | White        | White       | Black            | Black     | Black    | Black | Black             | Black              | Black | Black          | 149       |
| Li         | 2017 | 9           | Black | Black | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 113       |
| Brekelmans | 1994 | 10          | Grey  | White | White    | White     | White        | White       | Black            | Black     | Black    | Black | Black             | Black              | Black | Black          | 107       |
| Bellio     | 2017 | 11          | White | White | White    | White     | White        | White       | Black            | Black     | Black    | Black | Black             | Black              | Black | Black          | 107       |
| Collett    | 2005 | 12          | Black | Black | White    | Black     | Black        | Black       | Black            | Black     | Black    | Black | Black             | Black              | Black | Black          | 95        |
| Chen       | 2015 | 23          | Black | Grey  | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 75        |
| Meshkat    | 2015 | 13          | Black | Black | White    | Black     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 71        |
| Caldarella | 2013 | 14          | Grey  | White | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 66        |
| Gilliland  | 2000 | 24          | Grey  | White | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 64        |
| Cowan      | 2000 | 15          | Grey  | White | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 57        |
| Crosier    | 1999 | 16          | Black | Grey  | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 51        |
| Palka      | 2008 | 17          | Grey  | White | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 48        |
| Musolino   | 2012 | 18          | Black | Grey  | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 48        |
| Lowery     | 2011 | 19          | Grey  | White | White    | Black     | Black        | Black       | Black            | White     | White    | White | White             | White              | White | White          | 47        |
| Anttinen   | 2003 | 20          | White | White | Black    | White     | White        | White       | White            | White     | White    | White | White             | White              | White | White          | 39        |
| Domingo    | 2010 | 21          | Black | Black | Grey     | White     | Black        | Black       | Black            | Grey      | Grey     | Grey  | Grey              | Grey               | Grey  | Grey           | 34        |

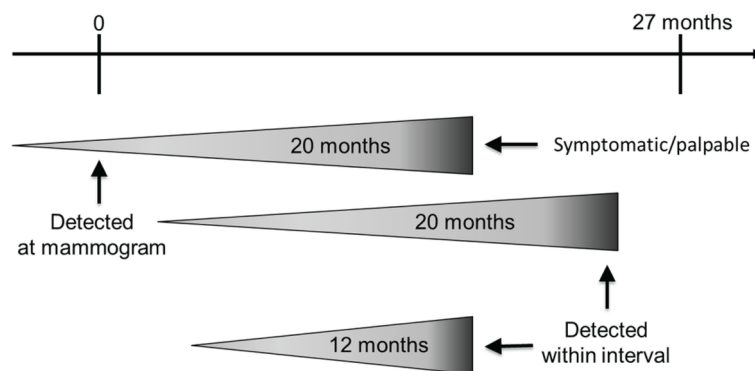




PATH\_5251\_Figure 1.tif



PATH\_5251\_Figure 2.tif



PATH\_5251\_Figure 3.tif