Current gene panels account for nearly all homologous recombination repair-associated multiple-case breast cancer families

Thibaut S. Matís1,2,3, Nadia Zayed1,2, Bouchra Labraki4, Manon de Ladurantaye5, Théophane A. Matis6, José Camacho Valenzuela7,8, Nancy Hamel2, Adrienne Atayan1, Barbara Rivera9,10, Yuval Tabach1,2,11, Alexandre Orthwein7,8, Anne-Marie Mes-Masson5,12, Zaki El Haffaf13, William D. Foulkes1,2,7,8,9, and Paz Polak3,15

It was hypothesized that variants in underexplored homologous recombination repair (HR) genes could explain unsolved multiple-case breast cancer (BC) families. We investigated HR deficiency (HRD)-associated mutational signatures and second hits in tumor DNA from familial BC cases. No candidates genes were associated with HRD in 38 probands previously tested negative with gene panels. We conclude it is unlikely that unknown HRD-associated genes explain a large fraction of unsolved familial BC.

Many multiple-case BC families remain unexplained by a pathogenic variant, despite routine clinical testing of an increasing number of recognized BC susceptibility genes (BCSGs). These cases pose a major clinical challenge for disease management of patients and cancer prevention.

To identify germline pathogenic variants (GPVs) in patients, Next-Generation Sequencing (NGS) of blood-derived DNA is used, often through pan-cancer panels containing up to 25 genes. The three major BCSGs – BRCA1, BRCA2, and PALB2 – are involved in homology-directed DNA repair (HR). Several other, less well-studied genes, also involved in DNA repair and all variably associated with increased risk for hereditary breast and ovarian cancer (HBOC), are also included on some panels. As PARP inhibitors have been approved for treating cancers with HR defects, it is important to identify all possible HR pathway genes reliably implicated in BC susceptibility. To this end, many clinical BC susceptibility genes (BCSGs) testing panels include putative BC-related HR pathway genes, but their candidacy remains unproven. This can lead to inappropriate use of expensive therapies.

We and others showed that tumor sequencing is a powerful tool to find genes or variants that cause HRD breast cancer. In BC cases due to known BCSGs, the inherited pathogenic variant confers genetic susceptibility, but tumorigenesis is usually driven by a somatic inactivating "second-hit" event resulting in loss of gene function, often via loss of heterozygosity (LOH) through deletion of the wild-type allele. In this tumor suppressor model of BCSG function, as a nascent tumor cell loses the ability to repair DNA, lesions accumulate, including in genes favoring tumor progression1. Tumor genomes of patients with BRCA1/2 inactivation were found to be enriched in copy number losses and rearrangements. This observation led to the development of an HRD index based on these somatic genomic abnormalities that can be used to detect HRD implicated tumors rather than on the specific mutational spectrum at the nucleotide level of the altered DNA sequence2-3. In addition, in tumors with anomalies in HR repair, single base-pair changes have been shown to accumulate in the tumor genome, giving rise to what is referred to as mutational signatures. We showed that using sequencing of paired tumor and normal tissues we can detect these HRD-associated mutational signatures as well as second hits4,5. COSMIC tumor mutational signature 3 (Sig 3) has been associated with homologous repair deficiency (HRD)6. Monoallelic inactivation of BRCA1/2 does not give rise to Sig 3, whereas Sig 3 is seen in tumors with biallelically inactivated BRCA1/2.

The mutational signature analysis helped to define the role of various HR genes directly from case-only data by association with Sig 3 levels. We showed that inactivation of RAD51C, RAD51D, PALB2, and BARD1 are all associated with elevated Sig 3 levels, indicating that GPVs in these genes lead to BRCA-like/HRD cancers. In contrast, GPVs in other established and candidate BCSG genes such as ATM, ATR, BRIP1, MRE11, NBN, and RAD50 do not lead to BRCA-like tumors7. Moreover, we developed a framework to reclassify germline BRCA1/2 variants of unknown significance (VUS) as benign or pathogenic variants in the context of HRD in tumors8. A variant was classified as a pathogenic or disease-causing variant if the tumor that harbored it showed a second hit and Sig 3. This approach had 100% sensitivity and 98.5% specificity in classifying 142 known benign and pathogenic missense mutations in BRCA1/2. Thus, Sig 3 levels can be used to identify new HR genes and new pathogenic variants in genes that lead to BRCA-like tumors.

In this study, we designed a pipeline encompassing variant calling and detection of second events involving somatic variants, combined with mutational signature analysis for HRD detection.

1Department of Human Genetics, McGill University, Montreal, QC, Canada. 2Cancer Research Program, Centre for Translational Biology, The Research Institute of the McGill University Health Centre, Montreal, QC, Canada. 3Cancer Genetics Unit, Institut Bergonié, Bordeaux, France. 4Division of Médecine Génétique, CRCHUM, Montreal, QC, Canada. 5Centre de recherche du Centre hospitalier de l’Université de Montréal and Institut du cancer de Montréal, Montreal, Quebec, Canada. 6Ecole Supérieure de Génie Informatique, Paris, France. 7Gerald Bronfman Department of Oncology, McGill University, Montreal, QC, Canada. 8Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada. 9IDIBELL Bellvitge Biomedical Research Institute, Barcelona, Spain. 10Department of Developmental Biology and Cancer Research, Institute for Medical Research-Israel-Canada, Hebrew University of Jerusalem, Jerusalem, Israel. 11Department of Medicine, McGill University, Montréal, QC, Canada. 12Department of Medicine, Université de Montréal, Montréal, QC, Canada. 13Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ✉email: william.foulkes@mcgill.ca; paz.polak@mssm.edu

Published in partnership with the Breast Cancer Research Foundation
The pipeline was optimized for archival formalin-fixed paraffin-embedded (FFPE) tumor tissues. We studied 38 unsolved BC patients with either early-onset BC or a strong family history of BC with no identified GPVs in known BCSGs (Supplementary Table 1). We tested the hypothesis that their tumors are linked to HRD due to alterations in genes that have not yet been definitively found associated with heritable risk to BC (Fig. 1).

To identify possible genetic causes that underly BC in our cases, we compiled a list of candidate genes implicated in HR which are not on current genetic testing panels (see methods). We analyzed whole-exome sequencing data (WES) of tumor/normal pairs to look for potentially pathogenic rare variants and to search for evidence of a second hit in the tumor based on a tumor suppressor gene model as well as for evidence of an HRD signature6 (Fig. 1).

Three identified variants were considered as pathogenic by ClinVar: two variants were in the FANCA and GJB2 genes, but neither showed a second hit. The third was a nonsense variant in LIG4 that showed LOH in the tumor. We observed LOH of the WT allele for 18 additional germline variants. We found five VUSs and three variants with Conflicting Interpretation of pathogenicity (CIOp) in APC, ATM, BRCA2, CHEK2, FLNC, and RAD51D; and the remaining variants were not classified in ClinVar (Fig. 2, Supplementary Fig. 1). No second somatic LOF alteration was found in combination with a germline candidate variant. All variants are summarized in Tables 1, 2.

To identify patients with HRD-related tumors, we used two HRD classifiers tools. SigMA is tailored for WES tumor data and is used to identify Sig 3 (Fig. 2, Supplementary Fig. 2). We also used ScarHRD, which is based on evaluating genomics scars due to HRD in tumor DNA and has been recently adapted for use with WES data. We identified 7 HRD tumors based on SigMA, and 12 based on ScarHRD (Supplementary Fig. 3). All 7 cases exhibiting an HRD signature by SigMA analysis were those also classified as having genomic scars consistent with HR defects by ScarHRD. Given that WES is not ideally suited to detect large genomic events, we considered the 7 tumors (18%) that were called positive by both tools for further analysis (Fig. 2). This percentage is within the range of detection reported by other HRD studies7,9-11. The only HRD tumor with a germline BCSG candidate variant was from a BC patient with the missense variant RAD51D p.S207L (case C10758). Using copy number analysis, we were able to identify duplication of the germline missense variant along with a trans allele loss, resulting in copy neutral LOH (CN-LOH) event at the gene locus as the second hit. We found a somatic deletion at a donor splice-site in RAD51C associated with duplication-LOH in another tumor with HRD (case JGH1), without germline events (Fig. 2). No other HRD tumor was found to be associated with our candidate genes either somatically or in the germline.

Variants in known BC drivers12 were found in 84% of tumors (Fig. 2, Supplementary Fig. 4). Of these, TP53, GATA3, and CBFB were considered as a driver (see methods). Three patients harbored a PIK3CA hotspot mutation. We also detected somatic alterations in three BC genes (PTEN, PALB2, and CDH1) in 4 other patients. Copy number profiles were similar to observations from previous BC cohort studies12,13.

Our pipeline worked well in calling variants in archival FFPE tumors, which are more accessible than fresh tumor material. Except for the RAD51D variant, which has a conflicting interpretation of pathogenicity (https://www.ncbi.nlm.nih.gov/clinvar/variation/142102/), we did not find supporting evidence to reclassify any ClinVar-described variants found in our HRD tumors as pathogenic. The fact that we were able to independently identify p.S207L RAD51D variant as a candidate pathogenic variant using our pipeline demonstrates the robustness of the genomic tools used in this study to reassess VUS. Although we sequenced a highly selected population of BC patients from multiple-case BC families, we did not find any other plausible GPVs in genes beyond...
those currently included in the clinical gene testing panels, which is in line with recent studies\textsuperscript{14}. We focused on HRD as a causal mechanism in BC, but our approach could be of use to identify germline drivers in other tumor types with mutational signatures that are associated with defects in DNA repair pathways, such as mismatch repair, hyper-mutated tumors, and base excision repair.

There are a number of caveats to our study. Our inability to detect potentially pathogenic variants in recently confirmed BCSGs (e.g. \textit{RAD51C} and \textit{BARD1})\textsuperscript{10,15} as well in largely unexplored HR genes could be due to the small sample size of our study. As WES is not ideally suited to identify large structural rearrangements associated with HR defects we could have missed tumor harboring anomalies due to HRD. Also, it is possible that undiscovered BCSGs do not operate via the HR pathway, and in this case our HRD-based framework will not detect variants in these genes. Finally, we cannot exclude the possibility that variants were missed for technical reasons given the age of some of the FFPE material, as one-quarter of our cases were diagnosed more than 10 years prior to analysis, which likely affected the quality of DNA for analysis. This limitation would not be as relevant for somatic testing of freshly processed tissues in a clinical context. Importantly, our study focused on WES data, so we could not assess other molecular mechanisms affecting gene function as potential contributors, such as complex structural variations\textsuperscript{16} or somatic methylation events.

Nones et al.\textsuperscript{17} took a similar approach using fresh frozen tumors and were able to show that all but one causally unexplained BCs with HRD were due to inactivation of \textit{BRCA1}, \textit{BRCA2}, or \textit{PALB2}. Taken together with the results we present here, we can conclude that the tail of missing heritability due to HRD genes must be long and thin. It also means that existing gene testing panels are sufficient to identify carriers of clinically important BC genes using a normal/tumor pair sequencing-based approach. We must therefore look elsewhere to find the missing inherited susceptibility present in these families.

**METHODS**

**Samples selection**

This study called "Genome-wide approaches in hereditary cancer families", study number A08-M61-09B, was created in 2011 and approved by the McGill University Research Ethics Board. Participants to this study were consented at two Montreal sites. Some participants were consented directly into the scientific study at McGill University and its affiliated hospitals. Others were recruited from a Montreal-based biobank entitled "Banque d’échantillons biologiques et de données (cliniques et biologiques) associées à des fins de recherche sur les cancers métastatiques du..."
| Gene | Patient ID | NM number | DNA changes | Amino Acid changes | Variant consequences | Gene category | ClinVar classification | ACMG classification | Variant retention type | Germline REF read counts | Germline ALT read counts | Tumor REF read counts | Tumor ALT read counts | p value |
|------|------------|-----------|-------------|-------------------|---------------------|---------------|----------------------|--------------------|----------------------|-----------------------|----------------------|----------------------|----------------------|---------|
| APC  | C00380     | NM_001127510 | c.8275C>T    | p.R2759C          | Missense            | CSG          | VUS                 | –                  | CN-LOH               | 80                    | 64                   | 134                  | 177                  | 0.013197 |
| ATM  | C11277     | NM_000051  | c.1538A>G    | p.G513R           | Missense            | BCSG         | VUS                 | –                  | CN-LOH               | 191                   | 142                  | 8                   | 24                   | 0.000446 |
| BRCA2 | C07766    | NM_000059  | c.974G>A     | p.M3248I          | Missense            | BCSG         | VUS                 | –                  | CN-LOH               | 114                   | 97                   | 13                  | 52                   | 0.000669 |
| BRDT | C09459     | NM_001242805 | c.272C>T     | p.A91V            | Missense            | C-HRG        | Unknown             | VUS                | CN-LOH               | 42                    | 35                   | 12                  | 38                   | 0.00129  |
| CABIN1 | C10860   | NM_012295  | c.2446C>T    | p.L816F           | Missense            | C-HRG        | Unknown             | Likely benign      | DEL-LOH              | 112                   | 107                  | 32                  | 55                   | 0.032096 |
| CHEK2 | C08017     | NM_001005735 | c.628G>A     | p.G210R           | Missense            | BCSG        | VUS                 | –                  | DEL-LOH              | 211                   | 155                  | 22                  | 103                  | <0.00001 |
| CHEK2 | C09211     | NM_001005735 | c.1165C>T    | p.R389C           | Missense            | BCSG        | VUS                 | –                  | DEL-LOH              | 114                   | 128                  | 36                  | 67                   | 0.037132 |
| CHEK2 | C11277     | NM_001005735 | c.1465A>G    | p.N489D           | Missense            | BCSG        | VUS                 | –                  | DEL-LOH              | 192                   | 180                  | 23                  | 49                   | 0.002237 |
| EIF2D | C09463     | NM_006993  | c.1538A>G    | p.Y513C           | Missense            | C-HRG        | Unknown             | VUS                | CN-LOH               | 35                    | 36                   | 136                 | 406                  | 0.000016 |
| FCHO1 | C07766     | NM_001161357 | c.557G>A     | p.R186Q           | Missense            | C-HRG        | Unknown             | VUS                | DUP-LOH              | 155                   | 148                  | 31                  | 87                   | <0.00001 |
| FLCN | JGH4       | NM_144997  | c.715C>T     | p.R239C           | Missense            | CSG          | CIOP                | –                  | CN-LOH               | 109                   | 120                  | 44                  | 185                  | <0.00001 |
| NCO1 | C102096    | NM_006311  | c.2794C>T    | p.H932Y           | Missense            | C-HRG        | Unknown             | VUS                | CN-LOH               | 138                   | 101                  | 34                  | 99                   | <0.00001 |
| POLQ | C00722     | NM_199420  | c.1981T>G    | p.N661G           | Missense            | C-HRG        | Unknown             | Likely Benign      | DUP-LOH              | 26                    | 26                   | 7                   | 21                   | 0.042667 |
| RAD51D | C10758    | NM_002878F | c.620G>T     | p.S207L           | Missense            | BCSG        | VUS                 | –                  | CN-LOH               | 129                   | 120                  | 13                  | 132                  | <0.00001 |
| REV1 | C08041     | NM_001321460 | c.2017G>A    | p.D673N           | Missense            | DNA         | Unknown             | VUS                | CN-LOH               | 88                    | 80                   | 60                  | 93                   | 0.01811  |

For the three CIOP variants, **FLCN** c.715C>T is trending towards benign [https://www.ncbi.nlm.nih.gov/clinvar/variation/134427/](https://www.ncbi.nlm.nih.gov/clinvar/variation/134427/) and is not a plausible BCSG. **RAD51D** c.620 G>T is discussed above and is recorded as likely pathogenic by most laboratories [https://www.ncbi.nlm.nih.gov/clinvar/variation/142102/](https://www.ncbi.nlm.nih.gov/clinvar/variation/142102/). **CHEK2** c.628 G>A (also known as **NM_007194.4**(CHEK2):c.499 G>A (p.Gly167Arg)) is reported as likely pathogenic by nearly all laboratories on ClinVar [https://www.ncbi.nlm.nih.gov/clinvar/variation/142524/](https://www.ncbi.nlm.nih.gov/clinvar/variation/142524/) and has been reported to result in a severe cancer phenotype in the homozygous state. While the presence of LOH in the case reported here could contribute to its final classification on ClinVar as Pathogenic, biallelic **CHEK2** mutations are not associated with any known mutational signature and therefore our approach cannot be used to reclassify this or any other **CHEK2** variant.

**BCSG** breast cancer susceptibility genes, **CSG** cancer susceptibility genes, **C-HRG** candidate homologous recombination genes, **DEL-LOH** LOH by deletion of the WT allele, **DUP-LOH** LOH by multiplication of the alternate variant and loss of the WT allele, **CN-LOH** LOH by copy neutral event, **VUS** variant of uncertain significance, **CIOP** variant with conflicting interpretations of pathogenicity.
Table 2. List of truncating variants retained in the tumor by LOH.

| Gene           | Variant type | Change       | Frequency of LoF | Allele frequency of LoF | Patient ID | NM_001370 | NM_007060 | NM_013770 | NM_013771-19 | NM_013771-18 | NM_013771-19-
|----------------|--------------|--------------|------------------|-------------------------|------------|------------|------------|------------|----------------|----------------|----------------|
| DNAH6          | frameshift-deletion | p.6184delG | 0.0007506        | 0.0007506               | C08041     | 79         | 34         | 48         | 36             | 0.0007506       | 0.0007506       |
| LIG4           | frameshift-deletion | p.6184delG | 0.0007506        | 0.0007506               | C08041     | 79         | 34         | 48         | 36             | 0.0007506       | 0.0007506       |
| TANGO2         | frameshift-deletion | p.6184delG | 0.0007506        | 0.0007506               | C08041     | 79         | 34         | 48         | 36             | 0.0007506       | 0.0007506       |

Somatic Copy Number Analysis and LOH detection

With GATK CNA, read coverage counts across the target were collected. We created a separate copy number variation (CNV) PoN to reduce read coverage data noise to obtain denoising copy ratios against the PoN. Then
we collected counts of reference versus alternate alleles for allelic imbalance. We used GISTIC\textsuperscript{26} to assess recurrent focal amplifications and deletions in tumors. We used FACTS\textsuperscript{27} and Sequenza\textsuperscript{28} to infer the integer copy number of each pair of alleles as well as LOH. Hemizygous deletions were defined as the loss of one allele in the tumor. Homozygous deletions were defined as the loss of both alleles in the tumor. Amplification was defined as a copy number status $\geq 6$ per tumor for ERBB2, MYC, and ZNF703.

**Variant visualization**

The Integrative Genomics Viewer\textsuperscript{29} was used for the manual examination and visualization of all germline candidate variants, and somatic events.

**Somatic cancer driver gene identification**

A combination of several tools was used to infer drivers of oncogenesis based on statistical methods through q-score to identify genes that were mutated more often than expected by chance: MutSigCV\textsuperscript{30}, OncodriveFMT\textsuperscript{31}, MutPanning\textsuperscript{32}, and FishHook\textsuperscript{33}. Genes that were found mutated more than once with a q-score < 0.05 were considered as cancer driver genes.

**Mutational signature and homologous recombination repair deficiency detection**

Different tools were used to assess the mutational signature from normal-tumor paired WES VCF data: MutationalPatterns\textsuperscript{34}, Maftools\textsuperscript{35}, DeconSigs\textsuperscript{36} for tumor signature clustering and the mutational signature contribution in each tumor. Different approaches were used: a de novo mutational signature extraction using a non-negative matrix factorization (NMF) and a COSMIC signature similarity using cosine similarity. We used SigMA\textsuperscript{37} to detect the mutational signature associated with HR defect designed for WES. We also used ScarHRD\textsuperscript{38} to detect HRD genomic scars-genomic abnormalities characteristic of HRD, which is based on the sum of LOH score, Large Scale Transition (LST) score, and Telomeric Allelic Imbalance (TAI) score. A score of $\geq 42$ classifies tumors as HRD (see Supplementary Fig. 3).

**Candidate variant selection algorithm**

A list of candidate pathogenic variants was generated using our variant detection algorithm (Supplementary Fig. 5). ClinVar annotations were used to classify variants already reported, and when no ClinVar data was available, we used the ACGM criteria. Variables classified as likely benign or benign by ClinVar were filtered out. The master list of variants was comprised of truncating, missense, or splice-site alterations found in gene candidates. For non BSCG, missense and splice-site variants predicted to be deleterious by at least 4 of 8 bioinformatic prediction tools (SIFT, PolyPhen2, MutationTaster, FATHM, Provean, MetaSVM, MetaLR, and CADD) were considered for further analysis. For known BSCG genes, we considered all variants regardless of scores as these can be evaluated by experts. Truncating variants were classified separately. The candidate list was comprised of 231 variants: 7 missense VUSs or variants with conflicting interpretations of pathogenicity in known BSCGs, 46 truncating variants, and 178 predicted pathogenic variants in C-HRGS, DNA repair genes, and in CSGs. From this list of candidate genes, we searched for evidence of allelic imbalance at each gene locus in the tumor using statistical tests (Chi-squared test) and filtered out variants where the p-value was $>0.05$ and confirmed with FACET results that there was a deletion of one haplotype (see above). We also looked for somatic second-hit point mutations. Our final list of candidate variants was comprised of genes that showed evidence of a possible second mutational hit in the tumor.

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**DATA AVAILABILITY**

The data generated and analyzed during this study are described in the following data record: https://doi.org/10.6084/m9.figshare.14802945. Tumor and germline datasets generated during the current study are not publicly available as the consent provided by study participants did not include a provision for widespread disclosure. However, direct requests can be made to the corresponding authors for access. The following files are openly available as part of the figshare data record: ‘Master_list_variant.xlsx’ (underlying Table 1), ‘Genes assessed initially.xlsx’ (underlying Supplementary Table 1), ‘Candidate Genes list.xlsx’ (underlying Supplementary Table 2). All other data are housed on institutional storage and are not openly available in order to protect patient privacy as informed consent to share participant-level data was not obtained prior to or during data collection. However, these data can be requested from Dr Foulkes. The files are: ‘all germline_sample.vcf’, ‘data.tumor.man’, ‘all_tumor_sample.vcf’, ‘all_sample.out.gz’small.seq.gz’.

**CODE AVAILABILITY**

The bioinformatics analyses were performed using open-source software, including Burrows–Wheeler alignment tool (v. 0.3.9), GATK (v.4.1.6.0), samtools (v.1.10), bcftools (v.1.10), Htslib (v.1.10), Picard (v.2.24.2), vcf2maf (v.1.6.17), R (v.4.0.3), CNV_facets (v.0.16.0), Maftools (v.0.9.30), FishHook (v.1.3.0), Bioconductor (v.3.12), MutationalPatterns (v.3.0.1), deconstructSigs (v.1.0), SigMA (v.1.0), Sequenza (v.3.0.0) and ScarHRD (v.1.0).

Received: 12 February 2021; Accepted: 26 July 2021; Published online: 25 August 2021

**REFERENCES**

1. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat. Rev. Cancer 12, 68–78 (2011).
2. Abkevich, V. et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br. J. Cancer 107, 1776–1782 (2012).
3. Popova, T. et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res. 72, 5454–5462 (2012).
4. Birkbak, N. J. et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. Cancer Discov. 2, 366–375 (2012).
5. Telli, M. L. et al. Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. Clin. Cancer Res. 22, 3764–3773 (2016).
6. Polak, P. et al. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. Nat. Genet. 49, 1476–1486 (2017).
7. Davies, H. et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational-signatures. Nat. Med. 23, 517–525 (2017).
8. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. Cell 149, 979–993 (2012).
9. Nguyen, L., W. M. Martens, J., Van Hoeck, A. & Cuppen, E. Pan-cancer landscape of homologous recombination deficiency. Nat. Commun. 11, 5544 (2020).
10. Riaz, N. et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. Nat. Commun. 8, 857 (2017).
11. Wang, Y. et al. Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. Nat. Genet. 49, 856–865 (2017).
12. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 534, 47–54 (2016).
13. Sircoulomb, F. et al. ZNF703 gene amplification at 8p12 specifies luminal B breast cancer. EMBO Mol. Med. 3, 153–166 (2011).
14. Narod, S. A. Which genes for hereditary breast cancer? N. Engl. J. Med. 384, 471–473 (2021).
15. Breast Cancer Association Consortium. et al. Breast cancer risk genes—association analysis in more than 113,000 women. N. Engl. J. Med. 368, 428–439 (2017).
16. Briandese, R. C. et al. BRCA1 deficiency is a recurrent event in early-onset triple-negative breast cancer: a comprehensive analysis of germline mutations and somatic promoter methylation. Breast Cancer Res. Treat. 167, 803–814 (2018).
17. Nones, K. et al. Whole-genome sequencing reveals clinically relevant insights into the aetiology of familial breast cancers. Ann. Oncol. 30, 1071–1079 (2019).
18. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010).
19. Chang, X. & Wang, K. wANNOVAR: annotating genetic variants for personal genomes via the web. J. Med. Genet. 49, 433–436 (2012).
20. Sillir-Rofe, D. et al. Mapping global and local coevolution across 600 species to identify key genes. Cell 173, 355–370e14 (2018).
21. Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443 (2020).
23. The 1000 Genomes Project Consortium. A global reference for human genetic
variation. Nature 526, 68–74 (2015).

24. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and
heterogeneous cancer samples. Nat. Biotechnol. 31, 213–219 (2013).

25. McLaren, W. et al. The ensembl variant effect predictor. Genome Biol. 17, 122 (2016).

26. Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the
targets of focal somatic copy-number alteration in human cancers. Genome Biol. 12, R41 (2011).

27. Shen, R. & Seshan, V. E. FACETS: allele-specific copy-number and clonal hetero-
geneity analysis tool for high-throughput DNA sequencing. Nucleic Acids Res. 44, e131 (2016).

28. Favero, F. et al. Sequenza: allele-specific copy number and mutation profiles from
tumor sequencing data. Ann. Oncol. 26, 64–70 (2015).

29. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer
(IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192 (2013).

30. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new
cancer-associated genes. Nature 499, 214–218 (2013).

31. Mularoni, L., Sabarinathan, R., Deu-Pons, J., Gonzalez-Perez, A. & López-Bigas, N.
OncoDriverFML: a general framework to identify coding and non-coding regions
with cancer driver mutations. Genome Biol. 17, 128 (2016).

32. Dietlein, F. et al. Identification of cancer driver genes based on nucleotide con-
text. Nat. Genet. 52, 208–218 (2020).

33. Imieliński, M., Guo, G. & Meyerson, M. Insertions and deletions target lineage-
deﬁning genes in human cancers. Cell 168, 460–472.e14 (2017).

34. Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. MutationalPatterns: compre-
hensive genome-wide analysis of mutational processes. Genome Med. 10, 33 (2018).

35. Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C. & Koehler, H. P. Maftools: efficient
and comprehensive analysis of somatic variants in cancer. Genome Res. 28, 1747–1756 (2018).

36. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C. DeconStrucT:
detecting mutational processes in single tumors distinguishes DNA repair defi-
ciencies and patterns of carcinoma evolution. Nucleic Acids Res. 47, 1–18 (2019).

37. Sztupinszki, Z. et al. Migrating the SNP array-based homologous recombination
deficiency measures to next generation sequencing data of breast cancer. NPJ Breast Cancer 4, 16 (2018).

38. Matis, T. S. et al. Metadata record for the article: Current gene panels account for
90% of cancer-associated genes. figshare https://doi.org/10.6084/m9.figshare.14802954 (2021).

39. Paperna, T. et al. Homozygosity for CHEK2 p.Gly167Arg leads to a unique cancer
syndrome with multiple complex chromosomal translocations in peripheral
blood karyotype. J. Med. Genet. 57, 500–504 (2020).

ACKNOWLEDGEMENTS
This study, including the acquisition of archived samples, was funded by a Quebec
Breast Cancer Foundation (QBCF) grant to W.D.F. and P.T. Tumor banking was
supported in part by the Banque de tissus et de données of the Réseau de recherche
sur le cancer of the Fonds de recherche du Québec Santé (FRQS) affiliated with the
Canadian Tumor Repository Network (CTRNNet). Y.T. is supported by Israel Science
Foundation grant. A.O. is supported by CCSRI Innovation Grant and Canada Research
Chair Tier 2 in Genome Stability and Hematological Malignancies. P.P. is supported by
the Jimmy V foundation scholar award.

AUTHOR CONTRIBUTIONS
T.S.M. developed the pipeline, analyzed the data, and drafted the manuscript. N.Z., B.L.,
Z.E.H., N.H., A.M.M.M., P.N.T., and W.D.F. provided clinical data or samples. N.Z. and M.D.
L. prepared the samples. T.A.M. and J.C.V. contributed to the pipeline development. A.
A. calculated the probability of pathogenic variants in BRCA1/2 and related genes. B.R.,
N.H., P.P., J.T., A.O., and W.D.F. validated the data. T.S.M., B.R., N.H., P.P., and W.D.F.
wrote the manuscript. All authors reviewed and edited the manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material
available at https://doi.org/10.1038/s41523-021-00315-8.

Correspondence and requests for materials should be addressed to W.D.F. or P.P.

Reprints and permission information is available at http://www.nature.com/
reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims
in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 Interna
tional License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give
appropriate credit to the original author(s) and the source, provide a link to the Creative
Commons license, and indicate if changes were made. The images or other third party
material in this article are included in the article’s Creative Commons license, unless
indicated otherwise in a credit line to the material. If material is not included in
the article’s Creative Commons license and your intended use is not permitted by statutory
regulation or exceeds the permitted use, you will need to obtain permission directly
from the copyright holder. To view a copy of this license, visit http://creativecommons.
org/licenses/by/4.0/.

© The Author(s) 2021

Published in partnership with the Breast Cancer Research Foundation

npj Breast Cancer (2021) 109