



OPEN ACCESS

EDITED BY

Paolo Peterlongo,
IFOM - The FIRC Institute of Molecular
Oncology, Italy

REVIEWED BY

Miguel De La Hoya,
San Carlos University Clinical Hospital,
Spain
Yiyang Wu,
Vanderbilt University Medical Center,
United States

*CORRESPONDENCE

Patricia N. Tonin
✉ patricia.tonin@mcgill.ca

SPECIALTY SECTION

This article was submitted to
Cancer Genetics,
a section of the journal
Frontiers in Oncology

RECEIVED 29 November 2022

ACCEPTED 06 February 2023

PUBLISHED 08 March 2023

CITATION

Alenezi WM, Fierheller CT, Serruya C,
Revil T, Oros KK, Subramanian DN, Bruce J,
Spiegelman D, Pugh T, Campbell IG,
Mes-Masson A-M, Provencher D,
Foulkes WD, Haffaf ZE, Rouleau G,
Bouchard L, Greenwood CMT, Ragoussis J
and Tonin PN (2023) Genetic analyses of
DNA repair pathway associated genes
implicate new candidate cancer
predisposing genes in ancestrally defined
ovarian cancer cases.
Front. Oncol. 13:1111191.
doi: 10.3389/fonc.2023.1111191

COPYRIGHT

© 2023 Alenezi, Fierheller, Serruya, Revil,
Oros, Subramanian, Bruce, Spiegelman,
Pugh, Campbell, Mes-Masson, Provencher,
Foulkes, Haffaf, Rouleau, Bouchard,
Greenwood, Ragoussis and Tonin. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Genetic analyses of DNA repair pathway associated genes implicate new candidate cancer predisposing genes in ancestrally defined ovarian cancer cases

Wejdan M. Alenezi^{1,2,3}, Caitlin T. Fierheller^{1,2}, Corinne Serruya²,
Timothée Revil^{1,4}, Kathleen K. Oros⁵, Deepak N. Subramanian⁶,
Jeffrey Bruce⁷, Dan Spiegelman^{1,8}, Trevor Pugh^{7,9,10},
Ian G. Campbell^{6,11}, Anne-Marie Mes-Masson^{12,13},
Diane Provencher^{12,14}, William D. Foulkes^{1,2,5,15,16,17},
Zaki El Haffaf^{12,18}, Guy Rouleau^{1,8}, Luigi Bouchard^{19,20,21},
Celia M. T. Greenwood^{1,5,17,22}, Jiannis Ragoussis^{1,4}
and Patricia N. Tonin^{1,2,16*}

¹Department of Human Genetics, McGill University, Montreal, QC, Canada, ²Cancer Research Program, Centre for Translational Biology, The Research Institute of McGill University Health Centre, Montreal, QC, Canada, ³Department of Medical Laboratory Technology, Taibah University, Medina, Saudi Arabia, ⁴McGill Genome Centre, McGill University, Montreal, QC, Canada, ⁵Lady Davis Institute for Medical Research of the Jewish General Hospital, Montreal, QC, Canada, ⁶Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia, ⁷Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada, ⁸Montreal Neurological Institute, McGill University, Montreal, QC, Canada, ⁹Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada, ¹⁰Ontario Institute for Cancer Research, Toronto, ON, Canada, ¹¹Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, VIC, Australia, ¹²Centre de recherche du Centre hospitalier de l'Université de Montréal and Institut du cancer de Montréal, Montreal, QC, Canada, ¹³Department of Medicine, Université de Montréal, Montreal, QC, Canada, ¹⁴Division of Gynecologic Oncology, Université de Montréal, Montreal, QC, Canada, ¹⁵Department of Medical Genetics, McGill University Health Centre, Montreal, QC, Canada, ¹⁶Department of Medicine, McGill University, Montreal, QC, Canada, ¹⁷Gerald Bronfman Department of Oncology, McGill University, Montreal, QC, Canada, ¹⁸Service de Médecine Génique, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada, ¹⁹Department of Biochemistry and Functional Genomics, Université de Sherbrooke, Sherbrooke, QC, Canada, ²⁰Department of Medical Biology, Centres intégrés universitaires de santé et de services sociaux du Saguenay-Lac-Saint-Jean hôpital Universitaire de Chicoutimi, Saguenay, QC, Canada, ²¹Centre de Recherche du Centre hospitalier l'Université de Sherbrooke, Sherbrooke, QC, Canada, ²²Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada

Not all familial ovarian cancer (OC) cases are explained by pathogenic germline variants in known risk genes. A candidate gene approach involving DNA repair pathway genes was applied to identify rare recurring pathogenic variants in familial OC cases not associated with known OC risk genes from a population exhibiting genetic drift. Whole exome sequencing (WES) data of 15 OC cases from 13 families tested negative for pathogenic variants in known OC risk genes were investigated for candidate variants in 468 DNA repair pathway genes. Filtering and prioritization criteria were applied to WES data to select top candidates for further analyses. Candidates were genotyped in ancestry defined study groups of 214 familial and 998 sporadic OC or breast cancer (BC) cases and 1025 population-matched controls and screened for additional

carriers in 605 population-matched OC cases. The candidate genes were also analyzed in WES data from 937 familial or sporadic OC cases of diverse ancestries. Top candidate variants in *ERCC5*, *EXO1*, *FANCC*, *NEIL1* and *NTHL1* were identified in 5/13 (39%) OC families. Collectively, candidate variants were identified in 7/435 (1.6%) sporadic OC cases and 1/566 (0.2%) sporadic BC cases versus 1/1025 (0.1%) controls. Additional carriers were identified in 6/605 (0.9%) OC cases. Tumour DNA from *ERCC5*, *NEIL1* and *NTHL1* variant carriers exhibited loss of the wild-type allele. Carriers of various candidate variants in these genes were identified in 31/937 (3.3%) OC cases of diverse ancestries versus 0–0.004% in cancer-free controls. The strategy of applying a candidate gene approach in a population exhibiting genetic drift identified new candidate OC predisposition variants in DNA repair pathway genes.

KEYWORDS

germline variants, familial ovarian cancer, cancer predisposing genes, whole exome sequencing, DNA repair pathways, Genetic drift

Introduction

Since the identification of *BRCA1* (1) and *BRCA2* (2) as breast cancer (BC) and ovarian cancer (OC) predisposing genes, which are involved in the homologous recombination (HR) DNA repair pathway (3), no other major high risk gene has been reported to account for the remaining familial cancer cases found to be negative for germline pathogenic variants (PVs) in these genes (4). Carriers of PVs in *MLH1*, *MSH2*, *MSH6* or *PMS2*, genes involved in the mismatch repair (MMR) pathway (3), have also been shown to have a significantly increased lifetime risk of developing OC (5, 6) often associated with hereditary non-polyposis colorectal cancer syndrome families (7). However, carriers of PVs in MMR genes are very rare accounting for fewer than 1% of sporadic OC cases, which is significantly lower than the 5–15% carrier frequency of PVs in *BRCA1* and *BRCA2*, depending on the population studied (8). Carriers of PVs in relatively new OC predisposing genes have been reported such as *RAD51C* (9), *RAD51D* (10) and *BRIPI* (11), genes also involved in the HR DNA repair pathway (3). The carrier frequency of PVs in each of these genes combined is estimated to be less than 2% of sporadic OC cases (8, 12–14). PVs in other DNA repair genes such as *PALB2* (15–19), *CHEK2* (16, 20) and *ATM* (16, 17), all associated with BC risk, were recently associated with OC, though risk has yet to be established. Other genes also playing a role in various DNA repair pathways have been proposed as candidate OC risk genes such as *FANCM* (15, 21), *POLE* (22), *MRE11* (17, 23), *RAD1* (17) and *FANCI* (24), and collectively, the frequency of carriers of PVs in these genes are also low relative to *BRCA1* and *BRCA2* carriers. Thus, research has consistently shown that a candidate gene approach investigating DNA pathway genes has successfully identified new and candidate OC predisposing genes (25, 26), though it is expected that the carrier frequency is significantly lower relative to carriers harbouring PVs in *BRCA1* and *BRCA2*.

Defining the contribution of moderate- to high-risk genes in OC remains a challenge as it is not clear that all monogenic cancer predisposing genes have been identified for this genetically

heterogeneous disease (4). Based on the family history of cancer and the population investigated, the proportion of OC families known to be negative for *BRCA1* or *BRCA2* PVs has a wide range of approximately 15–65% (8, 27). Indeed, a recent whole exome sequencing (WES) study of familial and sporadic OC cases revealed significant heterogeneity of candidate OC risk genes, representing diverse functional pathways with relatively few involved in the approximately 200 investigated DNA repair genes (17). However, this study focused only on investigating rare, protein-coding loss-of-function (LoF) variants (17). As there are at least 400 known or putative genes that are directly or indirectly involved in repairing DNA (3, 28–32), it is plausible that PVs (LoF or missense) in genes not previously investigated in OC could be associated with OC risk that have yet to be identified. As carriers of new candidate variants are likely to be rare, identifying them will be challenging.

We have proposed a strategy for identifying candidate variants in new cancer predisposing genes that involves the investigation of cancer families from populations exhibiting genetic drift (33). Over time, rare PVs in such populations could attain disproportionately high carrier frequencies of rare risk variants relative to the general population (34, 35). For example, *PALB2* (36) and *BRIPI* (11) were discovered as BC and OC predisposing genes by investigating cancer families and cases from the Finnish and Icelandic populations, respectively, both populations exhibiting genetic drift. Our research of French Canadians (FC) from the Quebec population of Canada, identified *RECQL* (37) and *FANCI* (24) as new candidate BC or OC predisposing genes, respectively. Genetic drift in the FCs of Quebec has been attributed to common ancestors as a result of the geographic isolation and multiple waves of expansion of European settlers from France since 1608 (33–35, 38, 39). Investigating these populations facilitates the characterization of deleterious variants in known or candidate cancer predisposing genes as all types of variants could be investigated and not only LoF variants (33). A small number of PVs in *BRCA1* and *BRCA2* (40, 41) and one in each of *PALB2* (42), *RAD51C*

(43) and *RAD51D* (43, 44) have been shown to be frequently occurring in FC OC and/or BC cases versus population-matched controls. Specific PVs in *MLH1* (45), *MSH2* (45) and *MSH6* (46) have also been reported in FCs in the context of hereditary non-polyposis colorectal cancer.

We recently reported that not all remaining *BRCA1* and *BRCA2* negative families with at least two close relatives with OC from the FC population of Quebec by WES analysis were due to PVs in *RAD51C*, *RAD51D* or *BRIP1* approximating that 40% of such cases are unaccounted for by known or emerging OC predisposing genes (33, 43, 47). Also, we reported that likely pathogenic variants (LPV) in *FANCI*, a proposed new OC predisposing gene from the Fanconi anemia (FA) pathway, were rarely implicated in familial and sporadic OC cases in this population (24). We posit that DNA repair pathway genes have not been fully explored as candidate OC risk genes. In this study, we report the identification of candidate LPVs in DNA repair pathway genes that were identified by applying a candidate gene approach focusing on an extensive list of 468 DNA repair pathway genes in available WES data derived from the germline of FC familial OC cases. Candidate variants prioritized based on our bioinformatic analyses were selected for targeted genotyping in larger groups of defined FC cases to determine carrier frequencies in 1212 FC and 937 non-FC familial and sporadic OC cases and 1025 population-matched controls. Available tumour DNA from our FC carriers also was investigated for loss of the wild-type allele of candidate loci.

Materials and methods

Study participants

FC cancer cases and controls are described in Figure 1 and Table S1. For clarity, cases diagnosed with primary fallopian tubes or peritoneal cancers were included as cases in recognition of common aetiology with OC, which are associated with hereditary cancer risk factors (48, 49). For the discovery of new candidate OC risk variants (study phase I; Figure 1A), WES data from peripheral blood lymphocytes (PBL) DNA was available from 15 OC cases from 13 families, each family having at least two first-, second- or third-degree relatives with OC. These cases were confirmed negative for PVs in the known OC risk genes: *BRCA1*, *BRCA2*, *BRIP1*, *RAD51C* or *RAD51D* as previously reported (47, 43). This group includes three index cases harbouring a LPV in *FANCI* c.1813C>T; p.Leu605Phe (24). As *FANCI* remains a candidate OC predisposing gene requiring further independent studies, we did not exclude *FANCI* variant carriers from any of our study groups for our investigation.

Targeted analyses of the candidate variants was performed to determine their carrier frequencies (study phase II; Figure 1B) on the PBL DNA from FC OC cases, regardless of their carrier status for *BRCA1* and *BRCA2* PVs, from 42 hereditary breast and ovarian cancer (HBOC) syndrome families having one OC and at least two BC cases in the same familial branch and 435 sporadic OC cases not selected for age at diagnosis with the disease or for family history of any cancers. Genetic data was available from 1025 population-

matched controls provided by three independent biobanks as previously described (43). As known OC predisposing genes are also involved in BC risk (nccn.org/guidelines/category_2), targeted analyses of the candidate variants was also performed on the PBL DNA from FC BC cases, regardless of their *BRCA1* and *BRCA2* PV carrier status, from 33 HBOC families, 139 hereditary breast cancer (HBC) syndrome families having at least three close relatives with BC within first-, second- or third-degree of relationship from the same familial branch and 563 sporadic BC cases not selected for age at diagnosis with the disease or for family history of any cancers.

Targeted analyses of PBL DNA from additional OC cases was performed to identify more OC carriers of our candidate variants (study phase III; Figure 1C). These groups were comprised of: 52 sporadic early-onset cases diagnosed with high-grade serous ovarian carcinoma (HGSC) before the age of 50 years who tested negative for PVs in *BRCA1* or *BRCA2*; and 553 OC cases, regardless of their *BRCA1* and *BRCA2* PV carrier status and not defined by any criteria as previously described in this study (43).

The majority of FC cancer cases self-reported FC ancestry of Quebec as described previously (24, 40, 41, 43, 47, 50–56). FC controls from Université de Sherbrooke-The Genetics of Glucose Regulation in Gestation and Growth (Gen3G) (57) and McGill University-Montreal Neurological Institute (MNI) (58) biobanks self-reported FC ancestry as described previously (43). FC controls from CARTaGENE biobank (cartagene.qc.ca) were born in the province of Quebec, reported being FC ancestry, having parents and all four grandparents born in Canada and French as first language learned as described previously (43, 59).

The cancer cases and controls not selected for being of FC ancestry of Quebec, mainly of European ancestry are referred to as non-FC groups in this study, were available from different resources. Genetic analyses to determine the spectrum and prevalence of candidate variants in genes that were identified in the study phase I were performed (study phase IV; Figure 1D) on available genetic data derived from PBL DNA from three independent groups with OC: 9 OC cases from 7 families with at least two close relatives with OC (MIX familial OC cases) (47); 516 OC familial or sporadic cases from the Australian population (AUS OC cases) (17) and 412 OC cases as part of The Cancer Genome Atlas (TCGA) Pan-Cancer Atlas project (not selected for ethnicity) (60) and cancer-free controls as part of the Genome Aggregation Database (gnomAD) v2.1.1. (61). The gnomAD v2.1.1. controls were also used to filter common variants as part of study phase I (Figure 1A and Table S1) which is described in the following section.

All biological samples, clinico-pathological, pedigree and relevant medical genetic information from the cancer cases and control groups that were investigated in this study are from biobanks where participants had been recruited in accordance with ethical guidelines of the biobanks respective Institutions Research Ethics Boards as described in Table S1. Where applicable, samples were anonymized at source by the providers and were assigned a unique identifier (PT followed by four digits) to further protect their identity. This project was conducted with approval and in accordance with the guidelines of The McGill University Health Centre Research Ethics Board (MP-37-2019-4783).

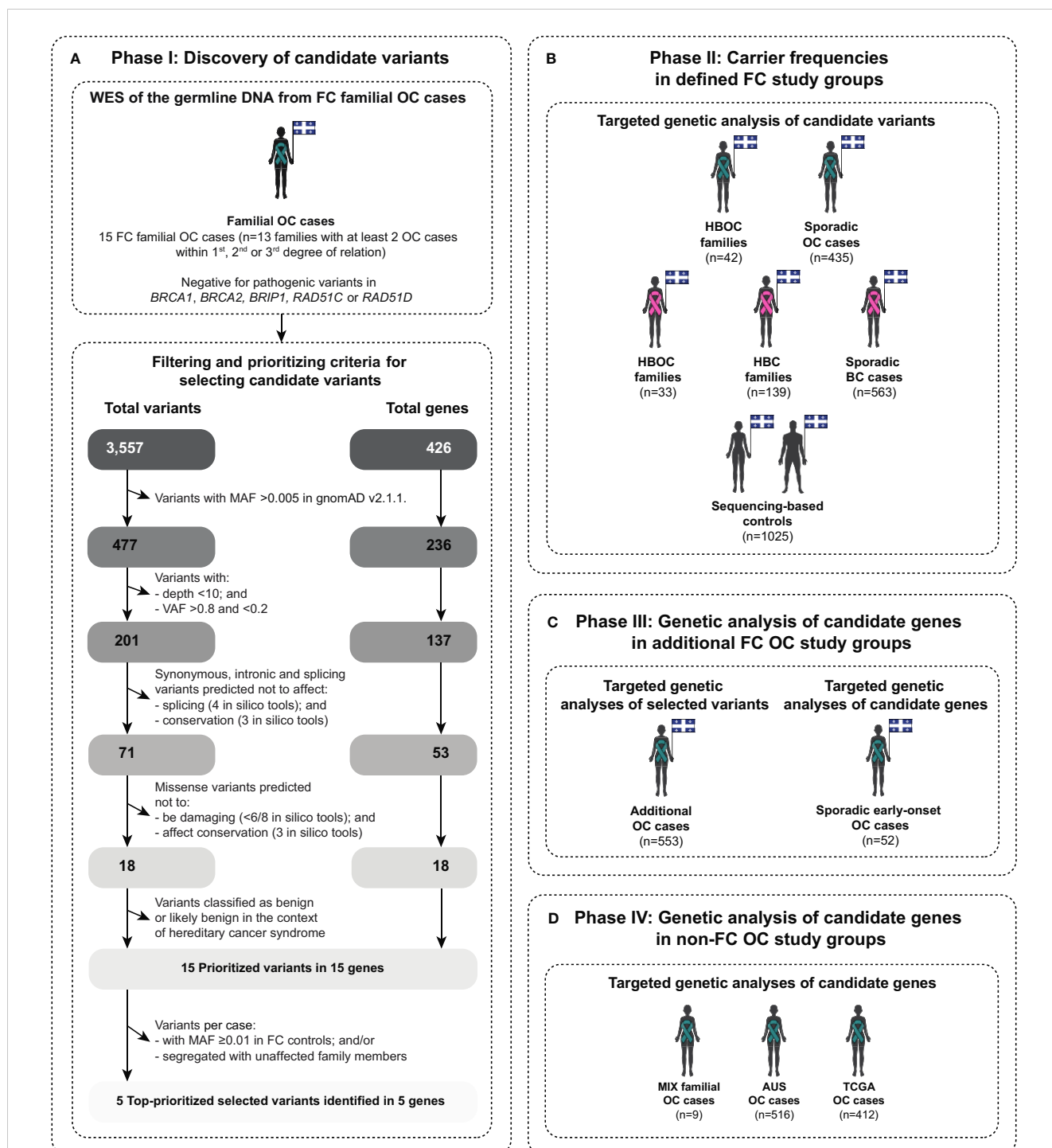


FIGURE 1

Scheme describing different phases of this study in identifying and evaluating candidate variants in genes involved in various DNA repair pathways. The diagram illustrates: (A) study phase I for identifying candidate variants by applying a candidate gene approach of known or putative DNA repair genes (see Table S2) on peripheral blood lymphocytes (PBL) DNA from familial ovarian cancer (OC) cases of French Canadians (FC) of Quebec by whole exome sequencing (WES) and bioinformatic analyses (see Table S1); (B) study phase II for determining the carrier frequencies of the top-prioritized candidate variants in FC familial and sporadic OC and BC cases, including hereditary breast and ovarian cancer (HBOC) syndrome and hereditary breast cancer (HBC) syndrome families, and population-matched controls by targeted genetic analyses (see Table S1); (C) study phase III for identifying additional carriers in FC OC cases by targeted genetic analyses (see Table S1); and (D) study phase IV for identifying candidate variants in the identified candidate DNA repair genes from phase I in non-FC OC cases, mainly of European origin, by targeted genetic analyses: (MIX, mixed ethnicity; AUS, Australian; and TCGA, The Cancer Genome Atlas) (see Table S1). Teal ribbon signifies women with OC and pink ribbon signifies women with BC, and diagrams contain the provincial flag of Quebec, Canada denoting the geographic ascertainment of cases and controls. MAF, Minor allele frequency; and VAF, Variant allele frequency.

Identifying and selecting for top candidate variants in FC cancer cases

For phase I of the study (Figure 1A), WES data was available from PBL DNA from 15 OC index cases from 13 cancer families that had at least one first-, second- or third-degree relative from the same familial branch with OC, and were confirmed being negative for PVs in *BRCA1*, *BRCA2*, *BRIP1*, *RAD51C* or *RAD51D* by WES analyses (43, 47). WES had been subjected to a customized bioinformatics pipeline for germline variant calling at the McGill Genome Center as previously reported by our group (24, 43). In brief, NimbleGen SeqCap[®] EZ Exome v3.0 library kit (Roche, US), followed by paired-end sequencing on different Illumina HiSeq platforms was performed. Reads were aligned to the human reference genome assembly GRCh37/hg19 using Burrows-Wheeler aligner v0.7.17, followed by PCR deduplication using Picard v2.9.0. Realignment around small insertions and deletions was performed, and germline variants were called using HaplotypeCaller using Genome Analysis Toolkit (GATK) v3.5. Variants were then filtered for base sequencing quality score ≥ 30 and annotated using Ensembl Variant Effect Predictor (VEP) and GEMINI v0.19.1.

Using a candidate gene approach, a curated list of 468 known or putative DNA repair genes (28–32) (ebi.ac.uk/QuickGO/term/GO:0006281) (Table S2) were investigated for candidate PVs in WES data from selected index OC cases (Figure 1A). Variants identified in these DNA repair genes were extracted from the annotated variant call format (VCF) files from the index OC cases (Figure 1A). Variants with minor allele frequency (MAF) > 0.005 in gnomAD v2.1.1.1. (61, 62), with total low coverage < 10 reads and/or those with variant allele frequency (VAF) < 0.2 and > 0.8 were filtered out and retained variants were subjected for further prioritization and selection. These thresholds have been tested previously under the assumption that new variants follow an autosomal dominant mode of inheritance (63). These variants were then verified by manual inspection in the aligned sequences in compressed binary alignment map (BAM) files by Integrative Genomics Viewer (IGV) v2.4.10. (64).

Top candidate variants were selected from this master list of variants for further analyses based on various prioritization criteria as shown in Figure 1A. First, we prioritized LoF variants (nonsense, frameshift and alternative splicing variants), inframe, missense and intronic variants, which were predicted to be conserved and damaging at the RNA or protein level by 15 selected in silico tools: (1) by at least one out of three prediction tools for conservation Genomic Evolutionary Rate Profiling v1.0 (GERP++ [score ≥ 2.0]) (65), Phylogenetic P value of 100 vertebrates v4.2 (PhyloP 100 way [score ≥ 0.2]) (66) and PAST Conservation of 100 vertebrates v4.2 (PhastCons 100 way [score ≥ 0.9]) (67); (2) by at least one out of four prediction tools for splicing Maximum Entropy Estimates of Splice Junction v2.0 (MaxEntScan) (68), two different Database Splicing Consensus Single Nucleotide Variant (dbSNV) in silico tools: AdaBoost v4.0 (ADA [score ≥ 0.4]) and Random Forest v4.0 (RF [score ≥ 0.4]) (69) and SpliceAI (score ≥ 0.4) (70); and (3) at least six out of eight prediction tools for damaging of protein function based on their best performance (71–74): Combined Annotation Dependent Depletion v1.4 (CADD [Phred

score ≥ 20]) (75), Eigen (score ≥ 0.0) (76), Meta-analytic Logistic Regression v4.2 (MetaLR [score ≥ 0.5]) (77), Meta-analytic support Vector Machine v4.2 (MetaSVM [score ≥ 0.0]) (77), MetaRNN 4.2 (score ≥ 0.5) (78), Rare Exome Variant Ensemble Learner v4.2 (REVEL [score ≥ 0.5]) (79), Variant Effect Scoring Test v4.2 (VEST [score ≥ 0.5]) (80) and Protein Variation Effect Analyzer v4.0 (PROVEAN v4.0 [score ≤ -2.5]) (81). Then, the variants having a clinical classification as benign or likely benign in the context of hereditary cancer syndromes in ClinVar (82, 83) and/or American College of Medical Genetics and Genomics (ACMG) guidelines (84, 85) were given a lower priority for further investigation.

The remaining prioritized variants were then subjected to further prioritization. Variants were surveyed in available genetic data generated from the germline of three FC study groups (Table S1): (1) WES data from the germline of 52 sporadic early-onset OC cases negative for PVs in *BRCA1* and *BRCA2* (43, 47); and (2) sequencing-based (WES or whole genome sequencing [WGS]) data and/or genotyping-based data from 1025 FC controls (43). Then, the variants were subjected for further selection and characterization for genetic analyses.

Selected top candidate variants were verified in the PBL DNA by bidirectional Sanger sequencing using customized primers (available upon request) performed at the McGill Genome Center as described previously (86, 43, 47; 24, 33). Sequencing chromatograms were visually inspected for variant heterozygosity using 4Peaks v1.8. (nucleobytes.com/4peaks/) (The Netherlands Cancer institute, Amsterdam, The Netherlands).

Determining carrier frequencies of selected candidate variants in FC cancer cases and controls

Selected top candidate variants were investigated for carrier frequencies in defined FC study groups (study phase II) comprised of 42 index OC and 33 index BC cases from 75 HBOC families, 139 index BC cases from 139 HBC families, 435 sporadic OC cases and 563 sporadic BC cases (Figure 1B and Table S1). PBL DNA from index cases were genotyped using customized TaqMan[®] (87), Sequenom iPLEX[®] Gold (88) or Fluidigm[®] SNP TypeTM (89) genotyping assays (primers available upon request) as described previously (24, 43, 90). Tumour DNA samples from the index case were genotyped where PBL DNA was no longer available from the biobank. Carriers of candidate variants were verified by bidirectional Sanger sequencing of PBL DNA as described above. Selected candidate variants were also investigated for carrier frequency in population-matched controls by surveying 1025 available sequencing-based data sets (Table S1) and/or 8493 single nucleotide polymorphism (SNP) genotyping-based data sets as previously described (43). For probes of variants not presented on the SNP array, pre-phasing and imputation were performed as described previously (24, 43).

Pair-wise comparisons were performed of carrier frequencies of candidate variants in different FC cancer groups versus sequencing-based controls. Two-tailed Fisher's exact test was used to compare

carrier frequencies in the cancer versus control groups where unadjusted *P* values <0.05 for multiple testing was considered significant.

Targeted genetic analyses of selected candidate variants or genes in FC cancer cases

To further characterize our candidate variants and genes in a population exhibiting genetic drift, we investigated carrier status in additional OC cases from the FC population. Selected top candidate variants were investigated (study phase III) in 52 sporadic early-onset FC HGSC cases and in an additional 553 FC OC cases by surveying available genetic data or targeted genotyping of PBL DNA (Figure 1C and Table S1). We also investigated other variants in our gene candidates that met our filtering and prioritizing criteria in the available WES data from the sporadic early-onset OC cases (Figure 1C).

Loss of heterozygosity analyses of candidate genes loci in OC tumour DNA from FC candidate variant carriers

To investigate evidence for inactivation of candidate genes in cancer cells, we performed loss of heterozygosity (LOH) analysis of tumour DNA from variant carriers. Bi-directional Sanger sequencing of available tumour DNA was performed using customized primers (available upon request) as described above. Extracted DNA from fresh-frozen (FF) or histopathological sections from formalin-fixed paraffin-embedded (FFPE) tumour tissues were provided by the RRCancer biobank for DNA extraction and LOH analysis (Promega, Canada). Sequencing chromatograms were inspected for loss of the wild-type allele using 4Peaks v1.8. (nucleobytes.com/4peaks/) (The Netherlands Cancer Institute, Amsterdam, The Netherlands).

Genetic analyses of candidate genes in non-FC cancer cases and controls

To further characterize the candidate variants and genes identified in our FC cancer cases, we investigated available genetic data from other populations that were not specifically selected for FC ancestry. The spectrum and prevalence of our candidate variants were investigated in genetic data from non-FC OC cases being predominantly of European ancestry and cancer-free controls were investigated for new variants in our candidate genes that met our filtering and prioritizing criteria (Figure 1D and Table S1). Variants were extracted from the annotated VCF files generated by WES data from the germline of: (1) 9 MIX familial OC cases; (2) 516 AUS OC cases; and (3) 412 OC cases from the Pan-Cancer TCGA project. Variants were extracted from the comma separated value (CSV) files downloaded directly from (<https://gnomad.broadinstitute.org>). All variants were annotated and subjected to our filtering and prioritizing criteria to identify candidate variants as described above (see Figure 1A).

Genetic analyses for co-occurring rare pathogenic variants in known OC risk genes in FC and non-FC candidate variant carriers

We investigated available WES data from OC cases, to determine whether the identified carriers of candidate variants (regardless of the ethnicity and study phase in which they were identified) also harbour rare PVs or LPVs in known OC risk genes (*n*=11): *BRCA1* (NM_007294.4), *BRCA2* (NM_000059.4), *MLH1* (NM_000249.4), *MSH2* (NM_000251.3), *MSH6* (NM_000179.3), *PMS2* (NM_000535.7), *BRIP1* (NM_032043.3), *RAD51C* (NM_058216.3), *RAD51D* (NM_001142571.2), *PALB2* (NM_024675.4) and *ATM* (NM_000051.4) based on the National Comprehensive Cancer Network (NCCN) Clinical Practice in Oncology Guidelines 2022 (Version 2.2022) —Genetic/Familial High-Risk Assessment: Breast, Ovarian and Pancreatic (nccn.org/guidelines/category_2). Variants were extracted from the annotated VCF files from carriers and subjected to our filtering and prioritizing criteria as previously described (47).

Results

Prioritization and selection of candidate variants: Phase I

We first extracted variants identified in a curated list of 468 DNA repair genes (Table S2) from VCF files generated from WES data from the 15 index FC OC cases from 13 families. In these index cases, we identified a total of 3,557 variants in 426 of 468 DNA repair genes (Figure 1). Based on their rarity and variant quality, we retained a total of 201 variants in 137 of 426 DNA repair genes where each index case harboured 3 to 25 (median=15) variants. From this list of 201 variants, we prioritized candidates that were predicted to be conserved or damaging at the level of RNA or protein using our selected *in silico* tools; and those classified benign or likely benign in the context of hereditary cancers using ClinVar and/or ACMG guidelines were not pursued further. Using these criteria, we retained a total of 15 of the 201 variants, each of which was found in a different gene: 3 nonsense variants, 1 canonical alternative splicing variant, 1 inframe and 10 missense variants (Table S3). These variants were identified in 10 of the 15 index cases from 8 out of the 13 OC families (Figures 2, S1). One of these variants was identified in two OC cases from the same family, two variants were identified in two OC cases from the same family and the remaining 12 variants were identified in one index case from independent OC families (Table S3). Two cases harboured either three or four variants, while the remaining nine cases harboured one to two variants.

To select our top candidates for further analyses, we reviewed the individual context wherein the 15 variants were identified as shown in Table S3 and Figures 2, S1. We estimated the allele frequencies of the 15 candidate variants in population-matched FC controls. Thereby, we did not pursue *ALKBH3* c.677A>G; p.Asn226Ser, which was identified in an index case (PT0136)

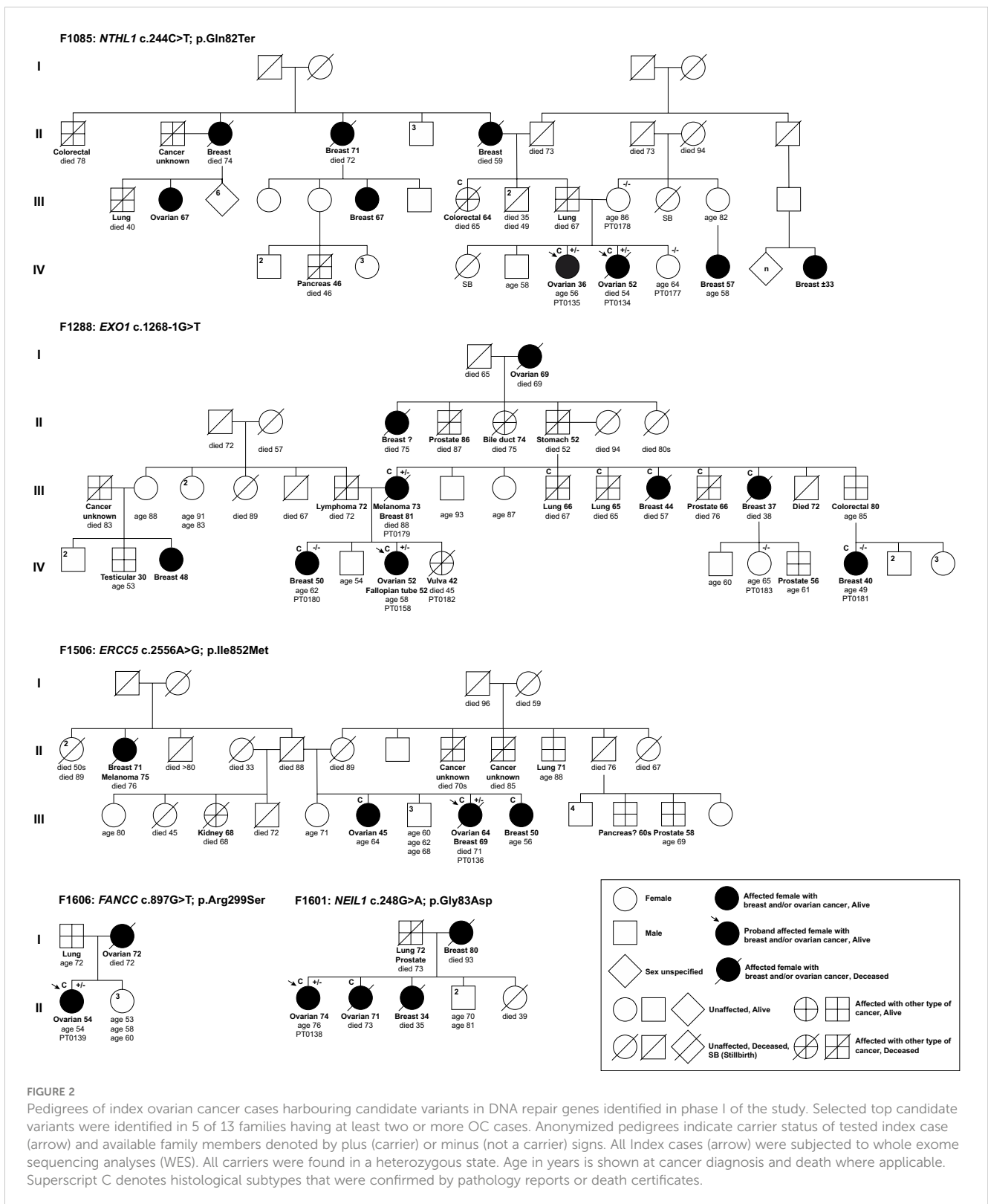


FIGURE 2

Pedigrees of index ovarian cancer cases harbouring candidate variants in DNA repair genes identified in phase I of the study. Selected top candidate variants were identified in 5 of 13 families having at least two or more OC cases. Anonymized pedigrees indicate carrier status of tested index case (arrow) and available family members denoted by plus (carrier) or minus (not a carrier) signs. All Index cases (arrow) were subjected to whole exome sequencing analyses (WES). All carriers were found in a heterozygous state. Age in years is shown at cancer diagnosis and death where applicable. Superscript C denotes histological subtypes that were confirmed by pathology reports or death certificates.

from family F1506, as it has a MAF ≥ 0.01 in FC controls. We also excluded the missense variant in *DNA2* c.836C>T; p.Thr279Ile for further analyses as it was identified in the index case (PT0128) and four unaffected members of the family F694, and was not inherited from the affected mother with the family history of OC and other cancers (Figure S1). We did not pursue *RBBP8* c.1941T>G;

p.Asp647Glu as it was not harboured by the other index OC case (PT0056) from the same family F1528 (Figure S1). Additional variants that were excluded for further analyses included: (1) *RHNO1* c.250C>T; p.Arg84Ter in the index OC case (PT0158) from family F1288; (2) *ATRX* c.4377_4379del; p.Glu1464del in one of the index OC case (PT0057) from family F1528 as they were

classified as benign in ClinVar and by ACMG guidelines, and the latter as not being harboured by the other index OC case of the same family F1528 (Figure S1); (3) *SMARCA2* c.3265C>T in the index OC case (PT0128) from family F694; p.Arg1089Trp; and (4) *KMT2C* c.6916C>T; p.Pro2306Ser in the index OC case (PT0047) from family F1490 as variants in these genes are associated with non-cancer related syndromes (Figure S1). Heterozygous germline variants in *SMARCA2* are linked with Nicolaides-Baraitser syndrome (MIM: 601358), which is characterized by intellectual disability, seizures, limited to absence of speech ability, short stature, dysmorphic facial features and sparse hair (91, 92; 93–97); and heterozygous germline variants in *KMT2C* are linked with Kleefstra syndrome, type 2 (MIM: 617768), which is characterized by delayed psychomotor development, variable intellectual disability and mild dysmorphic features (98–101). A genotype-phenotype of heterozygous variants located within exon 15–25 of *SMARCA2*, which encodes the ATPase domain, have been recently reported that over 80% of these variants were *de novo* based on WES analyses of 80 cases in trios with Nicolaides-Baraitser syndrome that have been documented worldwide so far (92, 95, 97). Our *SMARCA2* c.3265C>T; p.Arg1089Trp has never been reported in the literature, but it is located in exon 23 that encodes the ATPase domain (97). Whereas, *KMT2C* c.6916C>T; p.Pro2306Ser was reported in the context of Kleefstra syndrome (100). Finally, we did not pursue: *RECQL5* c.918G>A; p.Met306Ile, *ASCC3* c.3808C>T; p.Arg1270Ter and *UBB* c.569C>A; p.Pro190His as all were harboured by the same index case (PT0139) from family F1606 that also harbouring *FANCC* c.897G>T; p.Arg299Ser as this variant is a plausible and intriguing candidate where *FANCC* has been reported as a candidate BC predisposing gene (102). Thus, *FANCC* c.897G>T; p.Arg299Ser and the remaining variants identified in *ERCC5* c.2556A>G; p.Ile852Met, *EXO1* c.1268-1G>T, *NEIL1* c.248G>T; p.Gly83Asp and *NTHL1* c.244C>T; p.Gln82Ter from our list of most promising variants were selected for further analyses as candidates.

Characterization of the selected top candidate variants: Phase I

We selected five variants each identified in an OC family for further characterization and analyses (Table 1 and Figure 2): a nonsense variant *NTHL1* c.244C>T; p.Gln82Ter, a canonical splicing variant *EXO1* c.1268-1G>T, an exonic splicing variant *FANCC* c.897G>T; p.Arg299Ser and two missense variants *ERCC5* c.2556A>G; p.Ile852Met and *NEIL1* c.248G>T; p.Gly83Asp. Both missense variants were predicted to affect amino acid residues that are located in catalytic domains of their respective proteins that are critical to the biological function of ERCC5 (108) and NEIL1 (109) in the HR, nucleotide excision repair (NER) and base excision repair (BER) pathways.

Except for *EXO1* c.1268-1G>T, which was not found in the gnomAD v2.1.1. database, all other candidate variants were found to have MAFs between 0.002 and 0.00001 in the non-cancer non-Finnish European populations with variation in these frequencies across populations of different ancestry groups (Table S4). The loci

of all five candidate variants were predicted to be conserved by at least one of the selected *in silico* tools. The variants in *EXO1* and *FANCC* were predicted to affect splicing by all four selected *in silico* tools. The missense variants in *ERCC5* and *NEIL1* were predicted to be damaging by at least six selected *in silico* tools, including REVEL and VEST, which are two of the recently validated as top performing prediction *in silico* tools (73) (Table S3). Only *NTHL1* c.244C>T; p.Gln82Ter is classified as PV in ClinVar and by ACMG guidelines in the context of hereditary multi-cancer syndrome in an autosomal recessive mode of inheritance, and has recently been associated with BC risk in an autosomal dominant mode of inheritance (110, 111). Whereas, *FANCC* c.897G>T; p.Arg299Ser was classified as being of uncertain significance (VUS) in ClinVar in the context of FA, an autosomal recessive disorder (MIM: 227645) and as LPV by ACMG guidelines. As noted above, *FANCC* has been associated with BC predisposition in an autosomal dominant mode of inheritance (102). The remaining candidate variants have not been reported in ClinVar, but classified by ACMG guidelines as LPV for *EXO1* c.1268-1G>T, VUS for *ERCC5* c.2556A>G; p.Ile852Met and likely benign for *NEIL1* c.248G>T; p.Gly83Asp.

We genotyped PBL DNA samples from family members of the index carriers where possible to determine if the candidate allele segregated with disease (Figure 2). For family F1085, both the unaffected mother (PT0178) and sister (PT0177) of the index carrier cases did not carry *NTHL1* c.244C>T; p.Gln82Ter, suggesting that the variant allele may have been transmitted paternally. This observation is interesting as the paternal side of the family had numerous cancer cases including BC, OC, colorectal and pancreatic cancers. In family F1288, though the mother of the index *EXO1* c.1268-1G>T carrier case (PT0158) with BC and melanoma also carried the *EXO1* variant, her sibling (PT0180) and maternal female cousin (PT0181) both with BC were not carriers of the *EXO1* variant. These observations are interesting given the number of different types of cancer cases on the maternal side of the family. In family F1506, the index carrier harbouring *ERCC5* c.2556A>G also had a remarkable family history of diverse cancer types, whereas the index carrier of *NEIL1* c.248G>T; c.1268-1G>T from family F1601 had a cancer family history consistent with HBOC syndrome. The index carrier of *FANCC* c.897G>T; p.Arg299Ser from family F1606 reported a mother with OC and a father with lung cancer.

Identification of carriers of selected candidate variants in defined FC cancer study groups: Phase II

We genotyped or surveyed available genetic data of our candidate variants: *NTHL1* c.244C>T; p.Gln82Ter, *EXO1* c.1268-1G>T, *FANCC* c.897G>T; p.Arg299Ser, *ERCC5* c.2556A>G; p.Ile852Met and *NEIL1* c.248G>T; p.Gly83Asp in different FC OC and BC study groups and population-matched controls, regardless of their carrier status for *BRCA1* and *BRCA2* PVs (Figure 1B and Table S1). Carriers were identified in the sporadic OC study group with frequencies of 0.2% (1/435) for *EXO1* variant

TABLE 1 Characteristics of top-prioritized candidate variants identified in familial French Canadian cases with ovarian cancer.

Genomic features (hg19/GRCh37) ¹					
Gene	<i>NTHL1</i>	<i>EXO1</i>	<i>FANCC</i>	<i>ERCC5</i>	<i>NEIL1</i>
Transcript	NM_002528.7	NM_130398.4	NM_000136.3	NM_000123.4	NM_024608.4
Cytoband	16p13.3	1q43	9q22.32	13q33.1	15q24.2
Genome change	g.2096239G>A	g.242035333G>T	g.97887467C>G	g.103520485A>G	g.75641494G>A
Coding change	c.244C>T	c.1268-1G>T	c.897G>T	c.2556A>G	c.248G>T
Protein change	p.Gln82Ter	-	p.Arg299Ser	p.Ile852Met	p.Gly83Asp
Allele frequencies in gnomAD ²					
Non-cancer non-Finnish European	0.002 (235/118138)	-	8.8e-06 (1/113756)	9.74e-06 (1/102714)	0.001 (140/117290)
Clinical classification ³					
ClinVar (number of submissions)	PV (6); VUS (1)	-	VUS (2)	-	-
ACMG guidelines (implemented rule)	PV (PVS1)	LPV (PVS1/PP3/PM2)	LPV (PVS1/PP3/PM2)	VUS (PP3/PM2)	LB (PP3/BP1)
Predictions by <i>in silico</i> tools ⁴					
GERP++ v1.0	Conserved	Conserved	Conserved	Conserved	Conserved
PhyloP 100 way v4.2	Not conserved	Conserved	Not conserved	Not conserved	Conserved
PhastCons 100 way v4.2	Not conserved	Conserved	Not conserved	Conserved	Conserved
REVEL v4.2	-	-	Benign	Pathogenic	Pathogenic
MetaLR v4.2	-	-	Tolerated	Tolerated	Damaging
MetaSVM v4.2	-	-	Tolerated	Tolerated	Damaging
MetaRNN v4.2	-	-	Damaging	Damaging	Tolerated
CADD v1.6	Damaging	Damaging	Damaging	Damaging	Damaging
VEST v4.2	-	-	Damaging	Damaging	Damaging
EIGEN PC v4.2	-	-	Pathogenic	Pathogenic	Pathogenic
PROVEAN v4.2	-	-	Damaging	Damaging	Damaging
ADA v1.1	-	Affecting splicing	Affecting splicing	-	-
RF v1.1	-	Affecting splicing	Affecting splicing	-	-
MaxEntScan v2.0	-	Affecting splicing	Affecting splicing	-	-
SpliceAI	-	Affecting splicing	Affecting splicing	-	-

¹ Annotation of candidate variants based on the National Center for Biotechnology Information (NCBI) - Reference Sequence (RefSeq) database (tark.ensembl.org/web/manelist/) (103); ² Allele frequencies in the non-cancer, non-Finnish European controls from the Genome Aggregation Database (gnomAD) v2.1.1 database (gnomad.broadinstitute.org) (61); ³ Clinical classifications from ClinVar (ncbi.nlm.nih.gov/clinvar/) (82, 83) based on last revision in March 2022, and American College of Medical Genetics and Genomics (ACMG) guidelines (84, 104, 85, 105); ⁴ Applied *in silico* tools for conservation, damaging or affecting splicing selected based on their best performance (71, 74, 106, 107). Classification of variants by ACMG guidelines as: BP1: Benign Supporting Level 1; LB: Likely Benign; LPV: Likely Pathogenic Variant; PM2: Pathogenic Moderate Level 2; PP3: Pathogenic Supporting Level 3; PV: Pathogenic Variant; PVS1: Pathogenic Very Strong Level 1; VUS: Variant of Uncertain Significance; and (-): Not applicable/reported.

carriers and 0.5% (2/435) for carriers of each *NTHL1*, *ERCC5* or *NEIL1* variants, and one *NTHL1* carrier among sporadic BC cases (0.2%, 1/563). Carriers were not identified among index cases from HBOC and HBC families. *FANCC* variant carriers were not identified in any of these FC cancer study groups.

Carriers of all candidate variants in the FC controls are likely very rare as indicated by the observation that only one carrier was identified among 1025 FC sequencing-based controls (Table 2). Identifying a carrier of *NTHL1* c.224C>T; p.Gln82Ter was not surprising given the frequency of carriers of this variant = 0.002 in

the non-cancer non-Finnish European population in gnomAD v2.1.1 (Table S4). Overall, the carrier frequencies of our candidates are higher in cancer groups relative to our population-matched controls though the results were not significant (Table 2).

We investigated our variants in 8493 non-cancer SNP array genotyping-based controls from cancer-free FC population (see Table S1). None of the probes for variants in *EXO1*, *FANCC* and *ERCC5* were represented on any of the SNP arrays, nor was imputation possible as they were not represented in the Haplotype Reference Consortium (HRC.r1) release panel (43).

TABLE 2 Carrier frequency of candidate variants in French Canadian cancer cases and controls.

Gene	Variant	Study groups	Cancer cases tested	Number of tested participants (or families) per study group	Number of variant carriers (%)	p-value
<i>NTHL1</i>	c.244C>T; p.Gln82Ter	HBOC families	OC	42 (42)	0	–
		Sporadic OC cases	OC	435	2/435 (0.5)	0.213
		HBOC families	BC	33 (33)	0	–
		HBC families	BC	139 (139)	0	–
		Sporadic BC cases	BC	563	1/563 (0.2)	1.000
		FC sequencing-based controls	–	1025	1/1025 (0.1)	–
<i>EXO1</i>	c.1268-1G>T	HBOC families	OC	42 (42)	0	–
		Sporadic OC cases	OC	435	1/435 (0.2)	0.298
		HBOC families	BC	33 (33)	0	–
		HBC families	BC	139 (139)	0	–
		Sporadic BC cases	BC	563	0	–
		FC sequencing-based controls	–	1025	0	–
<i>FANCC</i>	c.897G>T; p.Arg299Ser	HBOC families	OC	42 (42)	0	–
		Sporadic OC cases	OC	435	0	–
		HBOC families	BC	33 (33)	0	–
		HBC families	BC	139 (139)	0	–
		Sporadic BC cases	BC	563	0	–
		FC sequencing-based controls	–	1025	0	–
<i>ERCC5</i>	c.2556A>G; p.Ile852Met	HBOC families	OC	42 (42)	0	–
		Sporadic OC cases	OC	435	2/435 (0.5)	0.213
		HBOC families	BC	33 (33)	0	–
		HBC families	BC	139 (139)	0	–
		Sporadic BC cases	BC	563	0	–
		FC sequencing-based controls	–	1025	0	–

(Continued)

TABLE 2 Continued

Gene	Variant	Study groups	Cancer cases tested	Number of tested participants (or families) per study group	Number of variant carriers (%)	p-value
NEIL1	c.248G>T; p.Gly83Asp	HBOC families	OC	42 (42)	0	-
		Sporadic OC cases	OC	435	2/435 (0.5)	0.213
		HBOC families	BC	33 (33)	0	-
		HBC families	BC	139 (139)	0	-
		Sporadic BC cases	BC	563	0	-
		FC sequencing-based controls	-	1025	0	-

Two-tailed p-values (not adjusted for multiple testing) calculated using Fisher's exact test in pair-wise comparisons between variant carriers in cancer study groups and population-matched controls. BC, Breast cancer; HBC, Hereditary breast cancer syndrome; HBOC, Hereditary breast and ovarian cancer syndrome; OC, Ovarian cancer; and (-), Not applicable.

However, we were able to determine the carrier frequency of the *NTHL1* (0.2%, 19/8493) and *NEIL1* (0.3%, 24/8493) variants (Table S5). The frequencies of these variants are consistent with those in the non-cancer non-Finnish European population in gnomAD v2.1.1. (Table S4), though we did not identify any carriers among the FC controls harbouring both *NTHL1* and *NEIL1* candidate variants.

Genetic analyses of other FC OC cases identified additional carriers of candidate variants: Phase III

Given genetic drift exhibited by the FC population that may result in higher frequency of candidate variant carriers with OC (34, 35, 43), we genotyped the germline of PBL DNA from additional 553 FC OC cases, which were recruited to the biobank but did not meet our criteria for the abovementioned defined OC study groups, and surveyed WES data available from 52 early-onset OC cases (Figure 1C and Table S1). We identified a total of six OC cases harbouring *NTHL1* c.244C>T (n=1), *EXO1* c.1268-1G>T (n=1), *FANCC* c.897G>T (n=1) and *NEIL1* c.248G>A (n=3) (Table S6).

Genetic analyses of sporadic early-onset FC OC cases identified other variants in our candidate genes: Phase III

Given the genetic heterogeneity observed in the FC population for rare PVs identified in *BRCA1*, *BRCA2*, *RAD51C* and *RAD51D* (33, 43, 44), we surveyed WES data available from 52 early-onset FC OC cases diagnosed at less than 50 years of age (Figure 1C and Table S1). The rationale for investigating this group is based on the plausibility that carriers of some of the known OC predisposing genes are more likely to develop OC before age of 60 as it is the median age of diagnosis of this disease in the general population with OC (112, 113). We identified a carrier of a rare variant in *NEIL1* c.569C>A; p.Pro276His that met our filtering and prioritizing criteria (Table S6). We genotyped this variant in our defined FC cancer study groups and controls, and we did not identify any other carriers of this *NEIL1* variant or in any of the additional 553 FC OC cases. We could not determine the variant carrier frequency in the 8493 genotyping-based FC cancer-free controls as it was not represented on the SNP array, and we could not impute this variant as it was not available in the HRC.r1 haplotype reference panel.

Evidence of loss of the wild-type allele in tumour DNA from carriers of candidate variants

As known OC risk genes behave as tumour suppressors where there is loss of the gene function in tumours is expected (114), we performed LOH analyses to investigate one of the classical mechanisms of inactivation of the loci of our candidate genes: *ERCC5*, *EXO1*, *FANCC*, *NEIL1* or *NTHL1*. We were able to perform LOH analyses on OC tumour DNA from our FC carriers of the candidate variants where

possible due to availability from the RRCancer biobank as follows: *NTHL1* c.244C>T; p.Gln82Ter (n=4), *FANCC* c.897G>T; p.Arg299Ser (n=1), *ERCC5* c.2556A>G; p.Ile852Met (n=2) and *NEIL1* c.248G>T; p.Gly83Asp (n=3) (Table S7). Chromatograms of bidirectional Sanger sequencing of OC tumour DNA and case matched normal were inspected for allelic content. We observed partial or complete loss of the wild-type alleles in the tumour DNA from two *NEIL1* variant carriers, two *NTHL1* variant carriers and one *ERCC5* variant carrier (Table S7); chromatogram of one example is shown in Figure 3. Moreover, we observed partial or complete loss of the wild-type alleles in the tumour DNA from the left and the right ovaries from both *NTHL1* variant carriers having bilateral OC. There was no clear evidence for loss of the wild-type allele in the remaining samples from tumour DNA from carriers of *FANCC* or *EXO1*. However, loss of the variant allele was observed in the tumour DNA from *FANCC* and *EXO1* variant carriers (Table S7).

Genetic analyses of non-FC cases identified other candidate variants in our gene candidates: Phase IV

To determine the relevance of our candidate genes to OC in non-FC populations, we investigated the spectrum and prevalence

of rare variants in our candidate genes in genetic data from three defined non-FC study groups (Figure 1D and Table S1). We applied our filtering and prioritizing criteria to WES data that was available from the germline of PBL DNA from: (1) 9 index OC cases from MIX familial OC cases; (2) 516 index AUS OC cases from HBOC and sporadic disease; and (3) 412 OC cases from Pan-Cancer – TCGA. In these study groups, we identified: one of the 9 MIX familial OC cases (11.1%) harbouring a *NTHL1* variant; 17 of 516 AUS OC cases (3.3%) harbouring 11 variants in *NTHL1* (n=6), *NEIL1* (n=3), *ERCC5* (n=1) and *EXO1* (n=1); 12 of 412 Pan-Cancer – TCGA OC cases (2.9%) harbouring 10 variants in *NEIL1* (n=4), *EXO1* (n=2), *FANCC* (n=2), *NTHL1* (n=1) and *ERCC5* (n=1) (Table S6 and Figure 4). Noteworthy, the frequency of all LoF rare (MAF \leq 0.005) variants in these genes in the cancer-free gnomAD v2.1.1 controls as follows: 112/1,563 (0.07) in *ERCC5*; 105/1,207 (0.09) in *EXO1*; 87/943 (0.09) in *FANCC*; 85/811 (0.1) in *NEIL1* and 47/629 (0.07) in *NTHL1*. Collectively, candidate variants we identified in cancer cases are comprised of three nonsense, four frameshift, three alternative splicing and nine missense variants. Variants in and *NEIL1* (n=3), *EXO1* (n=2), *NTHL1* (n=2), *ERCC5* (n=1) and *FANCC* (n=1) were LoF variants and classified as PVs or LPVs in ClinVar and/or by ACMG guidelines. The remaining variants were missense predicted to be PV or LPVs by our set of in silico tools. Some of these variants were those already identified

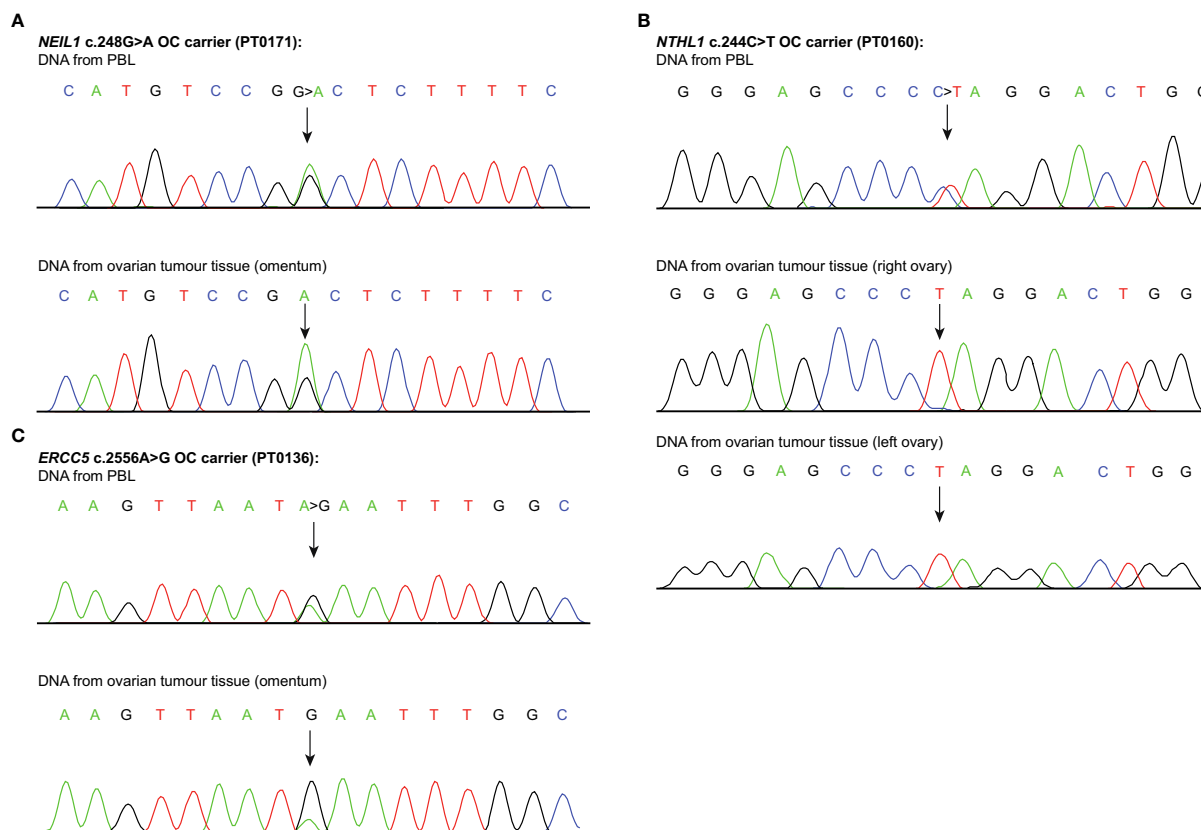


FIGURE 3

Loss of heterozygosity analyses of candidate genes loci. Sanger sequencing chromatograms showing loss of heterozygosity (LOH) analyses of the candidate variants (see Table S7), in genomic peripheral blood lymphocyte (PBL) DNA, ovarian tumour tissue DNA from carriers of (A) *NEIL1* c.248G>T; p.Gly83Asp (PT0171); (B) *NTHL1* c.244C>T; p.Gln82Ter (PT0160); and (C) *ERCC5* c.2556A>G; p.Ile852Met (PT0136). Each variant is indicated by an arrow. One example of such genetic event per candidate variant carrier is shown.

in our FC study groups: three of 516 (0.6%) AUS OC cases and two of 412 (0.5%) Pan-Cancer – TCGA OC cases carried *NTHL1* c.244C>T; p.Gln82Ter, while one each of 516 (0.2%) AUS OC cases and 412 (0.2%) Pan-Cancer – TCGA OC cases carried *NEIL1* c.248G>A; p.Gly83Asp.

We identified one Pan-Cancer – TCGA OC carrier of a synonymous variant in *NEIL1* c.159C>T; p.Gly53Gly that was predicted to affect splicing using SpliceAI (115) that may result in gain of a new donor splice site.

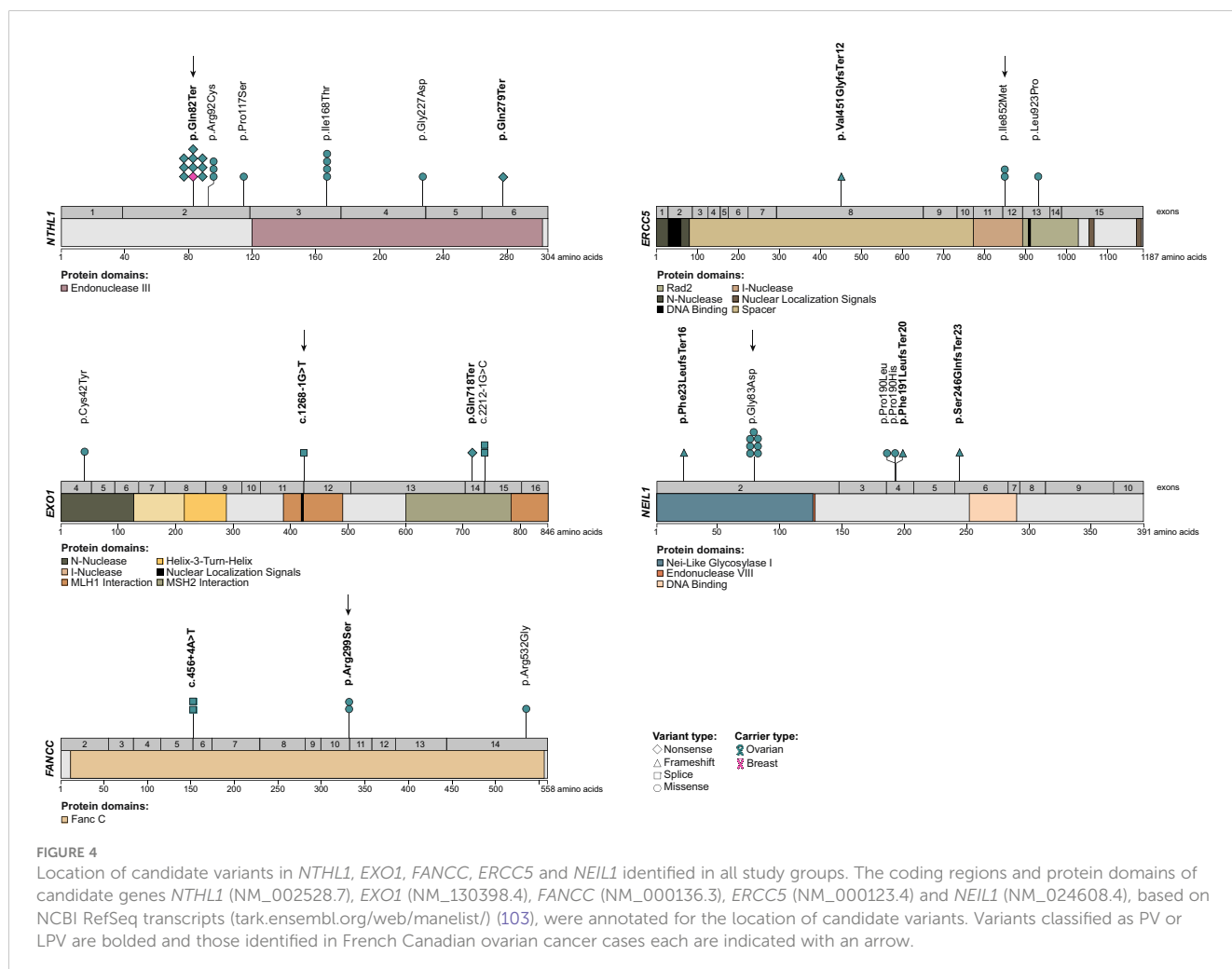
Most candidate variant carriers do not harbour co-occurring pathogenic variants in known OC predisposing genes

We investigated whether OC carriers harbouring any of our candidate variants may also harbour PVs in known OC predisposing genes (n=11) (nccn.org/guidelines/category_2) (Tables S8, S9). Only one of the 15 FC OC carriers investigated carried both *NEIL1* c.248G>A; p.Gly83Asp (PT0175) and *BRCA1* c.5102_5103del; p.Leu1701GlnfsTer14. This *BRCA1* variant is one of the most frequently occurring PVs in OC and BC cases from the FC population (33, 40, 41, 116).

Regarding non-FC OC carriers, one of the 17 OC AUS carriers of *NEIL1* c.248G>A; p.Gly83Asp (PT0314) also harboured a variant in a known OC risk gene *RAD51C* c.145+1_145+2insC, which was classified as PV by ACMG guidelines (Table S9) and was reported previously in an investigation of this study group (17). Two of the 13 carriers of our candidate variants from OC Pan-Cancer – TCGA project cases, harboured variants in known OC risk genes: a carrier of *NTHL1* c.244C>T; p.Gln82Ter (PT0261) also harboured *BRCA2* c.5065_5066insA; p.Ala1689AspfsTer6 and a carrier of *EXO1* c.2152C>T; p.Gln718Ter (PT0263) also harboured *BRCA2* c.1029del; p.Lys343AsnfsTer6. Both *BRCA2* variants have been classified as PVs in the ClinVar database and by ACMG guidelines.

Discussion

Our investigation of potentially deleterious variants in 468 genes that play a direct or associated role in various DNA repair pathways in the FC population exhibiting genetic drift identified LoF and potentially deleterious missense variants as candidates for OC predisposition in *ERCC5*, *EXO1*, *FANCC*, *NEIL1* or *NTHL1*. Genotyping analyses of independently ascertained FC cancer study groups identified multiple carriers with OC harbouring the same



variant, which is likely due to common ancestors within the FC population of Quebec (33–35). Overall, carriers of each variant are rare, each accounting for <1%, but collectively 9.6% of 52 familial OC cases with at least two or more OC cases and 1.6% of all 435 sporadic OC cases versus 0.1% of the population-matched controls. It is notable that none of the variants were found in known OC predisposing genes (nccn.org/guidelines/category_2), confirming prior findings from either clinical testing or our WES analyses of this group of cases (24, 43, 47).

Our candidate variants were identified by investigating a curated list of 468 known or putative genes involved in different DNA repair pathways. *ERCC5* is known to be involved in the NER pathway as an endonuclease, but it has been shown that this gene is also involved in the BER (117) and HR (108) pathways. *EXO1* is involved in the HR and MMR pathways as an exonuclease (118, 119). *FANCC* plays a role in the FA pathway as a member of the core complex (120). *NTHL1* and *NEIL1* are DNA glycosylases in the BER pathway (121). Although the role of these genes in conferring risk to hereditary OC requires further investigation with larger cohorts, a recent Australian study of familial and sporadic OC cases reported that there was a statistically significant difference in the frequency of germline LoF variants in the single-stranded DNA repair pathway genes involved in BER, NER and MMR in OC cases versus non-cancer controls (17). On the other hand, homozygous or compound heterozygous PVs in *NTHL1* have been linked to Familial adenomatous polyposis - 3 (MIM: 616415) (122) and most recently a multi-tumour phenotype (123). Homozygous or compound heterozygous PVs in *FANCC* and *ERCC5* are known to be linked to autosomal recessive disorders, Fanconi anemia complementation group C (MIM: 227645) and Xeroderma pigmentosum complementation group G (MIM: 278780), respectively, known to exhibit increased risk to cancer.

Three of our top candidate variants were predicted to be LoF: *NTHL1* c.244C>T; p.Gln82Ter, *EXO1* c.1268-1G>T and *FANCC* c.897G>T; p.Arg299Ser. *NTHL1* c.244C>T; p.Gln82Ter has been independently reported in the literature due to its frequency, while *FANCC* c.897G>T; p.Arg299Ser has been reported only in ClinVar database. The introduction of a termination codon in *NTHL1* p.Gln82Ter is predicted to affect *NTHL1* protein production, eliciting its classification as PV in ClinVar and by ACMG guidelines. *EXO1* c.1268-1G>T and *FANCC* c.897G>T; p.Arg299Ser were predicted to affect splicing by all of our selected in silico tools. As RNA was not available from carriers of these variants, we were unable to investigate their effect on the gene transcripts. We applied a stringent criteria for prioritizing missense variants using a selected set of high performance in silico prediction tools (71–74). *ERCC5* c.2556A>G; p.Ile852Met and *NEIL1* c.248G>T; p.Gly83Asp, which have been independently reported in the literature due to their frequencies, were among our top prioritized missense variants. A recent study assessed the performance of 44 in silico tools with 70 tool-threshold combinations in predicting missense variants using a curated dataset of over 9,000 missense variants in five OC and/or BC risk genes that were classified as deleterious or tolerated based on different functional assays (73). Two of the in silico tools that were selected in our analysis, REVEL with a threshold of >0.7 and VEST with a threshold of ≥0.5 prediction scores of a missense variant being deleterious, were shown to have the best performance of 79% and 74%, respectively (73).

Moreover, a combination of both tools with these prediction score thresholds boosts the prediction performance up to 81% (73). *NEIL1* c.248G>A; p.Gly83Asp had REVEL and VEST scores of >0.7, while *ERCC5* c.2556A>G; p.Ile852Met had a REVEL score at the threshold of 0.7 and VEST score of 0.9. The biological impact of *ERCC5* p.Ile852Met is unknown, though the variant alters a codon in the highly conserved I-Nuclease domain (see Figure 4), which may impact *ERCC5* endonuclease activity. Whereas, cells expressing *NEIL1* p.Gly83Asp have been shown to increase levels of stalled replication forks and double-strand breaks as compared to wild-type *NEIL1* (109). For the missense variants identified in the non-FC cases, eight of the nine missense variants were predicted to be deleterious by REVEL (>0.7) and/or VEST (≥0.5). One variant, *NTHL1* c.349C>T; p.Pro117Ser, was found to have a REVEL score of 0.6 and VEST score of 0.5, which is within the intermediate window of prediction scores (REVEL <0.7->0.4; VEST <0.5) where the threshold of predicting missense variants being tolerated is <0.4 (73). It is evident that in silico tools are being developed with increasing improvement in their performance and are useful alternatives to biological modelling of variants for selecting and prioritizing missense candidates for further characterization (73, 124).

We applied stringent criteria to select top candidates for further analyses as it was not feasible to perform WES on all our FC OC study cases. Though selecting for rare variants with MAF ≤0.005 aligns with our hypothesis for identifying candidate moderate- to high-risk variants with the assumption that new candidate genes are transmitted *via* an autosomal dominant mode of inheritance (63, 114, 125, 126), we filtered out our recently reported, LPV *FANCI* c.1813C>T; p.Leu605Phe (24). Notable is that the one family harbouring this *FANCI* variant among the 13 familial cases investigated in this study, did not harbour any of our top candidates. As our strategy selected but not eliminated top candidates, further research is required to determine their relevance to OC risk. Moreover, as we have shown in our studies of predicted missense identified in *RAD51C* and *RAD51D* (43) and *FANCI* (24) in OC cases from the FC population, modeling variants by in cellulo assays would provide further evidence for their relevance in OC biology.

Some of our candidate variants or others in these genes that met our selection criteria were identified in 3.3% of 937 non-FC familial or sporadic OC cases of mostly European ancestry 0-0.004% in gnomAD controls, suggesting that our gene candidates may be relevant in other populations. Though this observation was not unexpected, given that approximately 55% of our non-FC OC cases were from the same Australian study group (17), they are also consistent with our analysis of the Pan-Cancer TCGA OC cases. During the course of this study, a recent report investigated a set of DNA repair genes in 33 different cancers from Pan-Cancer TCGA, and they found that missense variants predicted damaging are statistically enriched in OC cases (127). Moreover, carriers of LoF variants in some of our candidate genes have been reported in HBOC families such as *ERCC5* (128, 129), *FANCC* (130–133) or *NTHL1* (111, 134, 135) from different populations. Interestingly, LoF variants in *ERCC5* (110), *FANCC* (102) or *NTHL1* (111, 110) have been associated with hereditary BC cases in the context of HBOC families. On the other hand, our literature search did not identify reports of germline PVs in *EXO1* or *NEIL1* in OC, although variants in these genes have been reported in

the context of other hereditary cancers such as colorectal cancer (136–138). However, common variants (MAF>1%) in *EXO1* have been associated with OC risk (139).

Although we are limited by sample size, we did not observe any striking clinical characteristics regarding age at diagnosis or histopathology of OC disease in carriers of our candidate variants. The average age at diagnosis with OC in FC variant carriers (average=58; median=60 years) is comparable to that of AUS variant carriers as well as Pan-Cancer – TCGA variant carriers (average=59; median=59 years), which in turn is comparable with that of carriers of *BRCA2* PVs (average=58 years) (41) and the general population (median=63 years) (112, 113). The majority of OC cases harbouring the candidate variants had HGSC (91.1%, 41/45), which is the most common subtype of OC reported in epithelial ovarian cancer (49) and thus is overrepresented in our study groups (56). We did observe three carriers of *EXO1* c.1268-1G>T, *NEIL1* c.248G>A; p.Gly83Asp and *NTHL1* c.244C>T; p.Gln82Ter with mixed histology (serous mixed with endometrioid or unspecified cell type; and endometrioid mixed with mucinous) (see Table S8). Interestingly, a survey of our candidate variants in the Ovarian Cancer Association Consortium (OCAC) database (ocac.ccge.medschl.cam.ac.uk/data-projects/, accessed on 15 June 2020), showed that *NEIL1* c.248G>T; p.Gly83Asp was significantly associated with OC overall (odds ratio [OR] = 1.5; p = 0.038), and this association was stronger with the endometrioid subtype (OR = 3.75; p = 0.00008) (see Table S10). *NTHL1* c.244C>T; p.Gln82Ter showed a higher OR = 1.5 in clear cell subtype but was not statistically significant (p = 0.36) (see Table S10). These observations are interesting as *NTHL1* and *NEIL1* are involved in repairing single stranded-DNA breaks via BER pathway. MMR genes as also involved in repairing single stranded-DNA breaks via MMR pathway that are associated with conferring an increased risk to the endometrioid and clear cell histological subtypes of OC (7). We could not investigate the other candidate variants *ERCC5* c.2556A>G; p.Ile852Met, *EXO1* c.1268-1G>T and *FANCC* c.897G>T; p.Arg299Ser from this genotyping-based database as they were not represented in the SNP arrays, which could be due to their rarity in the general population.

The role of our candidate genes in the etiology of OC is unknown, though LOH analyses suggest that loss of function of some of our candidate genes in tumour cells may be important in tumorigenesis of OC as has been demonstrated for known OC risk genes (114). We showed loss of the wild-type allele in tumours from carriers of *ERCC5* c.2556A>G, *NEIL1* c.248G>T; p.Gly83Asp or *NTHL1* c.244C>T; p.Gln82Ter. We also showed loss of the wild-type allele in the tumour DNA from the left and the right ovaries from two *NTHL1* c.244C>T carriers with bilateral OC. This suggests the possibility that loss of wild-type allele occurred at an early stage in tumorigenesis (140). However, we observed no LOH in tumour DNA from one carrier each of *ERCC5* or *NEIL1* variants and two carriers of *NTHL1* variant (see Table S7). In previous studies we have also demonstrated complete or partial loss of wild-type alleles in tumour DNA from FC carriers of *RAD51C* c.705G>T; p.Lys235Asn (43), *RAD51D* c.620C>T; p.Ser207Leu (43, 44) and *FANCI* c.1813C>T; p.Leu605Phe (24) also from the analyses of RRCancer biobank materials. We also observed no LOH in some of the tumour DNA from OC carriers of our *RAD51C* c.705G>T; p.Lys235AsnAs and *RAD51D* c.620C>T; p.Ser207Leu, as the DNA was extracted post-chemotherapy treatment, suggesting the

possibility of stromal cell contamination (43, 44). Indeed, all of the DNA tumours from the four cases analysed in this study were confirmed to be extracted after chemotherapy. Interestingly, our LOH analyses in this study provided evidence for loss of variant allele from other candidate variants *EXO1* c.1268-1G>T and *FANCC* c.897G>T carriers. It is not clear if this is due to contaminating normal tissues as our analyses were not performed on selected tumour cells as HGSC samples are often enriched in tumour cells (24). Srinivasan et al. (141) recently reported that approximately 20% of the 55 investigated cancers, including OC, showed a retention of the wild-type alleles in the high-penetrant genes such as *BRCA1* and *BRCA2* (141). A retention of the *RAD51D* c.620C>T; p.Ser207Leu was also observed in tumour DNA from an OC carrier (44). It is not clear if such cases reflect a reversion of variant to wild type allele as has been shown with *BRCA1* or *BRCA2* carriers in the context of developing resistance to cisplatin or the targeted therapy poly (ADP-ribose) polymerase (PARP) inhibitors (142, 143). Further research is required at the tumour cell level to determine biological impact of variants in the context of wild-type alleles in carriers.

There are several limitations in this original study that should be acknowledged. This study was limited in the sample size of the FC OC families, sporadic cases and cancer-free controls (study phase I and II). The small sample size in our cases and controls did not allow us to estimate the associated risk with OC of any of the candidate variants to support their candidacy as OC predisposing variants/genes (144) as such risk assessment requires thousands of cases and controls (12, 13, 1, 19). However, our highly selected OC families regardless of their status of PVs in *BRCA1* and *BRCA2* are very rare. The proportion of families with two first-degree relatives with OC is estimated to be less than 5%, and less than 1% for those with more than two first-degree relatives with OC (145–147). This study was also limited in lacking of biological samples that were required to confirm the predicted effect on RNA splicing of two of the candidate variants *EXO1* c.1268-1G>T and *FANCC* c.897G>T; p.Arg299Ser. Likewise, we were not able to perform LOH analysis on all OC carriers of the candidate variants as well as genotyping PBL DNA of the other family members to further support the candidacy of our variants. This was due to the fact that the extracted DNA from only PBL and/or tumour specimens were biobanked for almost all of the OC cases. Moreover, histopathological blocks for DNA and/or RNA extraction of these carriers were not available from the respective biobanks as some of the cancer families and cases date back to the early 2000s (see Table S1).

In conclusion, our WES and genetic analyses of 468 genes directly or associated with DNA repair pathways in study groups from a genetically defined population identified candidate variants in *ERCC5*, *EXO1*, *FANCC*, *NEIL1* or *NTHL1*. The genetic analyses of these variants and genes in non-FC OC study groups implicate these genes in other populations. Genetic epidemiology of variant carriers and functional assays to assess the biological impact of variant proteins could elucidate the effect of candidate variants to OC risk.

Data availability statement

The datasets presented in this article are not readily available because sequencing data for familial and sporadic ovarian cancer (OC)

cases, CARTaGENE, McGill University-Montreal Neurological Institute (MNI) and Université de Sherbrooke-The Genetics of Glucose Regulation in Gestation and Growth (Gen3G) will be returned to their respective biobanks at the conclusion of our study of OC predisposing genes which is still ongoing. For more information concerning these data contact Patricia N. Tonin at patricia.tonin@mcgill.ca. The data from the analyses of investigation of The Ovarian Cancer Association Consortium (OCAC) and The Genome Aggregation Database (gnomAD) are available from each of these data resource banks.

Ethics statement

This project has received approval from The McGill University Health Centre (MUHC) REB (MP-37-2019-4783). The patients/participants provided their written informed consent to participate in this study.

Author contributions

WA and PT conceived and designed the study. WA performed whole exome sequencing and bioinformatics analyses, conducted, and interpreted the genotyping and Sanger sequencing and wrote the initial drafts of this manuscript. TR performed the bioinformatics pipelines. CF aided in bioinformatics analysis. CS aided in collecting and reviewing all clinical pedigrees with associated information. KO performed imputation and provided required information from CARTaGENE genotyping-based data. DS and IC provided required information from in-house sequencing-based Australian ovarian cancer cases. JB and TP provided required information from The Cancer Genomic Atlas (TCGA) sequencing-based breast and ovarian cancer cases. DS, GR and LB provided required information from in-house sequencing-based control data. A-MM-M, DP, WF and ZH provided the study samples and clinical information. CG oversaw the statistical analyses. JR oversaw the WES analysis and bioinformatics pipeline. PT designed and oversaw all aspects of the study. All authors contributed to the article and approved the submitted version.

Funding

The work was supported in part by the Department of Medicine, McGill University to PT; Saudi Arabian Cultural Bureau to PT and WA; The Canadian Institute for Health Research (CIHR) operating grants (PCC-156736 to PT, CG and JR) and (PJT-156124 to PT and JR); Cancer Research Society and Ovarian Cancer Canada partnership grant (21123 to PT); Department of Medicine, McGill University Grant to PT; the Fond de la recherche du Québec en santé (FRQS) and Quebec Breast Cancer Foundation network grants to PT; Compute Canada resource allocation project wst-164 and Genome Canada Genome Technology Platform award to JR; Institut du cancer de Montréal Fonds Défi Spyder and Anne-Marie Chagnon which covers funds from platform to DP and A-MM-M; Gen3G has been supported

over the time by FRQS grant (20697 to Marie-France Hivert); CIHR grant (MOP-115071 to Dr. Marie-France Hivert) and (PJT-152989 to LB); American Diabetes Association (ADA) accelerator award (1-15-ACE-26 to Marie-France Hivert). LB is a senior research scholar from FRQS. The sequencing of the Gen3G offspring has been sponsored by Fonds de recherche du Québec, McGill University and Université de Sherbrooke. Ovarian tumour banking was supported by the Banque de tissus et de données of the Réseau de recherche sur le cancer of the FRQS affiliated with the Canadian Tumour Repository Network (CTRNet). WA was supported by a Scholarship Award by Ministry of Education in Saudi Arabia and CF was supported in part by RI-MUHC Scholarship Award and James O and Maria Meadows Award.

Acknowledgments

We acknowledge the Banque de tissus et de données of the RRCancer of the Fonds recherche Québec - Santé (FRQS), which is affiliated with the Canadian Tissue Repository Network (CTRNet) for providing the DNA samples from cancer families. We also acknowledge CARTaGENE biobank; and Drs. Patrice Peron, Luigi Bouchard, and Marie-France Hivert who established the Gen3G cohort and sequencing. Pierre-Étienne Jacques and Simon Gravel contributed to the generation and analysis of sequencing data in Gen3G. The authors acknowledge Manon De Ladurantaye and Lise Portelance for providing all clinical data and DNA from the RRCancer biobank; Liliane Meunier for providing the sections from histopathological blocks; Supriya Behl, Nancy Hamel and Celine Domecq for providing genetic data for sporadic breast cancer cases; and Suzanna Arcand for organizing samples for whole exome sequencing (WES). The authors acknowledge Mary Fujiwara for reviewing the final draft of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1111191/full#supplementary-material>

References

- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* (1994) 266(5182):66–71. doi: 10.1126/science.7545954
- Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* (1995) 378(6559):789–92. doi: 10.1038/378789a0
- Nielsen FC, van Overeem Hansen T, Sørensen CS. Hereditary breast and ovarian cancer: new genes in confined pathways. *Nat Rev Cancer* 2016 16:9 (2016) 16:599–612. doi: 10.1038/nrc.2016.72
- Turnbull C, Sud A, Houlston RS. Cancer genetics, precision prevention and a call to action. *Nat Genet* (2018) 50(9):1212–8. doi: 10.1038/s41588-018-0202-0
- Aarnio M, Sankila R, Pukkala E, Salovaara R, Aaltonen LA, De A, et al. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Wiley Online Library* (1999) 81:214–8. doi: 10.1002/(SICI)1097-0215(19990412)81:2<214::AID-IJC8>3.0.CO;2-L
- Stoll J, Rosenthal E, Cummings S, Willmott J, Bernhisel R, Kupferl SS. No evidence of increased risk of breast cancer in women with lynch syndrome identified by multigene panel testing. *JCO Precis Oncol* (2020) 4:51–60. doi: 10.1200/PO.19.00271
- Lynch HT, Casey MJ, Snyder CL, Bewtra C, Lynch JF, Butts M, et al. Hereditary ovarian carcinoma: Heterogeneity, molecular genetics, pathology, and management. *Mol Oncol* (2009) 3(2):97–137. doi: 10.1016/j.molonc.2009.02.004
- Pavanello M, Chan IH, Arif A, Pharoah PD, Gayther SA, Ramus SJ. Rare germline genetic variants and the risks of epithelial ovarian cancer. *Cancers* 2020 Vol. 12 Page 3046 (2020) 12:3046. doi: 10.3390/CANCERS12103046
- Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet* (2010) 42(5):410–4. doi: 10.1038/ng.569
- Loveday C, Turnbull C, Ramsay E, Hughes D, Ruark E, Frankum JR, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet* (2011) 43(9):879–82. doi: 10.1038/ng.893
- Rafnar T, Gudbjartsson DF, Sulem P, Jonasdottir A, Sigurdsson A, Jonasdottir A, et al. Mutations in BRIP1 confer high risk of ovarian cancer. *Nat Genet* (2011) 43(11):1104–7. doi: 10.1038/ng.955
- Ramus SJ, Song H, Dicks E, Tyrer JP, Rosenthal AN, Intermaggio MP, et al. Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. *J Natl Cancer Institute* (2015) 107(11):djv214. doi: 10.1093/jnci/djv214
- Song H, Dicks E, Ramus SJ, Tyrer JP, Intermaggio MP, Hayward J, et al. Contribution of germline mutations in the RAD51B, RAD51C, and RAD51D genes to ovarian cancer in the population. *J Clin Oncol* (2015) 33:2901–7. doi: 10.1200/JCO.2015.61.2408
- Suszynska M, Ratajska M, Kozlowski P. BRIP1, RAD51C, and RAD51D mutations are associated with high susceptibility to ovarian cancer: Mutation prevalence and precise risk estimates based on a pooled analysis of ~30,000 cases. *J Ovarian Res* (2020) 13(1):50. doi: 10.1186/s13048-020-00654-3
- Kanchi KL, Johnson KJ, Lu C, McLellan MD, Leiserson MDM, Wendl MC, et al. Integrated analysis of germline and somatic variants in ovarian cancer. *Nat Commun* (2014) 5:3156. doi: 10.1038/ncomms4156
- Lu HM, Li S, Black MH, Lee S, Hoiness R, Wu S, et al. Association of breast and ovarian cancers with predisposition genes identified by Large-scale sequencing. *JAMA Oncol* (2019) 5:51–7. doi: 10.1001/JAMAONCOL.2018.2956
- Subramanian DN, Zethoven M, McInerney S, Morgan JA, Rowley SM, Lee JEA, et al. Exome sequencing of familial high-grade serous ovarian carcinoma reveals heterogeneity for rare candidate susceptibility genes. *Nat Commun* (2020) 11:1–11. doi: 10.1038/s41467-020-15461-z
- Yang X, Leslie G, Dorozuk A, Schneider S, Allen J, Decker B, et al. Cancer risks associated with germline PALB2 pathogenic variants: An international study of 524 families. *J Clin Oncol* (2020) 38:674–85. doi: 10.1200/JCO.19.01907
- Song H, Dicks EM, Tyrer J, Intermaggio M, Chenevix-Trench G, Bowtell DD, et al. Population-based targeted sequencing of 54 candidate genes identifies PALB2 as a susceptibility gene for high-grade serous ovarian cancer. *J Med Genet* (2021) 58:305–13. doi: 10.1136/JMEDGENET-2019-106739
- Szymanska-Pasternak J, Szymanska A, Medrek K, Imyanitov EN, Cybulski C, Gorski B, et al. CHEK2 variants predispose to benign, borderline and low-grade invasive ovarian tumors. *Gynecologic Oncol* (2006) 102:429–31. doi: 10.1016/J.YGYNO.2006.05.040
- Dicks E, Song H, Ramus SJ, Van Oudenhove E, Tyrer JP, Intermaggio MP, et al. Germline whole exome sequencing and large-scale replication identifies <i>FANCM</i> as a likely high grade serous ovarian cancer susceptibility gene. *Oncotarget* (2017) 8(31):50930–40. doi: 10.18632/oncotarget.15871
- Zhu Q, Zhang J, Chen Y, Hu Q, Shen H, Huang RY, et al. Whole-exome sequencing of ovarian cancer families uncovers putative predisposition genes. *Int J Cancer* (2019) 146(8):2147–55. doi: 10.1002/ijc.32545
- Heikinen K, Karppinen SM, Soini Y, Mäkinen M, Winqvist R. Mutation screening of Mre11 complex genes: indication of RAD50 involvement in breast and ovarian cancer susceptibility. *J Med Genet* (2003) 40(12):e131. doi: 10.1136/JMG.40.12.E131
- Fierheller CT, Guitton-Sert L, Alenezi WM, Revil T, Oros KK, Gao Y, et al. A functionally impaired missense variant identified in French Canadian families implicates FANCI as a candidate ovarian cancer-predisposing gene. *Genome Med* (2021) 13(1):186. doi: 10.1186/S13073-021-00998-5
- Pharoah PDP, Dunning AM, Ponder BAJ, Easton DF. Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer* (2004) 4:850–60. doi: 10.1038/nrc1476
- Rahman N. Realizing the promise of cancer predisposition genes. *Nature* (2014) 505(7483):302–8. doi: 10.1038/nature12981
- Ramus SJ, Gayther SA. The contribution of BRCA1 and BRCA2 to ovarian cancer. *Mol Oncol* (2009) 3(2):138–50. doi: 10.1016/j.molonc.2009.02.001
- Milanowska K, Krwawicz J, Papaj G, Kosiński J, Poleszak K, Lesiak J, et al. REPAIRtoire—a database of DNA repair pathways. *Nucleic Acids Res* (2011) 39(Database issue):D788–92. doi: 10.1093/NAR/GKQ1087
- Chae YK, Anker JF, Carneiro BA, Chandra S, Kaplan J, Kalyan A, et al. Genomic landscape of DNA repair genes in cancer. *Oncotarget* (2016). doi: 10.18632/oncotarget.8196
- Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, et al. Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas. *Cell Rep* (2018) 23(1):239–254.e6. doi: 10.1016/j.celrep.2018.03.076
- Glentis S, Dimopoulos AC, Rouskas K, Ntrisots G, Evangelou E, Narod SA, et al. Exome sequencing in BRCA1- and BRCA2-negative Greek families identifies MDM1 and NBEAL1 as candidate risk genes for hereditary breast cancer. *Front Genet* (2019) 10:1005. doi: 10.3389/FGENE.2019.01005
- Sherill-Rofe D, Rahat D, Findlay S, Mellul A, Guberman I, Braun M, et al. Mapping global and local coevolution across 600 species to identify novel homologous recombination repair genes. *Genome Res* (2019) 29:439–48. doi: 10.1101/gr.241414.118
- Fierheller CT, Alenezi WM, Tonin PN. The genetic analyses of French Canadians of Quebec facilitate the characterization of new cancer predisposing genes implicated in hereditary breast and/or ovarian cancer syndrome families. *Cancers* (2021) 13(14):3406. doi: 10.3390/CANCERS13143406
- Scriver CR. Human genetics: lessons from Quebec populations. *Annual review of genomics and human genetics. Annu Rev Genomics Hum Genet* (2001) 2:69–101. doi: 10.1146/annurev.genom.6.080604.162340
- Laberge AM, Michaud J, Richter A, Lemyre E, Lambert M, Brais B, et al. Population history and its impact on medical genetics in Quebec. *Clin Genet* (2005) 68(4):287–301. doi: 10.1111/j.1399-0004.2005.00497.x
- Erkko H, Xia B, Nikkilä J, Schleutker J, Syrjäkoski K, Mannermaa A, et al. LETTERS a recurrent mutation in PALB2 in Finnish cancer families. *Nature* (2007) 446(7133):316–9. doi: 10.1038/nature05609
- Cybulski C, Carrot-Zhang J, Klunlich W, Rivera B, Kashyap A, Wokolorczyk D, et al. Germline RECQL mutations are associated with breast cancer susceptibility. *Nat Genet* (2015) 47:643–6. doi: 10.1038/NG.3284
- Charbonneau H, Desjardins B, Légaré J, Denis H. “The population of the st. Lawrence Valley 1608-1760 | bibliographie sur l’histoire de montréal.”. In: Haines MR, Steckel RH, editors. *A population history of north America*. Cambridge University Press (2000). p. 99–142.
- Gagnon A, Heyer E. Fragmentation of the québec population genetic pool (Canada): Evidence from the genetic contribution of founders per region in the 17th and 18th centuries. *Am J Phys Anthropology* (2001) 114:30–41. doi: 10.1002/1096-8644(200101)114:1<30::AID-AJPA1003>3.0.CO;2-L
- Tonin PN, Mes-Masson A-M, Futreal PA, Morgan K, Mahon M, Foulkes WD, et al. Founder BRCA1 and BRCA2 mutations in French Canadian breast and ovarian cancer families. *Am J Hum Genet* (1998) 63(5):1341–51. doi: 10.1086/302099
- Oros KK, Ghadirian P, Greenwood CMT, Perret C, Shen Z, Paredes Y, et al. Significant proportion of breast and/or ovarian cancer families of French Canadian descent harbor 1 of 5 BRCA1 and BRCA2 mutations. *Int J Cancer* (2004) 112(3):411–9. doi: 10.1002/ijc.20406
- Foulkes WD, Ghadirian P, Akbari MR, Hamel N, Giroux S, Sabbaghian N, et al. Identification of a novel truncating PALB2 mutation and analysis of its contribution to early-onset breast cancer in French-Canadian women. *Breast Cancer Res* (2007) 9(6):R83. doi: 10.1186/bcr1828
- Alenezi WM, Milano L, Fierheller CT, Serruya C, Revil T, Oros KK, et al. The genetic and molecular analyses of RAD51C and RAD51D identifies rare variants implicated in hereditary ovarian cancer from a genetically unique population. *Cancers* 2022 (2022) 14:2251. doi: 10.3390/CANCERS14092251
- Rivera B, Iorio M, Frankum J, Nadaf J, Fahiminiya S, Arcand SL, et al. Functionally null RAD51D missense mutation associates strongly with ovarian carcinoma. *Cancer Res* (2017) 77:4517–29. doi: 10.1158/0008-5472.CAN-17-0190
- Chong G, Jarry J, Marcus V, Thiffault I, Winocour S, Monczak Y, et al. High frequency of exon deletions and putative founder effects in French Canadian lynch syndrome families. *Hum Mutat* (2009) 30:E797–812. doi: 10.1002/humu.21056

46. Castellsagué E, Liu J, Volenik A, Giroux S, Gagné R, Maranda B, et al. Characterization of a novel founder MSH6 mutation causing lynch syndrome in the French Canadian population. *Clin Genet* (2015) 87:536–42. doi: 10.1111/CGE.12526
47. Alenezi WM, Fierheller CT, Revil T, Serruya C, Mes-Masson A-M, Foulkes WD, et al. Case review: Whole-exome sequencing analyses identify carriers of a known likely pathogenic intronic BRCA1 variant in ovarian cancer cases clinically negative for pathogenic BRCA1 and BRCA2 variants. *Genes* 2022 Vol. 13 Page 697 (2022) 13:697. doi: 10.3390/GENES13040697
48. Berek JS, Renz M, Kehoe S, Kumar L, Friedlander M. Cancer of the ovary, fallopian tube, and peritoneum: 2021 update. *Int J Gynecology Obstetrics* (2021) 155:61–85. doi: 10.1002/IJGO.13878
49. Köbel M, Kang EY. The evolution of ovarian carcinoma subclassification. *Cancers* (2022) 14:416. doi: 10.3390/CANCERS14020416
50. Oros KK, Ghadirian P, Maugard CM, Perret C, Paredes Y, Mes-Masson AM, et al. Application of BRCA1 and BRCA2 mutation carrier prediction models in breast and/or ovarian cancer families of French Canadian descent. *Clin Genet* (2006) 70(4):320–9. doi: 10.1111/j.1399-0004.2006.00673.x
51. Oros KK, Leblanc G, Arcand SL, Shen Z, Perret C, Mes-Masson AM, et al. Haplotype analysis suggest common founders in carriers of the recurrent BRCA2 mutation 3398delAAAAG, in French Canadian hereditary breast and/ovarian cancer families. *BMC Med Genet* (2006) 7:23. doi: 10.1186/1471-2350-7-23
52. Arcand SL, Maugard CM, Ghadirian P, Robidoux A, Perret C, Zhang P, et al. Germline TP53 mutations in BRCA1 and BRCA2 mutation-negative French Canadian breast cancer families. *Breast Cancer Res Treat* (2008) 108(3):399–408. doi: 10.1007/s10549-007-9608-6
53. Arcand SL, Akbari MR, Mes-Masson AM, Provencher D, Foulkes WD, Narod SA, et al. Germline TP53 mutational spectrum in French Canadians with breast cancer. *BMC Med Genet* (2015) 16:24. doi: 10.1186/s12881-015-0169-y
54. Osher DJ, De Leeneer K, Michils G, Hamel N, Tomiak E, Poppe B, et al. Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. *Br J Cancer* (2012) 106(8):1460–3. doi: 10.1038/bjc.2012.87
55. Ancot F, Arcand SL, Mes-Masson AM, Provencher DM, Tonin PN. Double PALB2 and BRCA1/BRCA2 mutation carriers are rare in breast cancer and breast-ovarian cancer syndrome families from the French Canadian founder population. *Oncol Letters*. (2015) 9(6):2787–90. doi: 10.3892/ol.2015.3123
56. Belanger MH, Dolman L, Arcand SL, Shen Z, Chong G, Mes-Masson AM, et al. A targeted analysis identifies a high frequency of BRCA1 and BRCA2 mutation carriers in women with ovarian cancer from a founder population. *J Ovarian Res* (2015) 8:1. doi: 10.1186/s13048-015-0124-8
57. Guillemette L, Allard C, Lacroix M, Patenaude J, Battista MC, Doyon M, et al. Genetics of glucose regulation in gestation and growth (Gen3G): A prospective prebirth cohort of mother-child pairs in sherbrooke, Canada. *BMJ Open* (2016) 6(2):e010031. doi: 10.1136/bmjopen-2015-010031
58. Gan-Or Z, Zhou S, Ambalavanan A, Leblond CS, Xie P, Johnson A, et al. Analysis of functional GLO1 variants in the BTBD9 locus and restless legs syndrome. *Sleep Med* (2015) 16:1151–5. doi: 10.1016/j.SLEEP.2015.06.002
59. Awadalla P, Boileau C, Payette Y, Idaghdour Y, Goulet JP, Knoppers B, et al. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *Int J Epidemiol* (2013) 42(5):1285–99. doi: 10.1093/ije/dys160
60. Huang K, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic germline variants in 10,389 adult cancers. *Cell* (2018) 173(2):355–370.e14. doi: 10.1016/j.cell.2018.03.039
61. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nat* 2020 581:7809 (2020) 581:434–43. doi: 10.1038/s41586-020-2308-7
62. Koch L. Exploring human genomic diversity with gnomAD. *Nat Rev Genet* (2020) 21(8):448. doi: 10.1038/s41576-020-0255-7
63. Pedersen BS, Brown JM, Dashnow H, Wallace AD, Velinder M, Tristani-Firouzi M, et al. Effective variant filtering and expected candidate variant yield in studies of rare human disease. *NPJ Genomic Med* 2021 6:1 (2021) 6:1–8. doi: 10.1038/s41525-021-00227-3
64. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. *Briefings Bioinf* (2013) 14(2):178–92. doi: 10.1093/bib/bbs017
65. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* (2010) 6(12):e1001025. doi: 10.1371/journal.pcbi.1001025
66. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res* (2010) 20:110–21. doi: 10.1101/GR.097857.109
67. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* (2005) 15(8):1034–50. doi: 10.1101/gr.3715005
68. Shamsani J, Kazakoff SH, Armean IM, McLaren W, Parsons MT, Thompson BA, et al. A plugin for the ensemble variant effect predictor that uses MaxEntScan to predict variant spliceogenicity. *Bioinformatics* (2019) 35(13):2315–7. doi: 10.1093/bioinformatics/bty960
69. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutation*. (2016) 37(3):235–41. doi: 10.1002/humu.22932
70. Ha C, Kim JW, Jang JH. Performance evaluation of SpliceAI for the prediction of splicing of NF1 variants. *Genes* (2021) 12:1308. doi: 10.3390/GENES12091308
71. Ghosh R, Oak N, Plon S. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. *Genome Biol* (2017) 12(9):1308. doi: 10.1186/s13059-017-1353-5
72. Niroula A, Vihinen M. How good are pathogenicity predictors in detecting benign variants? *PLoS Comput Biol* (2019) 15(2):e1006481. doi: 10.1371/JOURNAL.PCBL1006481
73. Cubuk C, Garrett A, Choi S, King L, Loveday C, Torr B, et al. Clinical likelihood ratios and balanced accuracy for 44 in silico tools against multiple large-scale functional assays of cancer susceptibility genes. *Genet medicine : Off J Am Coll Med Genet* (2021) 23:2096–104. doi: 10.1038/s41436-021-01265-Z
74. Gunning AC, Fryer V, Fasham J, Crosby AH, Ellard S, Baple EL, et al. Assessing performance of pathogenicity predictors using clinically relevant variant datasets. *J Med Genet* (2021) 58:547–55. doi: 10.1136/JMEDGENET-2020-107003
75. Rentsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* (2019) 47:D886–94. doi: 10.1093/nar/gky1016
76. Ionita-Laza I, McCallum K, Xu B, Buxbaum JD. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nat Genet* (2016) 48:214–20. doi: 10.1038/NG.3477
77. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* (2015) 24(8):2125–37. doi: 10.1093/hmg/ddu733
78. Li C, Zhi D, Wang K, Liu X. MetaRNN: Differentiating rare pathogenic and rare benign missense SNVs and InDels using deep learning. *bioRxiv* (2021) 2021.4.09.438706. doi: 10.1101/2021.04.09.438706
79. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* (2016) 99(4):877–85. doi: 10.1016/j.ajhg.2016.08.016
80. Douville C, Masica DL, Stenson PD, Cooper DN, Gyax DM, Kim R, et al. Assessing the pathogenicity of insertion and deletion variants with the variant effect scoring tool (VEST-indel). *Hum Mutat* (2016) 37:28–35. doi: 10.1002/HUMU.22911
81. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* (2015) 31(16):2745–7. doi: 10.1093/bioinformatics/btv195
82. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* (2016) 44(D1):D862–8. doi: 10.1093/nar/gkv1222
83. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* (2018) 46:1062–7. doi: 10.1093/nar/gkx1153
84. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
85. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Sherlock: A comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med* (2017) 19(10):1105–17. doi: 10.1038/gim.2017.37
86. Alenezi WM, Fierheller CT, Recio N, Tonin PN. Literature review of BARD1 as a cancer predisposing gene with a focus on breast and ovarian cancers. *Genes* (2020) 11:1–24. doi: 10.3390/GENES11080856
87. Fedick A, Su J, Jalas C, Northrop L, Devkota B, Ekstein J, et al. High-throughput carrier screening using TaqMan allelic discrimination. *PLoS One* (2013) 8(3):e59722. doi: 10.1371/journal.pone.0059722
88. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the sequenom massARRAY iPLEX platform. *Curr Protoc Hum Genet* (2009) Chapter 2:Unit 2.12. doi: 10.1002/0471142905.hg0212s60
89. De Wever J, Everaert H, Coppeters F, Rottiers H, Dewettinck K, Lefever S, et al. The development of a novel SNP genotyping assay to differentiate cacao clones. *Sci Rep* (2019) 9(1):9512. doi: 10.1038/s41598-019-45884-8
90. Behl S, Hamel N, de Ladurantaye M, Lepage S, Lapointe R, Mes-Masson A-M, et al. Founder BRCA1/BRCA2/PALB2 pathogenic variants in French-Canadian breast cancer cases and controls. *Sci Rep* (2020) 10(1):6491. doi: 10.1038/s41598-020-63100-w
91. Sousa SB, Abdul-Rahman OA, Bottani A, Cormier-Daire V, Fryer A, Gillissen-Kaesbach G, et al. Nicolaides-baraitser syndrome: Delineation of the phenotype. *Am J Med Genet* (2009) Part A 149A:1628–40. doi: 10.1002/AJMG.A.32956
92. Sousa SB, Hennekam RC, Abdul-Rahman O, Alders M, Azzarello-Burri S, Bottani A, et al. Phenotype and genotype in nicolaides-baraitser syndrome. *Am J Med Genetics Part C: Semin Med Genet* (2014) 166:302–14. doi: 10.1002/ajmg.c.31409
93. Gana S, Panizzon M, Fongaro D, Selicorni A, Memo L, Scandurra V, et al. Nicolaides-baraitser syndrome: Two new cases with autism spectrum disorder. *Clin Dysmorphology* (2011) 20:38–41. doi: 10.1097/MCD.0B013E31833EDAA9

94. Ejaz R, Babul-Hirji R, Chitayat D. The evolving features of nicolaides–baraitser syndrome – a clinical report of a 20-year follow-up. *Clin Case Rep* (2016) 4:351. doi: 10.1002/CCR3.425
95. Cappuccio G, Sayou C, Tanno P, Tisserant E, Bruel AL, Kennani S, et al. *De novo* SMARCA2 variants clustered outside the helicase domain cause a new recognizable syndrome with intellectual disability and blepharophimosis distinct from nicolaides–baraitser syndrome. *Genet Med* (2020) 22:1838–50. doi: 10.1038/s41436-020-0898-y
96. Cenik BK, Shilatifard A. COMPASS and SWI/SNF complexes in development and disease. *Nat Rev Genet* 2020 22:1 (2020) 22:38–58. doi: 10.1038/s41576-020-0278-0
97. Zhang X, Chen H, Song Y, Chen Z, Liu X, Rong P, et al. Ten-year follow-up of nicolaides–baraitser syndrome with a *de novo* mutation and analysis of 58 gene loci of SMARCA2-associated NCBRS. *Mol Genet Genomic Med* (2022) 10(9):e2009. doi: 10.1002/MGG3.2009
98. Koemans TS, Kleefstra T, Chubak MC, Stone MH, Reijnders MRF, de Munnik S, et al. Functional convergence of histone methyltransferases EHMT1 and KMT2C involved in intellectual disability and autism spectrum disorder. *PLoS Genet* (2017) 13(10):e1006864. doi: 10.1371/JOURNAL.PGEN.1006864
99. McRae JF, Clayton S, Fitzgerald TW, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of *de novo* mutations in developmental disorders. *Nat* 2017 542:7642 (2017) 542:433–8. doi: 10.1038/nature21062
100. Faundes V, Newman WG, Bernardini L, Canham N, Clayton-Smith J, Dallapiccola B, et al. Histone lysine methylases and demethylases in the landscape of human developmental disorders. *Am J Hum Genet* (2018) 102:175–87. doi: 10.1016/j.ajhg.2017.11.013
101. Lavery WJ, Barski A, Wiley S, Schorry EK, Lindsley AW. KMT2C/D COMPASS complex-associated diseases [KCDCOM-ADs]: An emerging class of congenital regulopathies. *Clin Epigenet* 2020 12:1 (2020) 12:1–20. doi: 10.1186/S13148-019-0802-2
102. Thompson ER, Doyle MA, Ryland GL, Rowley SM, Choong DYH, Tothill RW, et al. Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS Genet* (2012) 8(9):e1002894. doi: 10.1371/journal.pgen.1002894
103. Morales J, Pujar S, Loveland JE, Astashyn A, Bennett R, Berry A, et al. A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. *Nat* 2022 (2022) 604(7905):310–315. doi: 10.1038/s41586-022-04558-8
104. Richards CS, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Response to biesecker and Harrison. *Genet Med* (2018) 20:1689–90. doi: 10.1038/gim.2018.43
105. Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat* (2018) 39:1517–24. doi: 10.1002/humu.23626
106. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res* (2014) 42(22):13534–44. doi: 10.1093/nar/gku1206
107. Pirim D, Kaya N, Yildirim EU, Sag SO, Temel SG. Characterization and in silico analyses of the BRCA1/2 variants identified in individuals with personal and/or family history of BRCA-related cancers. *Int J Biol Macromolecules* (2020) 162:1166–77. doi: 10.1016/j.ijbiomac.2020.06.222
108. Trego KS, Groesser T, Davalos AR, Parplys AC, Zhao W, Nelson MR, et al. Non-catalytic roles for XPG with BRCA1 and BRCA2 in homologous recombination and genome stability. *Mol Cell* (2016) 61:535. doi: 10.1016/j.molcel.2015.12.026
109. Galick HA, Marsden CG, Kathe S, Dragon JA, Volk L, Nemecek AA, et al. The NEIL1 G83D germline DNA glycosylase variant induces genomic instability and cellular transformation. *Oncotarget* (2017) 8:85883–95. doi: 10.18632/oncotarget.20716
110. Li N, Lim BWX, Thompson ER, McNerny S, Zethoven M, Cheasley D, et al. Investigation of monogenic causes of familial breast cancer: Data from the BEACCON case-control study. *NPJ Breast Cancer* (2021) 7(1):76. doi: 10.1038/s41523-021-00279-9
111. Li N, Zethoven M, McNerny S, Devereux L, Huang YK, Thio N, et al. Evaluation of the association of heterozygous germline variants in NTHL1 with breast cancer predisposition: An international multi-center study of 47,180 subjects. *NPJ Breast Cancer* (2021) 7(1):52. doi: 10.1038/s41523-021-00255-3
112. Brenner D, Poirier A, Smith L, Aziz LS, Ellison L, Fitzgerald N, et al. Members of the Canadian cancer statistics advisory committee analytic leads additional analysis project management. (2021).
113. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin* (2021) 71:209–49. doi: 10.3322/CAAC.21660
114. Foulkes WD. Inherited susceptibility to common cancers. *New Engl J Med* (2008) 359:2143–53. doi: 10.1056/nejmra0802968
115. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. *Cell* (2019) 176:535–548.e24. doi: 10.1016/j.cell.2018.12.015
116. Metcalfe KA, Poll A, Royer R, Llacuachqui M, Tulman A, Sun P, et al. Screening for founder mutations in BRCA1 and BRCA2 in unselected Jewish women. *J Clin Oncol* (2010) 28:387–91. doi: 10.1200/JCO.2009.25.0712
117. Klungland A, Höss M, Gunz D, Constantinou A, Clarkson SG, Doetsch PW, et al. Base excision repair of oxidative DNA damage activated by XPG protein. *Mol Cell* (1999) 3:33–42. doi: 10.1016/S1097-2765(00)80172-0
118. Sertic S, Quadri R, Lazzaro F, Muzi-Falconi M. EXO1: A tightly regulated nuclease. *DNA Repair* (2020) 93:102929. doi: 10.1016/j.dnarep.2020.102929
119. Vali-Pour M, Lehner B, Supek F. The impact of rare germline variants on human somatic mutation processes. *Nat Commun* (2022) 13:1–21. doi: 10.1038/s41467-022-31483-1
120. Nalepa G, Clapp DW. Fanconi anaemia and cancer: An intricate relationship. *Nat Rev Cancer* (2018) 18:168–85. doi: 10.1038/NRC.2017.116
121. David SS, O'Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. *Nat* 2007 447:7147 (2007) 447:941–50. doi: 10.1038/nature05978
122. Weren RDA, Ligtenberg MJL, Kets CM, De Voer RM, Verwiel ETP, Spruijt L, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat Genet* (2015) 47(6):668–71. doi: 10.1038/ng.3287
123. Grolleman JE, de Voer RM, Elsayed FA, Nielsen M, Weren RDA, Palles C, et al. Mutational signature analysis reveals NTHL1 deficiency to cause a multi-tumor phenotype. *Cancer Cell* (2019) 35(2):256–266.e5. doi: 10.1016/j.ccell.2018.12.011
124. Loong L, Garrett A, Allen S, Choi S, Durkie M, Callaway A, et al. Reclassification of clinically-detected sequence variants: Framework for genetic clinicians and clinical scientists by CanVIG-UK (Cancer variant interpretation group UK). *Genet medicine : Off J Am Coll Med Genet* (2022) 24(9):1867–77. doi: 10.1016/j.gim.2022.05.002
125. Tung N, Domchek SM, Stadler Z, Nathanson KL, Couch F, Garber JE, et al. Counselling framework for moderate-penetrance cancer-susceptibility mutations. *Nat Rev Clin Oncol* (2016) 13:581–8. doi: 10.1038/NRCLINONC.2016.90
126. Liu YL, Breen K, Catchings A, Ranganathan M, Latham A, Goldfrank DJ, et al. Risk-reducing bilateral salpingo-oophorectomy for ovarian cancer: A review and clinical guide for hereditary predisposition genes. *JCO Oncol Pract* (2021) 18(3):201–9. doi: 10.1200/OP.21.00382
127. Hsiao YW, Lu TP. Race-specific genetic profiles of homologous recombination deficiency in multiple cancers. *J personalized Med* (2021) 11(12):1287. doi: 10.3390/JPM11121287
128. Bonache S, Esteban I, Moles-Fernández A, Tenés A, Duran-Lozano L, Montalban G, et al. Multigene panel testing beyond BRCA1/2 in breast/ovarian cancer Spanish families and clinical actionability of findings. *J Cancer Res Clin Oncol* (2018) 144:2495–513. doi: 10.1007/S00432-018-2763-9
129. Jarhelle E, Riise Stensland HMF, Hansen GÅM, Skarsfjord S, Jonsrud C, Ingebretsen M, et al. Identifying sequence variants contributing to hereditary breast and ovarian cancer in BRCA1 and BRCA2 negative breast and ovarian cancer patients. *Sci Rep* (2019) 9(1):19986. doi: 10.1038/s41598-019-55515-x
130. Hirasawa A, Imoto I, Naruto T, Akahane T, Yamagami W, Nomura H, et al. Prevalence of pathogenic germline variants detected by multigene sequencing in unselected Japanese patients with ovarian cancer. *Oncotarget* (2017) 8:112258–67. doi: 10.18632/oncotarget.22733
131. Pan ZW, Wang XJ, Chen T, Ding XW, Jiang X, Gao Y, et al. Deleterious mutations in DNA repair gene FANCC exist in BRCA1/2-negative Chinese familial breast and/or ovarian cancer patients. *Front Oncol* (2019) 9:169. doi: 10.3389/FONC.2019.00169
132. del Valle J, Rofes P, Moreno-Cabrera JM, López-Dóriga A, Belhadj S, Vargas-Parra G, et al. Exploring the role of mutations in fanconi anemia genes in hereditary cancer patients. *Cancers* (2020) 12(4):829. doi: 10.3390/cancers12040829
133. Westphalen CB, Fine AD, André F, Ganesan S, Heinemann V, Rouleau E, et al. Pan-cancer analysis of homologous recombination repair-associated gene alterations and genome-wide loss-of-Heterozygosity score. *Clin Cancer research : an Off J Am Assoc Cancer Res* (2022) 28:1412–21. doi: 10.1158/1078-0432.CCR.21-2096
134. Doddato G, Valentino F, Giliberti A, Papa FT, Tita R, Bruno LP, et al. Exome sequencing in BRCA1-2 candidate families: The contribution of other cancer susceptibility genes. *Front Oncol* (2021) 11:649435. doi: 10.3389/FONC.2021.649435
135. Salo-Mullen EE, Maio A, Mukherjee S, Bandlamudi C, Shia J, Kemel Y, et al. Prevalence and characterization of biallelic and monoallelic NTHL1 and MSH3 variant carriers from a pan-cancer patient population. *JCO Precis Oncol* (2021) 5:455–65. doi: 10.1200/PO.20.00443
136. Broderick P, Bagratuni T, Vijayakrishnan J, Lubbe S, Chandler I, Houlston RS. Evaluation of NTHL1, NEIL1, NEIL2MPG, TDG, UNG and SMUG1 genes in familial colorectal cancer predisposition. *BMC Cancer* (2006) 6:243. doi: 10.1186/1471-2407-6-243
137. Ciavarella M, Miccoli S, Prossomariti A, Pippucci T, Bonora E, Buscherini F, et al. Somatic APC mosaicism and oligogenic inheritance in genetically unsolved colorectal adenomatous polyposis patients. *Eur J Hum genetics : EJHG* (2018) 26:387–95. doi: 10.1038/S41431-017-0086-Y
138. Das R, Kundu S, Laskar S, Choudhury Y, Ghosh SK. Assessment of DNA repair susceptibility genes identified by whole exome sequencing in head and neck cancer. *DNA Repair* (2018) 66–67:50–63. doi: 10.1016/j.dnarep.2018.04.005
139. Lawrenson K, Iversen ES, Tyrer J, Weber RP, Concannon P, Hazelett DJ, et al. Common variants at the CHEK2 gene locus and risk of epithelial ovarian cancer. *Carcinogenesis* (2015) 36:1341–53. doi: 10.1093/CARCIN/BGV138
140. Kurman RJ, Shih IM. The dualistic model of ovarian carcinogenesis: Revisited, revised, and expanded. *Am J Pathol* (2016) 186:733–47. doi: 10.1016/j.ajpath.2015.11.011
141. Srinivasan P, Bandlamudi C, Jonsson P, Kemel Y, Chavan SS, Richards AL, et al. The context-specific role of germline pathogenicity in tumorigenesis. *Nat Genet* (2021) 53:1577–85. doi: 10.1038/s41588-021-00949-1

142. Damia G, Brogini M. Platinum resistance in ovarian cancer: Role of DNA repair. *Cancers* (2019) 11(1):119. doi: 10.3390/CANCERS11010119
143. Li H, Liu ZY, Wu N, Chen YC, Cheng Q, Wang J. PARP inhibitor resistance: the underlying mechanisms and clinical implications. *Mol Cancer* 2020 19:1 (2020) 19:1–16. doi: 10.1186/S12943-020-01227-0
144. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. *Nat* 2014 508:7497 (2014) 508:469–76. doi: 10.1038/nature13127
145. Ramsey SD, Yoon P, Moonesinghe R, Khoury MJ. Population-based study of the prevalence of family history of cancer: Implications for cancer screening and prevention. *Genet Med* (2006) 8:571–5. doi: 10.1097/01.gim.0000237867.34011.12
146. Zheng G, Yu H, Hemminki A, Försti A, Sundquist K, Hemminki K. Familial associations of female breast cancer with other cancers. *Int J Cancer* (2017) 141:2253–9. doi: 10.1002/ijc.30927
147. Momozawa Y, Sasai R, Usui Y, Shiraishi K, Iwasaki Y, Taniyama Y, et al. Expansion of cancer risk profile for BRCA1 and BRCA2 pathogenic variants. *JAMA Oncol* (2022) 8:871–8. doi: 10.1001/JAMAONCOL.2022.0476