Variants in *ATRIP* are associated with breast cancer susceptibility in the Polish population and UK Biobank

Authors

Cezary Cybulski, Neda Zamani, Wojciech Kluźniak, ..., Jean-Yves Masson, Jan Lubiński, Mohammad R. Akbari

Correspondence

mohammad.akbari@utoronto.ca

We report an association between *ATRIP* mutations and increased risk of breast cancer. Identifying individuals with germline *ATRIP* mutations could lead to diagnosing cancer in earlier stages through more intensive cancer screening and could encourage use of more efficient targeted therapies at the time of cancer diagnosis.

Cybulski et al., 2023, The American Journal of Human Genetics *110*, 648–662
April 6, 2023

https://doi.org/10.1016/j.ajhg.2023.03.002

ARTICLE

Variants in *ATRIP* are associated with breast cancer susceptibility in the Polish population and UK Biobank

Cezary Cybulski,[1] Neda Zamani,[2,3,12] Wojciech Kluźniak,[1,12] Larissa Milano,[4,12] Dominika Wokołorczyk,[1] Klaudia Stempa,[1] Helena Rudnicka,[1] Shiyu Zhang,[2] Maryam Zadeh,[2,3] Tomasz Huzarski,[5] Anna Jakubowska,[1,6] Tadeusz Dębniak,[1] Marcin Lener,[1] Marek Szwiec,[7] Paweł Domagała,[8] Amir Abbas Samani,[9,10] Steven Narod,[2,3,11] Jacek Gronwald,[1] Jean-Yves Masson,[4] Jan Lubiński,[1] and Mohammad R. Akbari[2,3,11,*]

Summary

Several breast cancer susceptibility genes have been discovered, but more are likely to exist. To identify additional breast cancer susceptibility genes, we used the founder population of Poland and performed whole-exome sequencing on 510 women with familial breast cancer and 308 control subjects. We identified a rare mutation in *ATRIP* (GenBank: NM_130384.3: c.1152_1155del [p.Gly385Ter]) in two women with breast cancer. At the validation phase, we found this variant in 42/16,085 unselected Polish breast cancer-affected individuals and in 11/9,285 control subjects (OR = 2.14, 95% CI = 1.13–4.28, p = 0.02). By analyzing the sequence data of the UK Biobank study participants (450,000 individuals), we identified *ATRIP* loss-of-function variants among 13/15,643 breast cancer-affected individuals versus 40/157,943 control subjects (OR = 3.28, 95% CI = 1.76–6.14, p < 0.001). Immunohistochemistry and functional studies showed the *ATRIP* c.1152_1155del variant allele is weakly expressed compared to the wild-type allele, and truncated ATRIP fails to perform its normal function to prevent replicative stress. We showed that tumors of women with breast cancer who have a germline *ATRIP* mutation have loss of heterozygosity at the site of *ATRIP* mutation and genomic homologous recombination deficiency. ATRIP is a critical partner of ATR that binds to RPA coating single-stranded DNA at sites of stalled DNA replication forks. Proper activation of ATR-ATRIP elicits a DNA damage checkpoint crucial in regulating cellular responses to DNA replication stress. Based on our observations, we conclude *ATRIP* is a breast cancer susceptibility gene candidate linking DNA replication stress to breast cancer.

Introduction

Breast cancer is the most common malignancy and the first leading cause of cancer-related mortality in women globally.[1] The etiology of breast cancer involves a complex interplay of various risk factors, among which genetic factors have a significant role in susceptibility to the disease. Hereditary breast cancer accounts for approximately 10% of all breast cancer cases.[2] *BRCA1* and *BRCA2* are the two primary breast cancer susceptibility genes based on their mutation frequencies and effect sizes. Deleterious mutations in these two genes are responsible for a significant number of hereditary breast cancer cases.[3,4] The lifetime risk of developing breast cancer is approximately 70% for individuals with a *BRCA1* or *BRCA2* mutation.[5,6] Individuals harboring pathogenic mutations in these genes are also at an increased risk for developing other malignancies such as ovarian, melanoma, prostate, and pancreatic cancers.[7,8]

In addition to *BRCA1/2*, a few other genes, including *ATM*, *BARD1*, *CHEK2*, *PALB2*, *RAD51C*, *RAD51D*, and *TP53*, are regarded as confirmed moderate to highly penetrant breast cancer susceptibility genes.[9,10] Like *BRCA1/2*, pathogenic mutations in these breast cancer susceptibility genes have also been linked to other cancer types such as ovarian, prostate, pancreatic, gastric, colon, and melanoma malignancies.[11,12] Together, deleterious mutations in known breast cancer susceptibility genes account for roughly half of the hereditary breast cancer cases,[13] and the genetic etiology for the remaining women with high familial clustering of breast cancer remains a major knowledge gap.

To address this gap, we applied whole-exome sequencing (WES) on the germline DNA of a cohort of women with familial breast cancer and unknown genetic etiology and unaffected women from the founder population of Poland. This approach has contributed to identifying

[1]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University in Szczecin, Szczecin, Poland; [2]Women's College Research Institute, Women's College Hospital, University of Toronto, Toronto, ON, Canada; [3]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada; [4]Genome Stability Laboratory, CHU de Québec Research Center, Oncology Axis; Department of Molecular Biology, Medical Biochemistry and Pathology; Laval University Cancer Research Center, Québec City, QC, Canada; [5]Department of Clinical Genetics and Pathology, University of Zielona Góra, Zielona Góra, Poland; [6]Independent Laboratory of Molecular Biology and Genetic Diagnostics, Pomeranian Medical University in Szczecin, Szczecin, Poland; [7]Department of Surgery and Oncology, University of Zielona Góra, Zielona Góra, Poland; [8]Department of Pathology, Pomeranian Medical University in Szczecin, Szczecin, Poland; [9]Department of Laboratory Medicine and Pathology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada; [10]Humber River Hospital, University of Toronto, Toronto, ON, Canada; [11]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[12]These authors contributed equally
*Correspondence: mohammad.akbari@utoronto.ca
https://doi.org/10.1016/j.ajhg.2023.03.002.

disease-associated variants.[14,15] There is less genetic heterogeneity in a founder population than in a large mixed population. Therefore, risk alleles segregating at extremely low frequencies in other populations may exist in higher frequencies within a genetically homogeneous population. The strategy for the discovery phase was to identify candidate genes that may be associated with an increased risk of breast cancer. Then, through a subsequent validation phase, we investigated the role of recurrent loss-of-function (LoF) variants in the identified candidate genes with breast cancer risk among a larger set of Polish women with breast cancer and healthy control subjects. Findings were also explored in predisposition to breast cancer among the UK Biobank study population.

Material and methods

Discovery phase

Study subjects

We recruited 510 probands with a strong family history of breast cancer and 308 unaffected women from the founder population of Poland. Affected individuals were diagnosed with invasive breast cancer between 2000 and 2017 at 16 different centers in Poland. The diagnosis of cancer was confirmed through the review of pathology reports. Women with breast cancer who had purely intraductal or intralobular cancers (ductal carcinoma *in situ* or lobular carcinoma *in situ*) were excluded. These women were selected from a registry of 3,519 familial breast cancer cases housed at the Hereditary Cancer Center (HCC) in Szczecin, Poland. Our team previously studied a cohort of 1,018 women with familial breast cancer from this registry for whom DNA sample was available.[16] Briefly, from 3,519 women with breast cancer in the registry, 1,018 breast cancer probands were selected from 1,018 families. Probands' relatives were not included and only women who had at least two affected women in their first- or second-degree relatives were considered. In the case of having only two affected family members, at least one had bilateral breast cancer or breast cancer below the age of 50. Interview was conducted by expert physicians at the time of enrollment to the familial breast cancer registry housed at HCC, and family history of cancer in all first- and second-degree relatives and their ages at onset were recorded. Among these 1,018 probands, 508 women with breast cancer tested positive for genetic variants in known genes associated with breast cancer risk (*BRCA1*, *BRCA2*, *BARD1*, *ATM*, *CHEK2*, *TP53*, *PALB2*, *PTEN*, *NBN*, *RECQL*, *RAD50*, or *CDH1*). The remaining 510 probands for whom genetic testing failed to detect genetic variants in the above-mentioned genes were recruited for the current study (age range 28–76, mean age 49.1 years). Control subjects were 308 cancer-free 40- to 89-year-old adults (mean age of 59.5 years) selected randomly from a registry of individuals who participated in a population-based study of 1.5 million residents of West Pomerania in Poland, designed to identify familial clustering of cancer.[17] Control subjects were not matched to the affected individuals by age, but both affected individuals and control subjects were self-reported ethnic Poles. Control subjects were selected randomly based on the criteria that they were cancer-free females $\geq$40 years of age and had no family history of cancer in their first-degree relatives. These subjects were interviewed between 2007 and 2012, and their blood samples were taken for DNA analysis. The study protocol was approved by the ethics committees of Pomeranian Medical University in Szczecin, Poland and Women's College Hospital, University of Toronto, Canada. All study subjects provided a signed consent form for participation in the study.

Exome sequencing

All 510 women with familial breast cancer and 308 unaffected control subjects were tested by whole-exome sequencing. Germline DNA was isolated using standard methods from peripheral blood leukocytes. The Agilent SureSelect human exome kit (V6) was used to capture sequence target regions. The kit captures 64 Mbp (2.2%) of the human genome and covers coding exons in CCDS and RefSeq databases, as well as exons annotated by the GENCODE project. This includes ~205,000 exons in ~35,000 genes, including protein-coding and noncoding RNA genes. The captured regions for each sample were barcoded, and every 16 samples were pooled and used for paired-end sequencing for 300 cycles (generating 150 bp reads) on a high-throughput sequencing cartridge of Illumina NextSeq 500. The Sentieon software (Sentieon Inc., www.sentieon.com) that includes an optimized implementation of the Burrows-Wheeler Aligner (BWA, https://bio-bwa.sourceforge.net/) and Genome Analysis Toolkit (GATK, https://gatk.broadinstitute.org) was used for secondary analysis of the sequencing data. The sequence reads were aligned to the human genome's reference build hg19. The mean depth of coverage was approximately 100× (ranging from 52× to 154×). On average, 97.4% (ranging from 91.2% to 99.1%) of the CCDS exons were covered at 20× depth of coverage or higher, which were used for variant calling. Regions with at least 20× depth of coverage were used for calling variants, and a different nucleotide from the reference sequence seen in at least 20% of the reads aligned to a given position was called as a variant. The SNP & Variation Suite (GoldenHelix Inc., https://www.goldenhelix.com) was used for annotating called variants. Annotation was used to determine the related genes and the effect of each variant on the encoded protein. We focused on loss-of-function (LoF) variants (frameshift indels, stop codon gain, stop codon loss, essential splicing site, and start codon loss mutations) with a minor allele frequency (MAF) of $\leq$1% among different groups reported in gnomAD and 1000 Genomes Project databases.

Sanger sequencing

Mutations of interest identified by WES were confirmed by Sanger sequencing before further evaluation. According to the manufacturer's protocol, sequencing reactions were performed using the BigDye Terminator v.3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). Sequencing products were analyzed on the ABI Prism 3500XL Genetic Analyzer (Thermo Fisher Scientific). All sequences were compared to the related gene RefSeq sequence for variant detection using Mutation Surveyor software (SoftGenetics Inc., https://softgenetics.com/products/mutation-surveyor/).

Statistical analysis

Genes with confirmed LoF variants among study subjects were selected for gene-based association analyses. We estimated age-adjusted odds ratio (OR) for each gene harboring LoF mutation to investigate its association with breast cancer risk. To do so, we compared total LoF variant frequencies in each gene among Polish women with familial breast cancer (affected individuals) to those observed in unaffected Polish individuals (control subjects). The age-adjusted OR with 95% CI and corresponding two-tailed p value were assessed using logistic regression analysis.

Validation phase

Study subjects

At the validation phase, study participants comprised 16,085 prospectively ascertained unselected Polish women with breast cancer and 9,285 unaffected Polish adults. These individuals were non-overlapping with the affected individuals and control subjects used in the discovery phase. Affected individuals were women aged 18–94 years (mean age of 56.2 years) diagnosed with invasive breast cancer between 1996 and 2013 at 18 different hospitals in Poland. Women with breast cancer who had purely intraductal or intralobular cancers (ductal carcinoma *in situ* or lobular carcinoma *in situ*) were excluded. All cases were unselected for family history and did not harbor any of the three common founder alleles in *BRCA1* (GenBank: NM_007294.4: c.181T>G [p.Cys61Gly]; GenBank: NM_007294.4: c.4035del [p.Glu1346Lysfs*20]; and GenBank: NM_007294.4: c.5266dup [p.Gln1756Profs*74]), which account for ~86% of all *BRCA1/2* mutations seen in the Polish population.[18] 5,524 women were 50 years old or younger at the time of diagnosis, and 10,561 women were older than 50 years at the time of diagnosis. Family history was collected by expert physicians from ~89% of the affected individuals by constructing a family tree or completing a questionnaire. Approximately 14% of women with family history data reported a first- or second-degree relative with breast cancer. Control subjects were cancer-free women aged 18–94 years (mean age of 51.2 years) who were derived from five sources. The first subgroup consisted of 979 women from the region of Szczecin (age range, 24 to 84 years) who were chosen for this study to be age-matched and geographically matched with a series of enrolled women with breast cancer diagnosed in Szczecin between 1996 and 2004. These women were part of a population-based study of the 1.5 million residents of West Pomerania (North-West Poland) designed to identify familial aggregations of cancer and were interviewed in 2007. The second control series consisted of 1,707 unselected women (age range 32–72) who participated in mammography screening at eight different centers all over Poland between 2009 and 2011 (Kielce, Legnica, Olsztyn, Poznan, Szczecin, Swidnica, Torun, Zielona Góra) and provided blood samples for DNA analysis. Women with breast cancer were excluded from this group. The third control group included 1,031 unselected women (age range, 20 to 94 years) selected at random from the computerized lists of individuals from family practices located in the region of Opole (South Poland). These women were invited to participate by mail and participated in 2012 and 2013. The fourth series included 985 Polish women (age range 50–66 years) who participated in the population colonoscopy screening program for colorectal cancer between 2007 and 2010 in Szczecin, Bialystok, and Lódź and provided blood samples for DNA analysis. The fifth control series consisted of 4,583 women (age range 22–92) who were interviewed in 17 genetic-oncology outpatient clinics all over Poland between 2013 and 2018 (Szczecin, Koszalin, Olsztyn, Białystok, Bydgoszcz, Toruń, Poznań, Konin, Warszawa, Łódź, Zielona Góra, Opole, Legnica, Bielsko-Biała, Świdnica, Kielce, Kraków) provided blood samples for DNA analysis and were selected as control subjects for this study because they were cancer free.

Almost half of the control subjects (48%; 4,430/9,285) were 50 years old or younger, and the other half (52%; 4,855/9,285) were older than 50 years. We have reviewed the demographic characteristics of our affected individuals and control subjects: 97.6% of 16,085 affected individuals and 97.4% of 9,285 of control subjects self-reported as Polish. In addition, all individuals with the *ATRIP* recurrent mutation self-reported as Polish, both among affected individuals and control subjects. Although affected individuals and control subjects were derived from 18 and 17 centers across Poland, respectively, their geographical distributions were similar. The distributions of affected individuals and control subjects were 44% for affected individuals and 62% for control subjects in northern, 16% for affected individuals and 14% for control subjects in central, and 40% for affected individuals and 24% for control subjects in southern Poland. Since Poland is populated by ethnic Slavs and it is a genetically homogeneous country, small differences in geographical distribution of affected individuals and control subjects probably have no effect on the results. Our team has previously shown that the frequency of Polish founder mutations (i.e., in *BRCA1*, *CHEK2*, *NBN*, and *RECQL*) is similar in different regions of Poland.[14] Detailed geographical distribution of affected individuals and control subjects including *ATRIP* c.1152_1155del mutation frequencies are shown in Table S1.

All individuals signed a consent form for participation in the study. The study protocol was approved by the ethics committee of Pomeranian Medical University in Szczecin, Poland.

TaqMan genotyping

For the candidate genes identified at the discovery phase, we excluded variants seen in one individual only and focused on studying the role of recurrent LoF variants in susceptibility to breast cancer. Germline DNA was isolated from 5 to 10 mL of the peripheral blood of study participants at the validation phase. The recurrent LoF variants were then genotyped using TaqMan assay on ABI 7500 Fast real-time system (Thermo Fisher Scientific). Laboratory technicians were blinded to case-control status. The overall genotyping call rate was 99.5%. The presence of mutations was confirmed by Sanger sequencing as described in the "discovery phase" section.

Statistical analysis

Breast cancer risk associated with carrying a potentially pathogenic founder variant was estimated by comparing its mutation frequency among affected individuals versus control subjects. The age-adjusted ORs for developing breast cancer with 95% CIs and corresponding two-tailed p values were assessed using logistic regression analysis.

ExAC exome data

The non-Finnish European population of the ExAC exome sequencing database was used to estimate the *ATRIP* genotype frequencies expected in the general population. The total number of individuals with LoF variants in *ATRIP* at the discovery phase was compared to the total number of individuals with a germline *ATRIP* LoF mutation reported among the 22,928 female population of the ExAC exome database.

Co-segregation analysis

Among all 42 women with breast cancer who had a germline *ATRIP* c.1152_1155del mutation, six women' relatives could be accessed and gave consent to provide their DNA samples for co-segregation analysis. Germline DNA was extracted from the peripheral blood of the family members and was screened for the *ATRIP* c.1152_1155del mutation by Sanger sequencing.

Functional studies of the *ATRIP* variant

Constructs

ATRIP (GenBank: NM_130384.3) Human Tagged ORF Clone (RC223562) was purchased from Origene on a pCMV6-with C-terminal Myc- DDK Tag. The *ATRIP* c.1152_1155del variant was obtained via site-directed mutagenesis using the QuickChange II Site-Directed Mutagenesis Kit (Agilent) on this *ATRIP* vector,

Table 1. *ATRIP* loss-of-function variants identified in Polish women with breast cancer

Mutation (DNA)[a]	Exon	Protein change	Frequency among affected women	Frequency among control subjects	Frequency among ExAC database[b]	OR[c] (95%CI)	p value[d]
Whole-exome sequencing of Polish women with breast cancer and healthy control subjects **(discovery phase)**							
c.69_75dup	1	p.Thr26Alafs*23	1/510	0/308	0/22,928	4.23 (NA)	0.29
c.1152_1155del	8	p.Gly385Ter	2/510	0/308	2/22,928		
Genotyping of recurrent *ATRIP* variants in Polish women with breast cancer and population control subjects							
c.1152_1155del	8	p.Gly385Ter	42/16,085	11/9,285	2/22,928	2.14 (1.13–4.28)	0.02

[a]Nucleotide positions are based on the GenBank: NM_130384.3 transcript of *ATRIP*.
[b]Frequencies are for the cancer-free non-Finnish European females of the ExAC database.
[c]Gene-based odds ratio was calculated for breast cancer risk at the discovery phase. Odds ratio for the discovery phase was calculated using Haldane-Anscombe correction. Odds ratio for the validation phase was adjusted for age using logistic regression analysis.
[d]Two-tailed p values were calculated.

previously modified to delete the C-terminal tag and insert of an N-terminal FLAG using the Q5 Site-Directed Mutagenesis Kit with primers listed on Table S2. To generate stable cells, genomic editing with the AAVS1 targeting system was used.[19] Sites NotI and PspXI were introduced into the pCMV6 FLAG *ATRIP* constructs using primers listed in Table 2. Digestion of the sites NotI and PspXI into the vector and the *ATRIP* insert were used to clone the *ATRIP* WT and *ATRIP* c.1152_1155del mutated version into the pAAVS1.V2 *neo* plasmid.

Cell lines

HEK293FT were purchased from Invitrogen and maintained in DMEM supplemented with 10% FBS. HeLa cells were obtained from ATCC and maintained in DMEM supplemented with 10% FBS and 1% P/S. All cell lines were grown at 37°C, 5% CO_2. RPE $p53^{-/-}$ cells were obtained from Daniel Durocher (SLRI) and maintained in DMEM supplemented with 10% FBS and 1% P/S. All cell lines were grown at 37°C, 5% CO_2, and routinely tested to be mycoplasma free. ATRIP KO cells were generated in the RPE $p53^{-/-}$ background using the Gene Knockout Kit v2 (Synthego), and the used target-specific multi-guide sgRNAs are listed in Table S3. Cas9 from *S. pyogenes* and sgRNA guides were transfected using CRISPRmax following manufacturer instructions. After three days, cells were isolated in single clones in 96-well plates. After expansion, genomic DNA was extracted using the Qiagen QIAMP DNA and Blood mini kit (QIAGEN) and genomic sequences were amplified using primers described in Table S3 and Sanger sequenced, and the sequences were submitted to the Synthego ICE analysis tool to confirm the indels. RPE $p53^{-/-}$ ATRIP KO clone 11B cells were stably complemented using the AAVS1genomic editing system.[19] Briefly, cells were transfected with the 4 μg of the AAVS1 construct containing either the WT *ATRIP* or the *ATRIP* c.1152_1155del, along with the 0.4 μg of the pZFN plasmid, using Lipofectamine 2000 (Invitrogen) for four hours. 24 h later, transfected cells were selected with G418 for seven days. Transient transfections in HeLa, HEK293FT, and RPE $p53^{-/-}$ ATRIP KO cells were performed with 1 μg of the pCMV6 FLAG *ATRIP* constructs using lipofectamine 2000 for 4 h. 24 h later, cells were treated with 25 μM of MG132 for 4 h when indicated. After treatment, cells were harvested for protein extraction and immunoblot.

Sensitivity assay

For the sensitivity assay in RPE $p53^{-/-}$ cells, cells were seeded in triplicates into a Corning 3603 black-sided clear-bottom 96-well microplate at a density of 2,000 cells per well. Hydroxyurea (HU) sensitivity assay was then performed as described in Rodrigue et al.[20] Cells were treated with HU for four days with concentrations ranging from 0 to 500 mM or for three days in combination different concentrations of MG132 ranging from 0 to 0.5 mM. Images of entire wells were acquired at 4× with a Cytation 5 Cell Imaging Multi-Mode Reader followed by quantification of Hoechst-stained nuclei with the Gen5 Data Analysis Software v.3.03 (BioTek Instruments). Cell viability was expressed as a percentage of survival in HU-treated cells relative to vehicle-treated cells. Results represent the mean ± SEM of at least three independent experiments, each performed in triplicate.

Protein extracts and immunoblot

Total soluble protein extracts and immunoblotting were performed as described in Castroviejo-Bermejo et al.[21] A polyclonal antibody (Abcam # ab19351) was used for ATRIP detection at a 1:1,000 dilution. A mouse monoclonal antibody (Sigma Aldrich #F3165) was used at a 1:1,000 dilution for FLAG detection. Anti-vinculin (1:200,000; Sigma Aldrich, #V9131) was the loading control. Horseradish peroxidase-conjugated anti-rabbit IgG or anti-mouse (1:10,000; Jackson Immuno Research) were used as secondary antibodies.

Immunofluorescence

Cells were seeded into Corning 96-Well Half Area High Content Imaging Film Bottom Microplate at 7,000 cells/well. Unless otherwise stated, all immunofluorescence dilutions were prepared in PBS and incubations performed at room temperature with intervening washes in PBS. Cell fixation was carried out by incubation with 4% paraformaldehyde for 10 min, followed by 100% ice-cold methanol for 5 min at −20°C. Cells were permeabilized in 0.2% Triton X-100 for 5 min, followed by a quenching step using 0.1% sodium borohydride for 5 min. After blocking for 1 h in a solution containing 10% goat serum and 1% BSA, cells were incubated for 1 h with primary antibody anti-ATRIP (1:200, Abcam # ab19351) diluted in 1% BSA. Secondary antibody labeling used Alexa Fluor 488 goat anti-rabbit (Invitrogen, # A-11008) diluted at 1:1,000 in 1% BSA for 1 h. Nuclei were stained for 10 min with 1 mg/mL 4,6-diamidino-2-phenylindole (DAPI). z stack images were acquired on a ZEISS Celldiscoverer 7 automated microscope using a 50× water immersion objective and analyzed for ATRIP nuclei intensity with ZEN Blue software 3.2 (ZEISS).

Loss-of-heterozygosity (LOH) analysis at the site of* ATRIP *mutation and homologous recombination deficiency (HRD) score

Ten available tumor samples from women with breast cancer who had the germline *ATRIP* c.1152_1155del variant were subjected to LOH analysis. Unstained formalin-fixed, paraffin-embedded (FFPE) sections were macro-dissected for DNA extraction. Expert

Table 2. Loss-of-heterozygosity analysis and homologous recombination deficiency scores for ten tumor samples of Polish women with breast cancer and the germline *ATRIP* c.1152_1155del

Tumor Sample	Tumor histology[a]	LOH-WGS[b] AD (VAF)	LOH-WES[b] AD (VAF)	LOH score	TAI score	LST score	HRD[c] score
1	ER+/HER2-	10.5 (33.3%)	N/A	3	1	20	24
2	ER+/HER2-	6.6 (50.0%)	N/A	2	5	13	20
3	ER+/HER2-	5.8 (61.5%)	4.18 (81.8%)	15	12	35	62
4	ER+/HER2-	7.8 (53.3%)	N/A	1	2	3	6
5	ER+/HER2-	0.6 (100%)	17.44 (72.1%)	18	19	49	86
6	ER+/HER2-	2.6 (75.0%)	57.70 (55.1%)	11	13	28	52
7	TN	8.14 (63.6%)	17.59 (77.6%)	10	17	53	80
8	ER+/HER2-	9.21 (70.0%)	N/A	11	8	45	64
9	TN	11.8 (42.1%)	N/A	4	13	60	77
10	ER+/HER2+	4.6 (60.0%)	91.98 (51.9%)	6	9	47	62

[a]Histology of the tested tumors based on presence of estrogen receptor (ER) and HER2; TN, triple negative.
[b]Loss of heterozygosity (LOH) at the site of *ATRIP* mutation was detected by whole-genome sequencing (WGS). For five samples for which either the variant allele frequency (VAF) was borderline (around 60%) or the allelic depth of coverage (AD, Reference Allele/Alternative Allele) at the mutation site was too low, whole-exome sequencing (WES) was carried out to confirm LOH status at the site of *ATRIP* mutation.
[c]Homologous recombination deficiency (HRD) score was calculated as the unweighted sum of the number of LOH (larger than 15 Mbp), telomeric allelic imbalance (TAI), and large-scale (at least 10 Mbp) state transition (LST) breakpoints across the tumor whole genome. None of the tumor DNA samples had any mutations in any of the 15 clinically important homologous recombination repair genes (*ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK1*, *CKEK2*, *CDK12*, *FANCL*, *PALB2*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD54L*, or *PPP2R2A*). N/A, not applicable.

pathologists had previously examined the hematoxylin and eosin (H&E) tumor slides to detect tumor/normal areas. Tumor DNA was extracted from tumor cells using the QIAamp DNA Micro Kit (Qiagen) and was subsequently quantified using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific Inc.). LOH analysis was conducted using NGS, and libraries were prepared using the xGen Prism DNA Library Prep Kit (Integrated DNA Technologies) according to the manufacturer's protocol. Every five indexed samples were pooled together and used for paired-end sequencing for 300 cycles (2 × 150) on the Illumina NovaSeq 6000 sequencer for whole-genome sequencing. We looked for another pathogenic mutation in the *ATRIP* in addition to the *ATRIP* c.1152_1155del germline mutation or the deletion of the entire wild-type allele. We considered LOH if the variant allele frequency (VAF) was greater than 60%. The VAF cut-off point of 60% is based on the formula VAF = (%tumor cells)/(%tumor cells +2 × (100 − %tumor cells)), which takes into consideration that the deletion of the wild-type allele in hereditary syndromes is an early phenomenon, and therefore, present in most tumor cells. This represents 75% purity of the tumor cells. For each tumor, the homologous recombination deficiency (HRD) score was calculated as the unweighted sum of three independent DNA-based measures of genomic instability that includes genomic LOH, telomeric allelic imbalance (TAI), and large-scale state transition (LST) scores. HRD-LOH score was defined as the number of LOH regions across the genome which are longer than 15 Mb.[22] HRD-TAI score was defined as the number of regions with an allelic imbalance extending to one of the sub telomeres but not crossing the centromere.[23,24] HRD-LST score was calculated as the number of breakpoints between regions longer than 10 Mb after filtering out regions shorter than 3 Mb.[25] NxClinical 6.2 (BioDiscovery LLC) was used for HRD score analysis. All tumor DNA samples were screened for mutations in 15 homologous recombination repair (HRR) genes (*ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK1*, *CKEK2*, *CDK12*, *FANCL*, *PALB2*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD54L*, and *PPP2R2A*) that could result in high HRD score if they are mutated.[26,27]

Immunohistochemistry (IHC) analysis

FFPE tissue block sections from 12 breast tumor tissues of women with the germline *ATRIP* c.1152_1155del mutation were placed on coated slides, washed in xylene to remove the paraffin, and rehydrated through serial dilutions of alcohol, followed by washings with a solution of PBS (pH 7.2). All subsequent washes were buffered via the same protocol. Treated sections were then placed in a citrate buffer (pH 6.0) and heated in a microwave for two 5-min sessions. The samples were then incubated with a monoclonal anti-human ATRIP polyclonal antibody (ab245632, abcam) for 60 min at 25°C. The tissue sections on slides were examined and scored using the German semi-quantitative scoring system considering the staining intensity and area extent.[28] Every tumor was given a score according to the intensity of the cytoplasmic and nucleic staining collectively (no staining, 0; weak staining, 1; moderate staining, 2; strong staining, 3) and the extent of stained cells (0%, 0; 1%–10%, 1; 11%–50%, 2; 51%–80%, 3; 81%–100%, 4). The IHC stain score was calculated by multiplying the intensity score by the extent of the positivity score (ranging from 0 to 12).

Evaluating the role of ATRIP *in susceptibility to breast cancer in the UK Biobank database*

To study the association of *ATRIP* mutations with breast cancer risk among populations other than the Polish population, we accessed individual-level exome-sequence and phenotypic data of the UK Biobank (UKB) study participants (application number: 85529). The UKB is a prospective population-based study involving more than 500,000 participants (229,134 men and 273,402 women) aged 40 to 69 years old at the time of recruitment. Participants were recruited between 2006 and 2010 from 22 assessment centers across England, Scotland, and Wales.[29] Breast cancer is the most prevalent malignancy and the most common incident cancer diagnosed in women within the UKB study. We included all White British women with invasive breast cancer for whom WES data were available and did not have *BRCA1/2* and *PALB2* mutations. In addition, all ethnicity-matched cancer-free women with available WES data that did not have a *BRCA1/2* and *PALB2* mutations were considered as the control group for association analysis.

Individuals who carried LoF variants in *ATRIP* were determined. Using exome sequence variant data, we estimated pairwise identical by descent (IBD) allele proportions for all individuals with a germline *ATRIP* mutation among affected women and control subjects. One sample from each pair of individuals whose proportion of their IBD alleles (PI_HAT value) was more than 0.125 (representing third-degree relatedness) were excluded from the dataset. PI_HAT value ranges between 0 and 1. It is the proportion of the genome of two individuals whose both alleles are identical by descent plus half of the proportion of their genome in which just one allele is identical by descent. This value for an identical twin pair is 1.0 and for a parent-offspring pair is 0.5. The lower the value, the more distantly related the individuals are.[30] To investigate the association of *ATRIP* with breast cancer risk, we performed a gene-based association analysis and compared total LoF mutations in *ATRIP* (MAF $\leq$ 1%) among affected women to those observed in control subjects. The age-adjusted OR with 95% CI and corresponding two-tailed p value were assessed using logistic regression analysis. All analyses were performed by SNP & Variation Suite V8 (Golden Helix Inc.). All the LoF variants identified in *ATRIP* among all the UK Biobank's affected women and control subjects were assessed using American College of Medical Genetics and Genomics (ACMG) guideline for classifying the variants.[31] Only variants with pathogenic or likely pathogenic classification were used for the association analysis.

Results

We performed whole-exome sequencing on the germline DNA of 510 Polish women with familial breast cancer and 308 unaffected individuals at the discovery phase. Affected individuals were selected based on having a strong family history of breast cancer and a young age of disease onset. All affected women did not have any mutations in known breast cancer susceptibility genes (*ATM, BRCA1*, *BRCA2*, *BARD1*, *CHEK2*, *TP53*, *PALB2*, *PTEN*, *NBN*, *RECQL*, *RAD50*, or *CDH1*).[16]

To identify additional breast cancer susceptibility genes, we identified genes in which a LoF mutation (stop codon gain, stop codon loss, essential splice-site, frameshift insertion/deletion, or start codon loss mutations) was observed among the studied subjects. We excluded LoF variants with a minor allele frequency (MAF) greater than 1% in the general population (either in in-house or publicly available mutation databases), assuming these to be benign or low-penetrance variants. Then we performed gene-based association analysis between affected women and control subjects for the genes with LoF variants. Among genes with LoF variants associated with a higher risk of breast cancer (odds ratio [OR] over 4 with any p value), we focused on the genes with known functions related to cancer pathogenesis and those with a recurrent LoF variant. We came up with a short list of nine breast cancer susceptibility genes for further evaluation. On top of the list, we had *ATRIP* with three individuals having LoF variants among 510 affected women and no individual with a mutation among 308 control subjects (OR = 4.23, p = 0.29) (Table 1). One affected woman had *ATRIP* GenBank: NM_130384.3: c.69_75dup (p.Thr26Alafs*23) frameshift insertion and two others had *ATRIP* GenBank: NM_130384.3: c.1152_1155del (p.Gly385Ter) frameshift deletion in this gene. The frequency of *ATRIP* LoF variants among 22,928 cancer-free non-Finnish European females of the ExAC database[32] is 1 in every 2,293 women.

Considering the observation of three LoF variants in *ATRIP* among 510 women with breast cancer and a very low frequency of *ATRIP* LoF variants in the general population, we decided to explore the frequency of the recurrent *ATRIP* c.1152_1155del variant in a large set of 16,085 Polish women with unselected breast cancer and 9,285 unaffected Polish women in the validation phase. The *ATRIP* c.1152_1155del variant was detected in 42 of 16,085 affected women (1 in every 383 women with breast cancer, 0.3%) compared to 11 of 9,285 control subjects (1 in every 844 individuals, 0.1%) (OR = 2.14, 95% CI = 1.13–4.28, p = 0.02) (Table 1).

The mean age at the time of breast cancer diagnosis was 55.5 years among all 42 Polish women who had the *ATRIP* c.1152_1155del mutation in our study, which was not different from the mean age at diagnosis among all women with breast cancer and no *ATRIP* mutation (56.2 years; p = 0.7). In addition, seven of 35 (20.0%) affected women with the *ATRIP* mutation for whom family history information was available reported a positive family history of breast cancer among their first- or second-degree relatives, in comparison to 1,964 of 14,320 (13.7%) women with breast cancer and no germline *ATRIP* mutation and available family history information (p = 0.28). Clinical characteristics of women with the ATRIP c.1152_1155del mutation and those without the mutation were compared in Table S4. Interestingly, ductal, grade 3 tumors were more common in women with a germline *ATRIP* mutation than in those with no mutation (36.8% vs. 19.9%; p = 0.02), and progesterone-receptor-positive cancers were less common in women with a germline *ATRIP* mutation compared to those without a mutation (52.6% vs. 71.2%; p = 0.02).

The pedigrees of six Polish women with familial breast cancer harboring the *ATRIP* c.1152_1155del mutation are shown in Figure 1. The relatives of these women with breast cancer for whom a DNA sample was available were tested by Sanger sequencing for the *ATRIP* mutation. We observed co-segregation of the *ATRIP* c.1152_1155del mutation with breast cancer in five families (Figure 1). Of the seven relatives (including four first-degree) with breast cancer who have been tested, six had the *ATRIP* family mutation in addition to the six probands. Only one second-degree relative did not have the mutation (individual II-4, family 4 in Figure 1).

The *ATRIP* 1152_1155del variant causes a frameshift mutation introducing a premature stop codon (p.Gly385Ter) which produces a truncated version of the protein. The estimated size of the truncated protein variant is 42 kDa, while the wild-type (WT) version of this protein has 85 kDa. First, to observe the effects of this variant in a cell system, different human cell lines were transiently

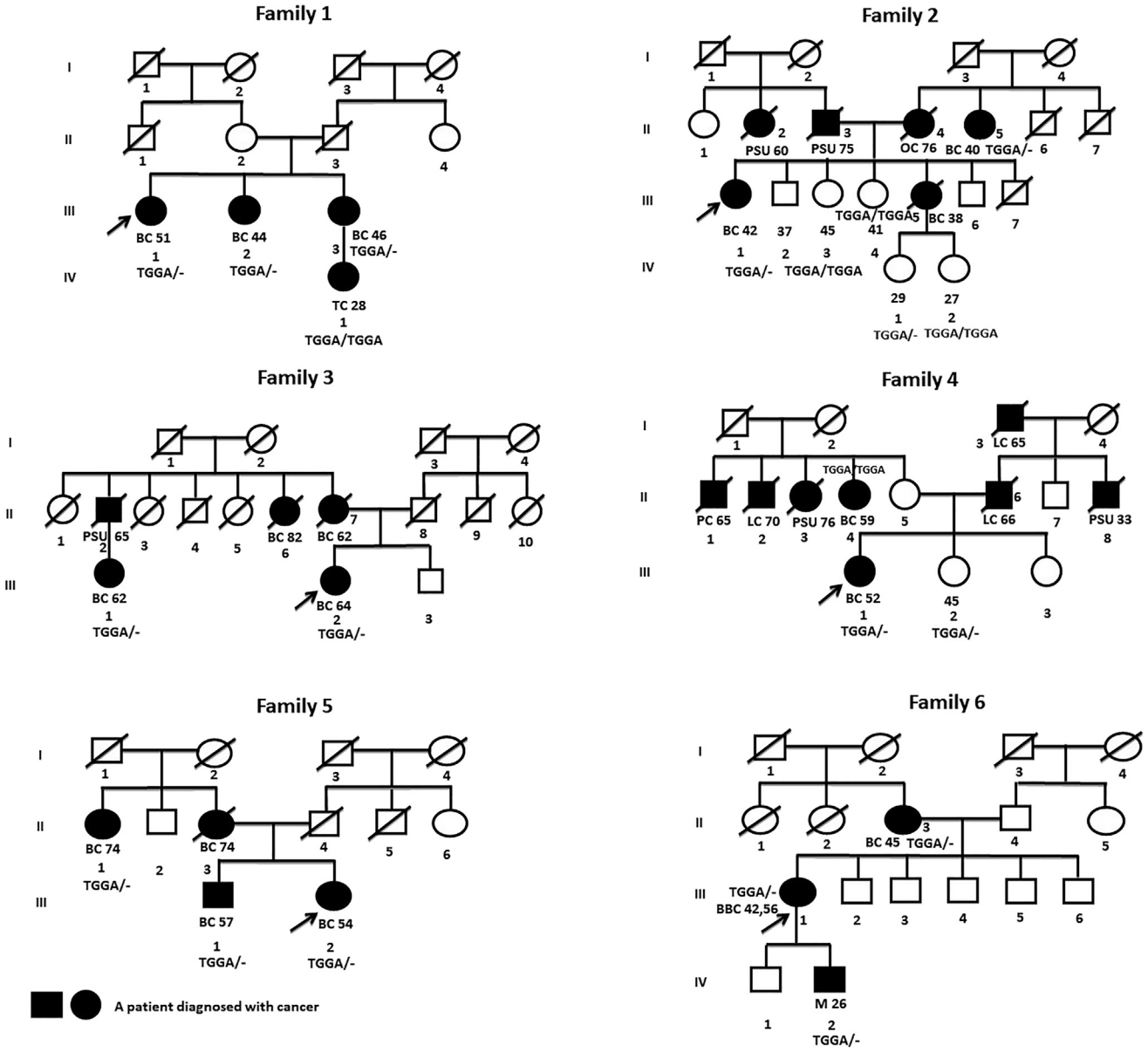

Figure 1. Pedigrees for six Polish women with familial breast cancer harboring the *ATRIP* c.1152_1155del mutation
The type of cancer and age at diagnosis are indicated for individuals with cancer diagnosis. A (TGGA/−) by the symbol indicates the presence of the mutation and a (TGGA/TGGA) symbol indicates the absence of the mutation. BC, breast cancer; OC, ovarian cancer; LC, lung cancer; PC, pancreatic cancer; M, medulloblastoma; TC, thyroid cancer; PSU, primary cancer site unknown.

transfected with the WT or 1152_1155del FLAG N-terminal *ATRIP* construct and *ATRIP* (Figures 2A–2C) or the FLAG epitope (Figures S1A–S1C) were detected by western blotting. Protein accumulation is depicted in HeLa cells (Figures 2A and S1A), in RPE p53$^{-/-}$ ATRIP KO cells (Figures 2B and S1B), and in HEK293FT cells (Figures 2C and S1C). While we could detect the accumulation of WT ATRIP, the ATRIP p.Gly385Ter variant showed reduced levels in Hela and RPE cells. Considering the premature codon introduced, we wondered whether this variant could be degraded through the proteasomal system. Hence, cells were treated with the proteasomal inhibitor MG132, and protein levels were detected. After 4 h of exposure to MG132 at 25 μM, accumulation of the truncated version of the protein was restored and detected with anti-ATRIP and anti-FLAG antibodies (Figures 2 and S1).

Next, to explore the functionality of the ATRIP p.Gly385Ter variant, we generated ATRIP CRISPR-Cas9 knockout cells in the RPE p53$^{-/-}$ background and stably complemented these cells with either the WT or the 1152_1155del constructs, using the AAVS1 genomic editing system (Figures 3A and 3B). ATRIP accumulation was detected in ATRIP KO complemented cells by immunoblot (Figure 3B) or immunofluorescence (Figures 3C and 3D). As previously observed in the transient system, the ATRIP p.Gly385Ter variant weakly accumulated compared to the WT, and protein levels were restored after MG132

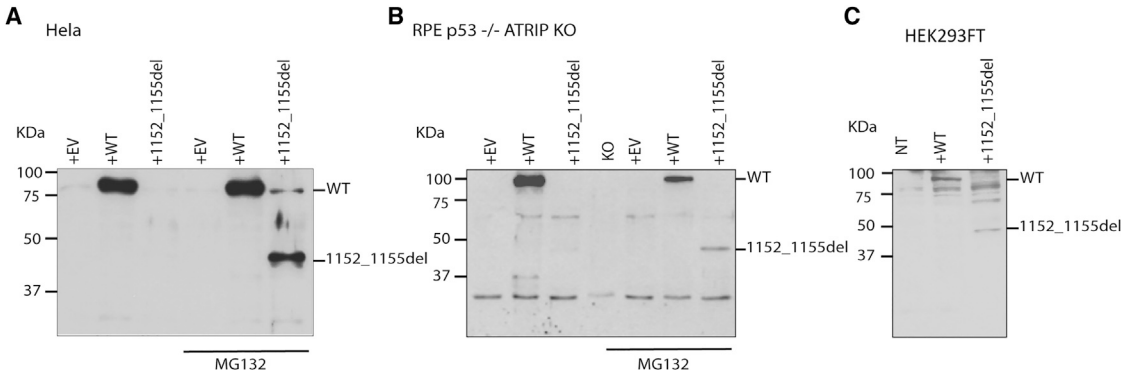

Figure 2. ATRIP levels produced from constructs with *ATRIP* c.1152_1155del in different cell lines
Human cell lines—Hela (A), RPE p53$^{-/-}$ ATRIP KO (B), and HEK293FT (C)—were transiently transfected with an empty vector (+EV) or either with the WT FLAG-*ATRIP* (+WT) or the variant c.1152_1155del FLAG-*ATRIP* (+1152_1155del). 24 h later, cells were exposed to 25 μM of the proteasomal inhibitor MG132 or DMSO (vehicle) for 4 h. ATRIP levels were assessed at the end of MG 132 exposure.

treatment (Figure 3B). ATRIP assists ATR recruitment to stressed replication forks[33]; we, therefore, evaluated the sensitivity of ATRIP-deficient and complemented cells exposed to hydroxyurea (HU), a replication-blocking agent that acts to reduce dNTP levels by inhibiting the ribonucleotide reductase enzyme.[34] RPE p53$^{-/-}$ ATRIP-deficient cells were more vulnerable to HU and complementation with *ATRIP* WT restored viability to the baseline levels (Figure 3E). As expected, the *ATRIP* c.1152_1155del variant failed to rescue the viability of the ATRIP-deficient cells, showing HU sensitivity similar to cells complemented with the empty vector (EV) construct. To further investigate whether the recovery in the accumulation of the ATRIP p.Gly385Ter variant after the MG132 treatment could have an impact on the sensitivity of cells harboring this variant to HU, we exposed the cells to a combined treatment of a non-cytotoxic dose of MG132 (Figure S2) with different concentrations of HU (Figures 3F and S2). On the one hand, the combined treatment with MG132 led to a slight increase in resistance to 500 μM of HU in RPE p53$^{-/-}$ and RPE p53$^{-/-}$ ATRIP KO+WT cells, although this result was not statistically significant. On the other hand, sensitivity to HU alone or combined with MG132 was significantly different in RPE p53$^{-/-}$ ATRIP KO+EV and RPE p53$^{-/-}$ ATRIP KO+p.Gly385Ter cells when compared to RPE p53$^{-/-}$ cells, but MG132 had no impact in the survival of cells harboring the ATRIP p.Gly385Ter variant. These results demonstrate that the increased accumulation of the truncated version of ATRIP after MG132 exposure is insufficient to re-establish protein function and prevent replicative stress. This lack of function is possibly due to the loss of interaction with ATR, as ATRIP p.Gly385Ter does not bear its C terminus required for interaction.[35]

After showing the higher frequency of *ATRIP* c.1152_1155del variant among breast cancer cases and its co-segregation with breast cancer in families of a few probands and also confirming the deleterious effect of this variant on gene function, we performed whole-genome sequencing of the breast tumor DNA (mean depth of coverage of 40×) for ten women with breast cancer and available samples who had the *ATRIP* c.1152_1155del variant in their germline DNA. Sequence data showed evidence of loss of heterozygosity (LOH) for the *ATRIP* c.1152_1155del variant in six samples (Table 2). However, for five samples, either the variant allele frequency (VAF) was borderline (around 60%) or the depth of coverage at the mutation site was too low. To further evaluate LOH in these five samples, we performed whole-exome sequencing on the tumor DNA (mean depth of coverage of 200×), and LOH was confirmed in three of the five samples (Table 2). All four samples with confirmed LOH through whole-genome and whole-exome sequencing had a homologous recombination deficiency (HRD) score greater than 42 indicating the presence of genomic instability. Three of the other six samples with no LOH also showed high HRD scores, two of the three were the samples with borderline LOH in WGS, which was not confirmed in WES data (Table 2). In addition, none of the ten samples have any mutations in any of the 15 homologous recombination repair (HRR) genes, including *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK1*, *CKEK2*, *CDK12*, *FANCL*, *PALB2*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD54L*, and *PPP2R2A*.[26,27]

Next, we stained ATRIP on the tumor sections of the ten available tumor tissues from women with breast cancer and the *ATRIP* c.1152_1155del germline variant and on tumor sections of eight affected women and no germline *ATRIP* mutation. In normal breast glands, ATRIP was localized in the cytoplasm and nuclei of epithelial cells, myoepithelial cells, and stromal cells to variable degrees (Figure 4A). All invasive ductal carcinomas in women without the germline *ATRIP* c.1152_1155del showed diffuse staining with moderate (Figure 4B) to strong (Figure 4C) intensity (mean score = 9.0, SD = 1.9). However, tumors from women with the *ATRIP* c.1152_1155del showed a wide range of staining for ATRIP from weak and focal (Figure 4D) to diffuse and strong staining (mean score = 4.9, SD = 3.2). In women with the germline

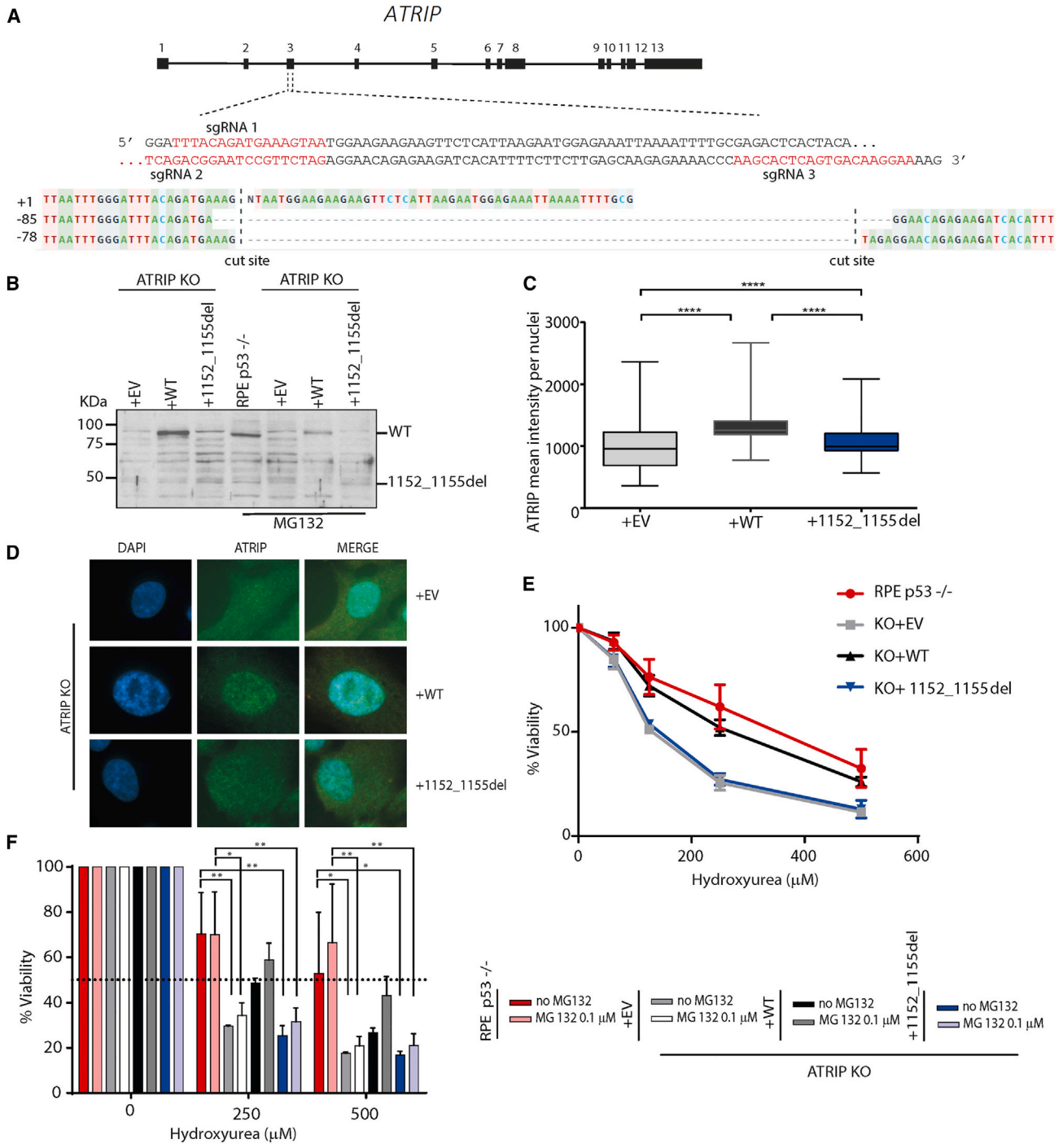

Figure 3. Assessing the function of *ATRIP* c.1152_1155del variant

(A) Generation of ATRIP-knockout (KO) cells using the CRISPR-Cas9 system in the RPE p53$^{-/-}$ background. Diagram showing the sgRNAs targeting the third exon of the human *ATRIP* (top) and contributions showing the inferred sequences present in the edited cells (bottom). Cut sites are represented as black dotted lines. RPE p53$^{-/-}$ cells were stably complemented with an empty vector (+EV), the wild type (+WT), or the truncated variant (+1152_1155del) using the AAVS1 system.

(B) Western blot analysis of the levels of ATRIP in KO complemented cells and WT RPE p53−/− cells exposed or not to MG132.

(C) Immunofluorescence analysis of the ATRIP levels in *ATRIP* KO cells complemented with +EV, +WT, or +1152_1155del. The boxplot shows the mean intensity of ATRIP in the nucleus of each cell type. The box, in the box and whiskers plot, extends from the 25th to 75th percentiles.Whiskers range from the smallest to the largest value, where the middle line represents the median. Statistical significance was determined by Kruskal-Wallis test with Dunn's multiple comparison post-test. ****$p < 0.0001$.

(D) Representative immunofluorescence images of ATRIP staining (green).

(E) Viability curves contrasting the abilities of *ATRIP* WT and the c.1152_1155del variant to rescue HU resistance in ATRIP KO cells. RPE p53$^{-/-}$ ATRIP KO cells and complemented cells were exposed to increasing concentrations of HU and cell viability was assessed 96 h

(legend continued on next page)

ATRIP c.1152_1155del who had a weak and focal staining of ATRIP, large areas of the total protein loss were evident (Figure 4E). In none of the samples was ATRIP completely null (i.e., no score = 0). The mean ATRIP accumulation score difference between those with and without the *ATRIP* c.1152_1155del mutation was statistically significant (9.0 vs. 4.9, p = 0.03).

To further explore the association of the *ATRIP* LoF variants with the risk of breast cancer, we looked at the UKB database. On May 30, 2022, we accessed the database, and there were 15,643 White British females affected with invasive breast cancer with available exome sequence data and no pathogenic mutation in any of the high-penetrance breast cancer susceptibility genes of *BRCA1*, *BRCA2*, and *PALB2*. There were 13 individuals with a germline *ATRIP* LoF variants (Table 3) among these 15,643 women with breast cancer from the UK Biobank (1 in every 1,203 women with breast cancer). We compared that with the *ATRIP* LoF variant frequency among 157,943 White British females with no personal history of cancer and available exome-sequencing data and no pathogenic mutations in *BRCA1*, *BRCA2*, and *PALB2* that existed in the UKB database. There were 40 women with a germline *ATRIP* LoF variants (Table 3) among the 157,943 individuals in the control group (1 in every 3,948 individuals). None of the 53 individuals with a germline *ATRIP* LoF mutation among affected women and control subjects had close cryptic relatedness (pairwise PI_HAT less than 0.05 for all possible pairs). LoF variants in *ATRIP* were associated with a higher risk of breast cancer with an OR of 3.28 (95% CI = 1.76–6.14, p < 0.001).

Discussion

We showed that the founder *ATRIP* c.1152_1155del variant was associated with an increased breast cancer risk (OR = 2.14, 95% CI = 1.13–4.28, p = 0.02) through whole-exome sequencing of 510 highly familial cases of Polish women with breast cancer and 308 healthy Polish women at the discovery phase, followed by validation among 16,085 unselected Polish women with breast cancer and 9,285 unaffected Polish women. We also showed that *ATRIP* LoF variants were associated with a higher risk of breast cancer among the British population (OR = 3.28, 95% CI = 1.76–6.14, p < 0.001). Although LoF variants in *ATRIP* seem relatively rare (1 in every 1,000–3,000 of the general population), these mutations are associated with a significant risk of breast cancer, suggesting *ATRIP* is a moderately penetrant gene for hereditary breast cancer. We showed that the *ATRIP* c.1152_1155del variant is deleterious for the gene function using *in cellulo* functional assays. We observed LOH in the tumor cells of at least 40% of the women with the germline *ATRIP* c.1152_1155del. We also showed homologous recombination deficiency in the tumor cells of most individuals with the germline *ATRIP* c.1152_1155del, including all those with a LOH.

ATRIP is on chromosome 3p21 and encodes an essential component of the DNA damage checkpoint. Single-stranded DNA (ssDNA) is coated with high-affinity replication protein A (RPA) at sites of DNA damage or stalled replication forks, suppressing further resections via negative feedback.[36,37] RPA-coated ssDNA recruits ATRIP (ATR-interacting protein) to the site of ssDNA damage. Subsequently, ATR (ataxia telangiectasia and Rad3 related) localizes to the RPA-coated ssDNA via forming a complex with ATRIP. Proper localization of the ATR-ATRIP complex on damaged ssDNA triggers the ATR-mediated CHEK1 activation for DNA damage checkpoint, which arrests the cell cycle and coordinates a replicative stress response.[35,38] Homozygous pathogenic variants in *ATRIP* and *ATR* are associated with Seckel syndrome, characterized by growth defects, neurodevelopmental malformations, and short life expectancy.[39,40] In addition, several cancers have been reported among individuals heterozygous for *ATR* pathogenic variants, including oropharyngeal, nonmelanoma skin, cervical, and breast cancers.[41]

Depleting ATRIP in human cell lines using small interfering RNA (siRNA) results in the loss of ATR destabilization and ATR-mediated checkpoint responses to DNA damage. This observation suggests that the interaction between ATRIP and RPA-ssDNA has a critical role in recruiting ATR to sites of DNA damage and initiating DNA damage signals.[38,42] Multiple regions of ATRIP have been shown to associate with RPA-ssDNA.[35] ATRIP mutant cell lines lacking the N-terminal 107 amino acids fail to bind to RPA-ssDNA.[37] *In vitro* translation of different ATRIP fragments and testing their binding capability to ssDNA have shown the presence of two RPA-ssDNA interacting domains in ATRIP outside of its N terminus, ATRIP 108–390 and ATRIP 390–791.[43] Studying random fragment libraries of ATRIP has shown that ATR-interacting domains exist at the C terminus of ATRIP.[35,44] The removal of *ATRIP* exon 11 in mammalian cells that codes for amino acids 658–684 severely compromised ATRIP binding to ATR. Thus, amino acids 658–684 of ATRIP truncated in the *ATRIP* c.1152_1155del variant seem essential for efficient binding to ATR.[35,38]

Conditional ATRIP knockout cell lines lacking the ATR binding domains of ATRIP at the C terminus show a considerably reduced monoubiquitination of FANCD2 and FANCI and undetectable FANCI phosphorylation, suggesting that ATR-ATRIP complex may trigger activation of

post-treatment. Quantification of surviving Hoechst-stained nucleus is represented as percent survival relative to the control condition. Error bars represent the standard error of the mean of three independent experiments.
(F) The panel of cells was treated with a combination of the indicated doses of MG132 and HU and viability was assessed 72 h post-treatment. Error bars represent the standard error of the mean of three independent experiments. Statistical significance was determined by two-way ANOVA with Tukey's multiple comparison post-test. *p < 0.05; **p < 0.01.

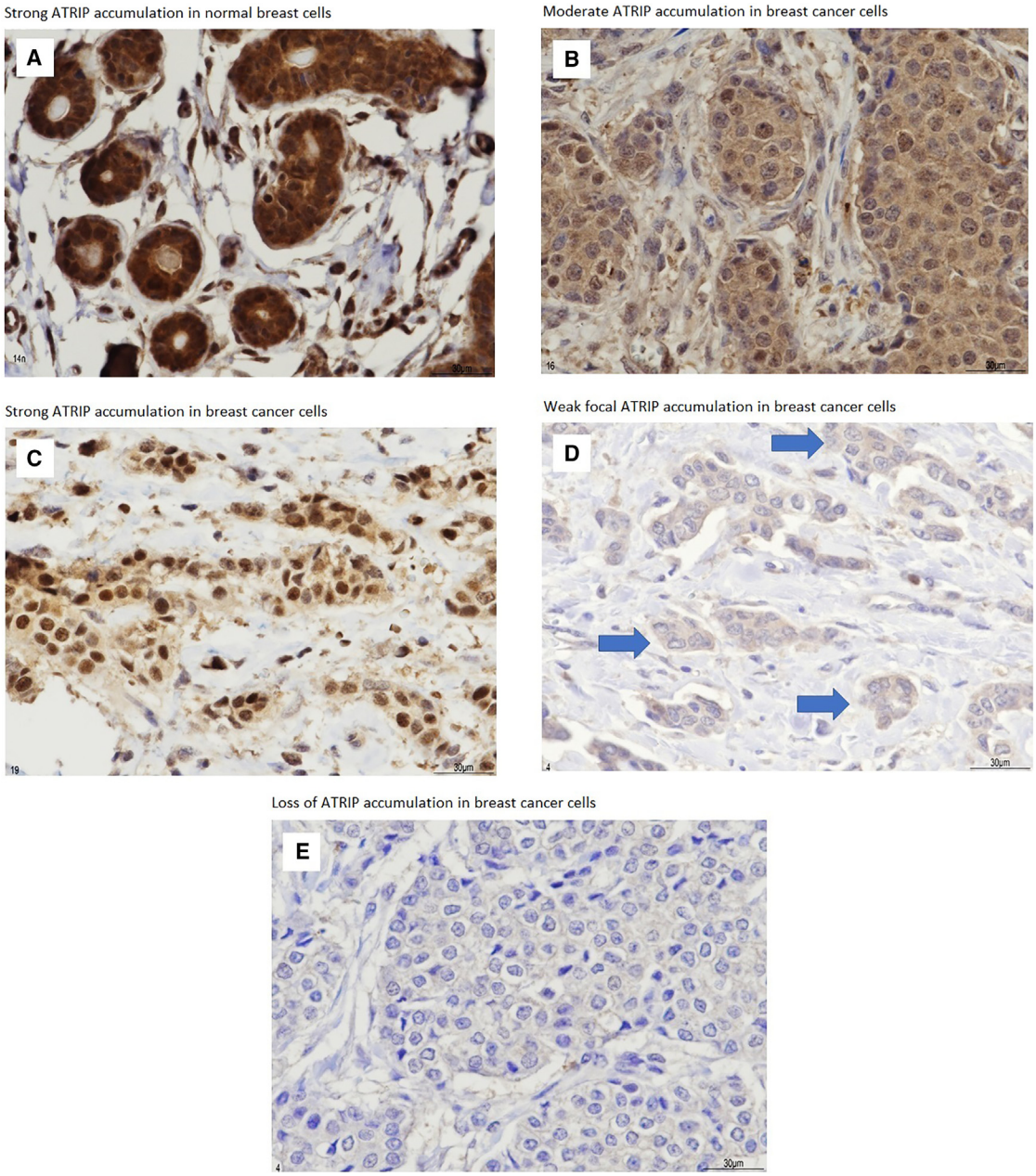

Figure 4. Staining ATRIP by immunohistochemistry
(A) Moderate to strong cytoplasmic and nuclear localization in epithelial, myoepithelial, and stromal cells of normal breast tissue.
(B and C) Moderate (B) to strong (C) accumulation in invasive ductal carcinoma tissue without a germline *ATRIP* mutation.
(D and E) Small foci of weak and focal accumulation mainly in cytoplasm of cancer cells as shown by arrow (D) adjacent to large area with loss of accumulation (E) in some tumors with a germline *ATRIP* mutation.

the Fanconi Anemia DNA Repair Pathway.[45,46,47] We also showed that many *ATRIP*-mutated tumors, including all samples with LOH at the site of germline mutated *ATRIP*, had HRD. However, the direct association of *ATRIP* or *ATR* mutations with HRD in tumor cells needs to be explored further in future studies, although dysfunctional CHEK1 (the protein activated by ATR-ATRIP complex) is known to be associated with HRD.[48,49,50] It is known that ATR inhibitors result in replication stress and genomic instability by inducing double-strand breakage of DNA at the site of stalled replication fork and inhibiting the repair of the broken DNA.[51,52] ATR inhibitors make tumor cells sensitive to poly (ADP-ribose) polymerase (PARP) inhibitors and sensitize again the cancer cells with acquired resistance to PARP inhibitor due to *BRCA1/2* reversion mutations.[53,54]

One limitation of our study is the small cohort size at the Polish discovery phase which can adversely affect the study power to detect association. We also did our best to contact the 42 Polish probands with the germline *ATRIP*

Table 3. *ATRIP* loss-of-function variants identified among 15,643 women with breast cancer and 157,943 control subjects from UK Biobank

DNA variant[a]	Protein effect	Exon	**Number of** affected women **with *ATRIP* mutation**	**Number of** control subjects **with *ATRIP* mutation**
c.16_20delinsATGCAAG	p.Ala6Metfs*67	1	2	2
c.51_63del	p.Pro18Argfs*50	1	0	2
c.217C>T	p.Gln73Ter	1	2	4
c.243_244del	p.Ser82Glnfs*2	1	1	3
c.247+2T>C	–	1	1	0
c.253del	p.His85Ilefs*7	2	0	1
c.269T>G	p.Leu90Ter	2	0	1
c.290dup	p.Asn97Lysfs*13	2	3	4
c.337G>T	p.Glu113Ter	2	1	3
c.376_379del	p.Glu126Argfs*2	2	0	2
c.436C>T	p.Arg146Ter	3	0	2
c.514dup	p.Gln172Profs*5	3	0	2
c.672-1G>C	–	5	1	0
c.721C>T	p.Gln241Ter	5	0	2
c.829+1G>C	–	5	0	1
c.830-2A>G	–	6	1	0
c.1103_1112del	p.Gly368Alafs*5	8	0	1
c.1152_1155del	p.Gly385Ter	8	0	1
c.1270C>T	p.Gln424Ter	8	0	1
c.1703del	p.Leu568Trpfs*2	8	0	1
c.1710del	p.Lys570Asnfs*2	8	0	1
c.2055+1G>A	–	11	0	2
c.2073_2074del	p.Val692Aspfs*57	12	0	1
c.2123dup	p.Arg709Lysfs*41	12	0	1
c.2192C>A	p.Ser731Ter	12	0	1
c.2263del	p.Val755Serfs*11	12	0	1
c.2278C>T	p.Arg760Ter	12	1	0

[a]Nucleotide positions are based on the GenBank: NM_130384.3 transcript of *ATRIP*.

c.1152_1155del and their family members for co-segregation analysis, but we were able to reach only six probands for the co-segregation analysis.

In summary, we identified *ATRIP* as a breast cancer susceptibility gene candidate in the founder population of Poland. Our result suggests that the *ATRIP* c.1152_1155del mutation is associated with an increased risk of breast cancer in this founder population. This observation prompted us to investigate further the association of *ATRIP* with breast cancer risk among other populations. By analyzing the UK Biobank sequence data, we showed that germline LoF mutations in *ATRIP* also predispose to an increased risk of breast cancer among the UK Biobank study participants. Functional and immunohistochemistry studies also provided evidence for the pathogenicity of the Polish *ATRIP* founder mutation. In principle, the strategy to recruit and study affected women with breast cancer from a founder population such as Poland provides a more homogeneous cohort with the greatest potential for discovering breast cancer susceptibility genes. Replication and validation of our findings in future studies may prompt consideration of clinical testing for mutations in this breast cancer susceptibility gene candidate.

Data and code availability

As per the consent obtained from the Polish studied subjects, we are not able to share their individual sequence data with any third parties. The exome-sequencing data of the UK Biobank participants is available through the cohort's data portal (UK Biobank database, https://www.ukbiobank.ac.uk) and we are not allowed

to distribute them. Other data such as mutation frequencies generated for this study could be shared upon request. This study did not generate any code.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.03.002.

Acknowledgments

We acknowledge the pathology lab at University Health Network, University of Toronto, Canada, for the IHC staining of the tumor tissue sections used in this study. This study was funded by National Science Centre, Poland, project# 2015/17/B/NZ5/02543, and the Canadian Institutes of Health Research, grant# 152939 and FDN-388879.

Author contributions

QA Conceptualizition: C.C., M.R.A., J.-Y.M.; Data Curation and Analysis: C.C., M.R.A., L.M., N.Z., W.K., A.A.S.,; Funding Acquisition: C.C., M.R.A., J.-Y.M., J.L.; Methodology: C.C., M.R.A., L.M., N.Z., J.-Y.M.; Project Administration: C.C., M.R.A., D.W., K.S., H.R., S.Z., M.Z.; Resources: C.C., M.R.A., T.H., A.J., T.D., M.L., P.D., S.N., J.-Y.M., J.L., M.S., J.G.; Supervision: C.C., M.R.A.; Visualization: L.M., A.A.S., M.R.A., C.C.; Writing-original draft: N.Z., M.R.A.; Writing-review and editing: C.C., L.M., J.-Y.M., S.N.

Declaration of interests

M.R.A. has equity ownership in Genewsie Inc.

Received: December 6, 2022
Accepted: March 2, 2023
Published: March 27, 2023

Web resources

GenBank, https://www.ncbi.nlm.nih.gov/genbank/

References

1. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 Countries*71*, pp. 209–249.
2. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J. Clin. *68*, 394–424.
3. Pujol, P., Barberis, M., Beer, P., Friedman, E., Piulats, J.M., Capoluongo, E.D., Garcia Foncillas, J., Ray-Coquard, I., Penault-Llorca, F., Foulkes, W.D., et al. (2021). Clinical practice guidelines for BRCA1 and BRCA2 genetic testing. Eur. J. Cancer *146*, 30–47.
4. Narod, S.A., and Foulkes, W.D. (2004). BRCA1 and BRCA2: 1994 and beyond. Nat. Rev. Cancer *4*, 665–676.
5. Narod, S.A., and Salmena, L. (2011). BRCA1 and BRCA2 mutations and breast cancer. Discov. Med. *12*, 445–453.
6. Kuchenbaecker, K.B., Hopper, J.L., Barnes, D.R., Phillips, K.A., Mooij, T.M., Roos-Blom, M.J., Jervis, S., van Leeuwen, F.E., Milne, R.L., Andrieu, N., et al. (2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA *317*, 2402–2416.
7. Levy-Lahad, E., and Friedman, E. (2007). Cancer risks among BRCA1 and BRCA2 mutation carriers. Br. J. Cancer *96*, 11–15.
8. Finch, A., Beiner, M., Lubinski, J., Lynch, H.T., Moller, P., Rosen, B., Murphy, J., Ghadirian, P., Friedman, E., Foulkes, W.D., et al. (2006). Salpingo-oophorectomy and the risk of ovarian, fallopian tube, and peritoneal cancers in women with a BRCA1 or BRCA2 Mutation. JAMA *296*, 185–192.
9. Breast Cancer Association Consortium, Dorling, L., Carvalho, S., Allen, J., González-Neira, A., Luccarini, C., Wahlström, C., Pooley, K.A., Parsons, M.T., Fortuno, C., Wang, Q., Bolla, M.K., and Dennis, J. (2021). Breast cancer risk genes — association analysis in more than 113,000 women. N. Engl. J. Med. *384*, 428–439.
10. Hu, C., Hart, S.N., Gnanaolivu, R., Huang, H., Lee, K.Y., Na, J., Gao, C., Lilyquist, J., Yadav, S., Boddicker, N.J., et al. (2021). A population-based study of genes previously implicated in breast cancer. N. Engl. J. Med. *384*, 440–451.
11. Solomon, S., Das, S., Brand, R., and Whitcomb, D.C. (2012). Inherited pancreatic cancer syndromes. Cancer J. *18*, 485–491.
12. Angeli, D., Salvi, S., and Tedaldi, G. (2020). Genetic predisposition to breast and ovarian cancers: how many and which genes to test? Int. J. Mol. Sci. *21*, 1128.
13. Couch, F.J., Nathanson, K.L., and Offit, K. (2014). Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science *343*, 1466–1470.
14. Cybulski, C., Carrot-Zhang, J., Kluźniak, W., Rivera, B., Kashyap, A., Wokołorczyk, D., Giroux, S., Nadaf, J., Hamel, N., Zhang, S., et al. (2015). Germline RECQL mutations are associated with breast cancer susceptibility. Nat. Genet. *47*, 643–646.
15. Rivera, B., Di Iorio, M., Frankum, J., Nadaf, J., Fahiminiya, S., Arcand, S.L., Burk, D.L., Grapton, D., Tomiak, E., Hastings, V., et al. (2017). Functionally Null *RAD51D* missense mutation associates strongly with ovarian carcinoma. Cancer Res. *77*, 4517–4529.
16. Cybulski, C., Kluźniak, W., Huzarski, T., Wokołorczyk, D., Kashyap, A., Rusak, B., Stempa, K., Gronwald, J., Szymiczek, A., Bagherzadeh, M., et al. (2019). The spectrum of mutations predisposing to familial breast cancer in Poland. Int. J. Cancer *145*, 3311–3320.
17. Kluźniak, W., Wokołorczyk, D., Kashyap, A., Jakubowska, A., Gronwald, J., Huzarski, T., Byrski, T., Dębniak, T., Gołąb, A., Gliniewicz, B., et al. (2013). The G84E mutation in the HOXB13 gene is associated with an increased risk of prostate cancer in Poland. Prostate *73*, 542–548.
18. Szwiec, M., Jakubowska, A., Górski, B., Huzarski, T., Tomiczek-Szwiec, J., Gronwald, J., Dębniak, T., Byrski, T., Kluźniak, W., Wokołorczyk, D., et al. (2015). Recurrent mutations of BRCA1 and BRCA2 in Poland: an update. Clin. Genet. *87*, 288–292.
19. Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J.P., Roques, C., Pandita, R.K., Paquet, E., Herst, P., Gingras, A.C., Pandita, T.K., et al. (2016). The TIP60 complex regulates bivalent chromatin recognition by 53BP1 through direct H4K20me binding and H2AK15 acetylation. Mol. Cell *62*, 409–421.
20. Rodrigue, A., Margaillan, G., Torres Gomes, T., Coulombe, Y., Montalban, G., da Costa E Silva Carvalho, S., Milano, L., Ducy, M., De-Gregoriis, G., Dellaire, G., et al. (2019). A global functional analysis of missense mutations reveals two major

hotspots in the PALB2 tumor suppressor. Nucleic Acids Res. *47*, 10662–10677.
21. Castroviejo-Bermejo, M., Cruz, C., Llop-Guevara, A., Gutiérrez-Enríquez, S., Ducy, M., Ibrahim, Y.H., Gris-Oliver, A., Pellegrino, B., Bruna, A., Guzmán, M., et al. (2018). A RAD 51 assay feasible in routine tumor samples calls PARP inhibitor response beyond BRCA mutation. EMBO Mol. Med. *10*, e9172.
22. Abkevich, V., Timms, K.M., Hennessy, B.T., Potter, J., Carey, M.S., Meyer, L.A., Smith-McCune, K., Broaddus, R., Lu, K.H., Chen, J., et al. (2012). Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br. J. Cancer *107*, 1776–1782.
23. Birkbak, N.J., Wang, Z.C., Kim, J.Y., Eklund, A.C., Li, Q., Tian, R., Bowman-Colin, C., Li, Y., Greene-Colozzi, A., Iglehart, J.D., et al. (2012). Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. Cancer Discov. *2*, 366–375.
24. Timms, K.M., Abkevich, V., Hughes, E., Neff, C., Reid, J., Morris, B., Kalva, S., Potter, J., Tran, T.V., Chen, J., et al. (2014). Association of BRCA1/2defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. Breast Cancer Res. *16*, 475.
25. Popova, T., Manié, E., Rieunier, G., Caux-Moncoutier, V., Tirapo, C., Dubois, T., Delattre, O., Sigal-Zafrani, B., Bollet, M., Longy, M., et al. (2012). Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res. *72*, 5454–5462.
26. Qing, T., Wang, X., Jun, T., Ding, L., Pusztai, L., and Huang, K.L. (2021). Genomic determinants of homologous recombination deficiency across human cancers. Cancers *13*, 4572.
27. Hirsch, S., Gieldon, L., Sutter, C., Dikow, N., and Schaaf, C.P. (2021). Germline testing for homologous recombination repair genes—opportunities and challenges. Genes Chromosomes Cancer *60*, 332–343.
28. Koo, C.L., Kok, L.F., Lee, M.Y., Wu, T.S., Cheng, Y.W., Hsu, J.D., Ruan, A., Chao, K.C., and Han, C.P. (2009). Scoring mechanisms of p16INK4a immunohistochemistry based on either independent nucleic stain or mixed cytoplasmic with nucleic expression can significantly signal to distinguish between endocervical and endometrial adenocarcinomas in a tissue microarray study. J. Transl. Med. *7*, 25.
29. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209.
30. Shemirani, R., Belbin, G.M., Avery, C.L., Kenny, E.E., Gignoux, C.R., and Ambite, J.L. (2021). Rapid detection of identity-by-descent tracts for mega-scale datasets. Nat. Commun. *12*, 3546.
31. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). 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. *17*, 405–424.
32. Karczewski, K.J., Weisburd, B., Thomas, B., Solomonson, M., Ruderfer, D.M., Kavanagh, D., Hamamsy, T., Lek, M., Samocha, K.E., Cummings, B.B., et al. (2017). The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res. *45*, D840–D845.
33. Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542–1548.
34. Koç, A., Wheeler, L.J., Mathews, C.K., and Merrill, G.F. (2004). Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. J. Biol. Chem. *279*, 223–230.
35. Ball, H.L., Myers, J.S., and Cortez, D. (2005). ATRIP binding to replication protein a-single-stranded DNA promotes ATR–ATRIP localization but is dispensable for Chk1 phosphorylation. Mol. Biol. Cell *16*, 2372–2381.
36. Liu, T., and Huang, J. (2016). Replication protein A and more: single-stranded DNA-binding proteins in eukaryotic cells. Acta Biochim. Biophys. Sin. *48*, 665–670.
37. Krasikova, Y.S., Rechkunova, N.I., and Lavrik, O.I. (2016). Replication protein A as a major eukaryotic single-stranded DNA-binding protein and its role in DNA repair. Mol. Biol. *50*, 649–662.
38. Cortez, D., Guntuku, S., Qin, J., and Elledge, S.J. (2001). ATR and ATRIP: partners in checkpoint signaling. Science *294*, 1713–1716.
39. Ogi, T., Walker, S., Stiff, T., Hobson, E., Limsirichaikul, S., Carpenter, G., Prescott, K., Suri, M., Byrd, P.J., Matsuse, M., et al. (2012). Identification of the First ATRIP–deficient patient and novel mutations in ATR define a clinical spectrum for ATR–ATRIP Seckel syndrome. PLoS Genet. *8*, e1002945.
40. O'Driscoll, M., Ruiz-Perez, V.L., Woods, C.G., Jeggo, P.A., and Goodship, J.A. (2003). A splicing mutation affecting expression of ataxia–telangiectasia and Rad3–related protein (ATR) results in Seckel syndrome. Nat. Genet. *33*, 497–501.
41. Tanaka, A., Weinel, S., Nagy, N., O'Driscoll, M., Lai-Cheong, J.E., Kulp-Shorten, C.L., Knable, A., Carpenter, G., Fisher, S.A., Hiragun, M., et al. (2012). Germline mutation in ATR in autosomal- dominant oropharyngeal cancer syndrome. Am. J. Hum. Genet. *90*, 511–517.
42. Yang, X.H., and Zou, L. (2006). Recruitment of ATR-ATRIP, Rad17, and 9-1-1 Complexes to DNA Damage. Methods Enzymol. *409*, 118–131.
43. Namiki, Y., and Zou, L. (2006). ATRIP associates with replication protein A-coated ssDNA through multiple interactions. Proc. Natl. Acad. Sci. USA *103*, 580–585.
44. Falck, J., Coates, J., and Jackson, S.P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. Nature *434*, 605–611.
45. Tomida, J., Itaya, A., Shigechi, T., Unno, J., Uchida, E., Ikura, M., Masuda, Y., Matsuda, S., Adachi, J., Kobayashi, M., et al. (2013). A novel interplay between the Fanconi anemia core complex and ATR-ATRIP kinase during DNA cross-link repair. Nucleic Acids Res. *41*, 6930–6941.
46. Shigechi, T., Tomida, J., Sato, K., Kobayashi, M., Eykelenboom, J.K., Pessina, F., Zhang, Y., Uchida, E., Ishiai, M., Lowndes, N.F., et al. (2012). ATR–ATRIP kinase complex triggers activation of the fanconi anemia DNA repair pathway. Cancer Res. *72*, 1149–1156.
47. Beumer, J.H., Fu, K.Y., Anyang, B.N., Siegfried, J.M., and Bakkenist, C.J. (2015). Functional analyses of ATM, ATR and Fanconi anemia proteins in lung carcinoma: ATM, ATR and FA in lung carcinoma. BMC Cancer *15*, 649.
48. Zhang, C., Zhu, D., Qu, Y., Shi, M., Ma, J., Peng, Y., Zhu, B., Tao, H., Ma, T., and Hou, T. (2023). Profiling of the genetic features of Chinese patients with gastric cancer with

HRD germline mutations in a large-scale retrospective study. J. Med. Genet. https://doi.org/10.1136/jmg-2022-108816.

49. Hu, B., Wang, H., Wang, X., Lu, H.R., Huang, C., Powell, S.N., Huebner, K., and Wang, Y. (2005). Fhit and CHK1 have opposing effects on homologous recombination repair. Cancer Res. *65*, 8613–8616.
50. Sørensen, C.S., Hansen, L.T., Dziegielewski, J., Syljuåsen, R.G., Lundin, C., Bartek, J., and Helleday, T. (2005). The cell-cycle checkpoint kinase Chk1 is required for mammalian homologous recombination repair. Nat. Cell Biol. *7*, 195–201.
51. Tercero, J.A., and Diffley, J.F. (2001). Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. Nature *412*, 553–557.
52. Buisson, R., Niraj, J., Rodrigue, A., Ho, C.K., Kreuzer, J., Foo, T.K., Hardy, E.J.L., Dellaire, G., Haas, W., Xia, B., et al. (2017). Coupling of homologous recombination and the checkpoint by ATR. Mol. Cell *65*, 336–346.
53. Ogiwara, H., Ui, A., Shiotani, B., Zou, L., Yasui, A., and Kohno, T. (2013). Curcumin suppresses multiple DNA damage response pathways and has potency as a sensitizer to PARP inhibitor. Carcinogenesis *34*, 2486–2497.
54. Yazinski, S.A., Comaills, V., Buisson, R., Genois, M.M., Nguyen, H.D., Ho, C.K., Todorova Kwan, T., Morris, R., Lauffer, S., Nussenzweig, A., et al. (2017). ATR inhibition disrupts rewired homologous recombination and fork protection pathways in PARP inhibitor-resistant BRCA-deficient cancer cells. Genes Dev. *31*, 318–332.