



HAL
open science

Altered regulation of BRCA1 exon 11 splicing is associated with breast cancer risk in carriers of BRCA1 pathogenic variants

Gorka Ruiz de Garibay, Ignacio Fernandez-Garcia, Sylvie Mazoyer, Flavia Leme de Calais, Pietro Ameri, Sangeetha Vijayakumar, Haydeliz Martinez-Ruiz, Francesca Damiola, Laure Barjhoux, Mads Thomassen, et al.

► To cite this version:

Gorka Ruiz de Garibay, Ignacio Fernandez-Garcia, Sylvie Mazoyer, Flavia Leme de Calais, Pietro Ameri, et al.. Altered regulation of BRCA1 exon 11 splicing is associated with breast cancer risk in carriers of BRCA1 pathogenic variants. *Human Mutation*, 2021, 42 (11), pp.1488-1502. 10.1002/humu.24276 . hal-04808367

HAL Id: hal-04808367

<https://hal.science/hal-04808367v1>

Submitted on 28 Nov 2024






HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Altered regulation of *BRCA1* exon 11 splicing is associated with breast cancer risk in carriers of *BRCA1* pathogenic variants

Gorka Ruiz de Garibay¹  | Ignacio Fernandez-Garcia² | Sylvie Mazoyer³  | Flavia Leme de Calais⁴ | Pietro Ameri² | Sangeetha Vijayakumar² | Haydeliz Martinez-Ruiz² | Francesca Damiola⁵ | Laure Barjhoux⁵ | Mads Thomassen⁶ | Lars v. B. Andersen⁶ | Carmen Herranz¹ | Francesca Mateo¹ | Luis Palomero¹ | Roderic Espín¹ | Antonio Gómez⁷ | Nadia García¹ | Daniel Jimenez¹ | Núria Bonifaci¹ | Ana I. Extremera¹ | Julio Castaño⁸ | Angel Raya^{8,9,10} | Eduardo Eyra^{10,11} | Xose S. Puente^{12,13} | Joan Brunet¹⁴ | Conxi Lázaro^{13,14} | GEMO¹⁵ | CIMBA¹⁶ | Paolo Radice¹⁷ | Daniel R. Barnes¹⁶ | Antonis C. Antoniou¹⁶ | Amanda B. Spurdle¹⁸  | Miguel de la Hoya^{13,19}  | Diana Baralle^{4,20} | Mary Helen Barcellos-Hoff^{2,21} | Miquel A. Pujana¹ 

¹ProCURE, Oncobell, Catalan Institute of Oncology, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, Barcelona, Catalonia, Spain

²Department of Radiation Oncology, New York University School of Medicine, New York, New York, USA

³Equipe GENDEV, INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Université Lyon 1, Université St Etienne, Lyon, France

⁴School of Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK

⁵Department of Biopathology, Pathology Research Platform, Centre Léon Bérard, Lyon, France

⁶Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark

⁷Gene Regulation, Stem Cells and Cancer, Center for Genomic Regulation (CRG), Barcelona, Catalonia, Spain

⁸Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL) and Program for Clinical Translation of Regenerative Medicine in Catalonia (P-CMRC), L'Hospitalet del Llobregat, Barcelona, Spain

⁹Centre for Networked Biomedical Research on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain

¹⁰Catalan Institution for Research and Advanced Studies, Barcelona, Spain

¹¹Department of Genome Sciences, The John Curtin School of Medical Research, EMBL Australia Partner Laboratory Network, Australian National University, Canberra, Australia

¹²Department of Biochemistry and Molecular Biology, University Institute of Oncology, University of Oviedo, Oviedo, Spain

¹³Biomedical Research Centre in Cancer (CIBERONC), Instituto Salud Carlos III, Madrid, Spain

¹⁴Hereditary Cancer Program, Catalan Institute of Oncology, Oncobell, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, and Girona Biomedical Research Institute (IDIBGI), Girona, Catalonia, Spain

¹⁵Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon/Centre Léon Bérard, Lyon, France

¹⁶Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK

¹⁷Unit of Molecular Bases of Genetic Risk and Genetic Testing, Research Department, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

¹⁸Genetics and Computational Division, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia

¹⁹Molecular Oncology Laboratory, Hospital Clínico San Carlos, Health Research Institute of the Hospital Clínico San Carlos (IdISSC), Madrid, Spain

²⁰Wessex Clinical Genetics Service, Southampton University Hospital NHS Trust, Southampton, UK

²¹Department of Radiation Oncology, School of Medicine, University of California San Francisco, San Francisco, California, USA

Correspondence

Miquel A. Pujana, ProCURE, Catalan Institute of Oncology, Bellvitge Institute for Biomedical Research (IDIBELL), Gran Via 199, Hospital Duran i Reynals, L'Hospitalet del Llobregat, Barcelona 08908, Catalonia, Spain.
Email: mapujana@iconcologia.net

Present address

Pietro Ameri, IRCCS (Scientific Institute for Research and Healthcare) Ospedale Policlinico San Martino, Genova, Italy.

Antonio Gómez, Rheumatology Department and Rheumatology Research Group, Vall d'Hebron Hospital Research Institute (VHIR), Barcelona, Spain.

Julio Castaño, Cellular and Advanced Therapy Service, Banc de Sang i Teixits (BST), Barcelona, Spain.

Cellular and Advanced Therapy Service, Banc de Sang i Teixits (BST), Barcelona, Spain.

Funding information

European Commission, Grant/Award Number: FEDER funds –a way to build Europe–; Instituto de Salud Carlos III, Grant/Award Numbers: CIBERONC, CIBERBBN, PI16/00563, PI18/01029, PI19/00553, PI21/01306; National Health and Medical Research Council, Grant/Award Number: APP-1177524; U.S. Department of Defense, Grant/Award Numbers: W81XWH-11-1-0779, W81XWH-11-1-780; Associazione Italiana per la Ricerca sul Cancro, Grant/Award Number: 22093; Novartis Pharma, Grant/Award Number: SOM230BUS03; NIHR Health Services and Delivery Research Programme, Grant/Award Number: RP-2016-07-011; Departament d'Innovació, Universitats i Empresa, Generalitat de Catalunya, Grant/Award Numbers: PERIS MedPerCan, PERIS PFI-Salut SLT017-20-000076, SGR 2017-449, SGR 2017-899, SGR 2017-1282, URDCat, CERCA Programme; Ministry of Economy and Competitiveness-MINECO, Grant/Award Numbers: RTI2018-095377-B-100, RD16/0011/0024

Abstract

Germline pathogenic variants in *BRCA1* confer a high risk of developing breast and ovarian cancer. The *BRCA1* exon 11 (formally exon 10) is one of the largest exons and codes for the nuclear localization signals of the corresponding gene product. This exon can be partially or entirely skipped during pre-mRNA splicing, leading to three major in-frame isoforms that are detectable in most cell types and tissue, and in normal and cancer settings. However, it is unclear whether the splicing imbalance of this exon is associated with cancer risk. Here we identify a common genetic variant in intron 10, rs5820483 (NC_000017.11:g.43095106_43095108dup), which is associated with exon 11 isoform expression and alternative splicing, and with the risk of breast cancer, but not ovarian cancer, in *BRCA1* pathogenic variant carriers. The identification of this genetic effect was confirmed by analogous observations in mouse cells and tissue in which a *loxP* sequence was inserted in the syntenic intronic region. The prediction that the rs5820483 minor allele variant would create a binding site for the splicing silencer hnRNP A1 was confirmed by pull-down assays. Our data suggest that perturbation of *BRCA1* exon 11 splicing modifies the breast cancer risk conferred by pathogenic variants of this gene.

KEYWORDS

BRCA1, breast cancer, isoform, risk, splicing, variant

1 | INTRODUCTION

Germline pathogenic variants in *BRCA1* are associated with a cumulative risk of female breast cancer of 40%–87% by the age of 70 years (Ford et al., 1998; Kuchenbaecker et al., 2017). The variable penetrance of *BRCA1* pathogenic variants is partially due to environmental, lifestyle, and individual biological factors (Howell et al., 2014). Dozens of common genetic variants can modify risk through the combination of their relatively small effects (Milne & Antoniou, 2016). Combined analyses of these modifiers and other factors offer the opportunity of improving breast cancer risk estimation and prevention (Barnes et al.,

2020). It was earlier hypothesized that genetic variation in the *BRCA1* wild-type allele contributes to risk modification in individuals with a *BRCA1* pathogenic variant (Cox et al., 2011). This study identified common variants in the *BRCA1* promoter that influence gene expression and, thus, are a potential source of differences in risk.

The human *BRCA1* gene contains 23 coding exons known to be subject to a plethora of alternative splicing events in normal and cancer cells (Colombo et al., 2014; de Jong et al., 2017; Romero et al., 2015). The functional significance of most of these events is unknown, but given the expected differences in the assembly of protein functional domains, some are predicted to have profound

consequences for tumorigenesis (Rebeck et al., 2015). Exon 11 (formally exon 10, due to historical misannotation of an additional "exon 4") is the largest in *BRCA1* (3426 base pairs [bp]) and codes for protein signals that are necessary for nuclear localization of the gene product and therefore considered relevant to the control of cell cycle and DNA damage repair (Li et al., 2019). Exon 11 is alternatively spliced in most cell and tissue types, including normal and breast cancer cells (Colombo et al., 2014; de Jong et al., 2017; Romero et al., 2015; Tamaro et al., 2012), and across cell-cycle phases (Orban & Olah, 2001). This splicing regulation results in the production of three main in-frame transcripts: a full-length isoform (including exon 11); $\Delta 11$ isoform (skipping of the entire exon 11); and $\Delta 11q$ isoform (partial skipping of exon 11). The $\Delta 11q$ isoform arises from the use of an alternative splice donor site located 117 bp downstream from the start of exon 11, and typically shows a higher level of expression than that of $\Delta 11$ isoform (Orban & Olah, 2001; Raponi et al., 2014; Tamaro et al., 2012; Wilson et al., 1997). Both skipping isoforms generate *BRCA1* products that lack nuclear localization signals. However, it is not clear whether alteration of the proportions of these isoforms (i.e., isoform imbalance) relative to total and/or full-length *BRCA1* expression influences breast cancer risk. Evidence from mouse studies indicates that isoform imbalance perturbs important cellular processes and can promote tumorigenesis: $\Delta 11$ over-expression alters mitosis and causes apoptosis (Bachelier et al., 2002); $\Delta 11$ exclusive expression impairs the correct DNA damage response, and triggers aneuploidy and senescence (Cao et al., 2003; Huber et al., 2001; Wang et al., 2004; Xu et al., 2001); and $\Delta 11$ -targeted deletion causes mammary ductal dilatation, glandular hyperplasia, and spontaneous tumor development in the context of *Trp53* deletion (S. S. Kim et al., 2006). Here, we identify genetic and molecular factors that influence splicing of exon 11 in mouse and human cells, and that appear to be associated with specific modification of breast cancer risk in *BRCA1* pathogenic variant carriers.

2 | MATERIALS AND METHODS

2.1 | Mice

Transgenic mice carrying exon 11 of *Brca1* flanked by *loxP* sites (*Brca1^{loxP/loxP}*) were obtained from the Mouse Repository of the National Cancer Institute (strain code 01XC8). The genetic background of *Brca1^{loxP/loxP}* mice is >90% C57BL/6. Most animals were bred as homozygotes in the VA facility and, where necessary, age-matched C57BL/6 wild-type mice, purchased from Charles River Laboratories (Charles River), were used as controls. The VA Institutional Animal Care and Use Committee (IACUC) of the School of Medicine, New York University, approved the study. Mice were maintained in accordance with the National Institute of Health (NIH) guidelines for the Care and Use of Laboratory Animals in an AALAC approved facility, in compliance with the US National Research Council's Guide for the Care and Use of Laboratory Animals, US Public Health Service's Policy on Humane Care and Use of

Laboratory Animals, and Guide for the Care and Use of Laboratory Animals.

2.2 | Mammary tissue studies

Inguinal mammary glands were analyzed at 21 and 28 days, and 2, 4, 6, 9, and 12 months of age. Whole mounts were imaged on a dissecting microscope at $\times 10$ magnification using Nikon ACT-1 software and analyzed with Image-Pro Plus software (version 4.5; Media Cybernetics). For immunostaining and histology, mammary glands were fixed in 10% neutral-buffered formalin before paraffin-embedding and sectioning. Hematoxylin-eosin-stained sections were reviewed by a pathologist without knowledge of the mouse genotype. Standard immunohistochemistry was conducted using anti-Ki67 (NeoMarkers; Thermo Fisher Scientific) primary antibody, horseradish peroxidase-conjugated goat anti-rabbit antibody (Dako), and 3'-diaminobenzidine (DAB) solution (Vector Laboratories). Omission of the latter was used as a negative control. For each animal, Ki67 staining was quantified by counting the percentage of positive cells from all the epithelial cells of various glandular structures in different sections.

2.3 | DNA analyses

DNA was extracted from mouse tails using the QIAamp DNA Minikit (Qiagen), and amplified with primers for the *loxP* insert upstream of exon 11 (B004 and B005) and deleted exon 11 (B004 and B006) provided by the National Cancer Institute (Table S1). The primers were purchased from Sigma-Aldrich. Results of polymerase chain reactions (PCRs) of tail DNA were verified as being the same as for PCRs of mammary tissue DNA by direct comparison in several animals. The PCR products were sequenced using the Sanger method. The rs5820483 genotype of K562 cells was determined by analyzing whole-genome sequencing data from ENCODE (ENCODE Project Consortium, 2012); experiment ENCSR045NDZ.

2.4 | Cell studies

Mammary epithelial cells and fibroblasts were isolated from the inguinal glands of 4-month-old nulliparous C57/BL6 and *Brca1^{loxP/loxP}* female mice, as previously described (Illa-Bochaca et al., 2010). Fibroblasts were cultured in DMEM/F-12 supplemented with 10% fetal calf serum, penicillin (100 IU/ml), streptomycin (100 μ g/ml), 4 mM L-glutamine, 5 μ g/ml bovine pancreatic insulin, 10 ng/ml cholera toxin, and 10 ng/ml epidermal growth factor. The lymphoblastoid cell lines originated from 23 *BRCA2*-pathogenic variant women carriers of the Genetic Modifiers of *BRCA1* and *BRCA2* (GEMO) group (Lesueur et al., 2018) analyzed in the Consortium of Investigators of *BRCA1/2* Modifiers (CIMBA) and balanced with respect to the rs582043 genotypes; *BRCA1*-pathogenic variant carriers were

excluded to avoid putative interferences of the mutation with the usual splicing pattern.

2.5 | RNA and BRCA1 expression analyses

RNA was extracted using the TRIzol reagent (Thermo Fisher Scientific) and complementary DNA (cDNA) synthesized from 1 µg of RNA by means of the SuperScript III First-Strand Synthesis System (Thermo Fisher Scientific) and employing a combination of random hexamers and oligodTs. cDNA was amplified using the primers shown in Table S1 and sequenced by Sanger. DNA sequences were aligned using SeqMan (DNASTAR). Expression of genes and isoforms of interest were assessed in at least three independent biological replicates, normalized against defined controls, and analyzed when appropriate with the delta cycle threshold (ΔC_t) method for quantifying transcripts. *BRCA1* exon 11 alternative splicing in RNA extracted from lymphoblastoid cell lines was quantified in triplicate in multiplex dual-color real-time polymerase chain reaction (PCR) assays, performed using a LightCycler 480 Instrument (Roche Diagnostics). We used the 4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein (VIC)-labeled custom-designed TaqMan assay specific to exon 10–12 junctions to detect $\Delta 11$, the VIC-labeled Hs01556199 assay specific to 11p-12 junction to detect $\Delta 11q$, and a 6-carboxyfluorescein (FAM)-labeled custom-designed TaqMan assay specific to exon 11q-12 junction to detect the full-length isoform (Table S1). To determine total *BRCA1* expression, we used the FAM-labeled Hs01556193_m1 TaqMan probe that targeted the boundaries of exons 23 and 24. The stability of the *HMMR* probe Hs01063280_m1 as a reference was measured against the expression of the 18S ribosomal housekeeping gene using the Hs03003631_g1 probe. For confirmation, digital real-time PCR was performed using pools of all the samples from each rs582043 genotype. The use of digital PCR for precise quantification of *BRCA1* isoforms has been described in detail elsewhere (de la Hoya et al., 2016); in this study, it was performed on a QuantStudio 3D dPCR 20K platform according to the manufacturer's instructions (Applied Biosystems). All quantification assays were performed combining FAM- and VIC-labeled assays in individual digital PCR chips. The chips were analyzed with QuantStudio 3D Analysis Suit Cloud software v2.0 (Applied Biosystems) with VIC as the declared target. Default settings were used in all cases. After reviewing the software's automatic assessment of the chip quality, only green-flag chips (i.e., those for which the data met all the quality thresholds, and for which a review of the analytical result was not required) and yellow-flag chips (i.e., those for which the data met all quality thresholds, but for which manual inspection was recommended) were considered for further analyses. Depending on the FAM and VIC assay combination, we used the target/total ratio, VIC/(FAM+VIC), calculated by the software, as a proxy for the exclusion rates of exon 11 and 11q, and/or for the exon 11 inclusion rate. For the study of P95H SRSF2, sequence read archive (SRA) data of the Gene Expression Omnibus (GEO) reference GSE123774 were downloaded, transformed to fastq using the fastq-dump NCBI tool (version 2.9.6), and aligned against mouse genome GRCm38.75 using STAR

(<https://github.com/alexandobin/STAR>). Exon counts were retrieved from aligned BAM using featureCounts (<http://bioinf.wehi.edu.au/featureCounts/>). Counts per million (CPM) were calculated from exon raw counts.

2.6 | Percent spliced in (PSI) and genotype data

PSI values of *BRCA1* isoforms in primary tumors of The Cancer Genome Atlas (TCGA) were obtained from the Genomics Data Commons repository (<https://gdc.cancer.gov/about-data/publications/PanCanAtlas-Splicing-2018>) (Kahles et al., 2018). Permission to use the individual-level data was granted by the dbGaP Data Access Committee (project #11689). *BRCA1/2* mutation status was taken from previous annotations (Kraya et al., 2019) and cBioPortal (<https://www.cbioportal.org/>).

2.7 | Geuvadis data analysis

To analyze RNA-seq data from the Geuvadis project (ArrayExpress reference E-GEUV-1), BAM files from 465 lymphoblastoid cell lines of the 1000 Genomes Project were examined to quantify the total number of reads supporting the *BRCA1* full-length isoform (corresponding to splicing between exons; Ensembl references ENSE00003522602 and ENSE00003547126), $\Delta 11q$ (between ENSE00001936588 and ENSE00003547126), or $\Delta 11$ (between ENSE00003787101 and ENSE00003547126).

2.8 | Western blot

Mouse *Brca1* was detected using antibodies that target the N-terminus (*Brca1-N*; a gift from Dr. Chodosh) and the C-terminus of *Brca1* (R&D Systems). Briefly, log-phase mammary fibroblasts were lysed in RIPA buffer and 100 µg of protein extracts were probed with *Brca1-N*. The relative density of the *Brca1* bands was normalized with respect to actin, which was used as the loading control. Immunoprecipitation was done with control (anti-GFP) and *Brca1* antibody (R&D Systems) from 3 mg extracts of mammary fibroblast. Proteins were detected using the LI-COR Odyssey system.

2.9 | Minigene assays

The pcDNA-Dup and *BRCA1* exon 10–12 minigenes have been described previously (Anczuków et al., 2008; Tournier et al., 2008). The constructs were transfected into cells using Lipofectamine LTX reagent, and RNA extraction and cDNA synthesis were performed as detailed above. The *loxP* fragments were generated by annealing complementary 5'-phosphorylated oligonucleotides designed with 5'-*EcoRI* and 3'-*BamHI* compatible ends, and cloned into the pcDNA-Dup middle exon. Exon inclusion was evaluated by performing

reverse-transcription PCR (RT-PCR) assays using cDNA obtained from transfected NIH-3T3-murine cells and the primers T7-Pro and Dup-2R (Table S1). The pEGFP10-12 *BRCA1* minigene (Anczuków et al., 2008) was used to assess the effect of rs582043 on *BRCA1* exon 11 alternative splicing. Defined alleles were introduced in the minigene through site-directed mutagenesis using the QuickChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Alternative splicing was quantified using multiplex semi-quantitative (q) RT-PCR with 25 amplification cycles and FAM-labeled primers full-length-R, isoform-R, and GFP-F (Table S1). The amount of each amplicon was quantified by measuring the area of each peak detected by capillary electrophoresis using an Applied Biosystems 3130xl Genetic Analyzer (Thermo Fisher Scientific). The area of an individual peak was referenced as the percentage of the sum of the areas of all detected peaks. Transfections and quantifications were performed in triplicate.

2.10 | Cancer risk modification

The data used in this study of women carrying *BRCA1/2* pathogenic variants were obtained by protocols approved by the ethics committees of the CIMBA contribution centers in their respective countries. Details of centers and data collection have been reported elsewhere (Antoniou et al., 2007). Information on the CIMBA consortium and participating centers can be found at <http://cimba.ccge.medschl.cam.ac.uk/> (Chenevix-Trench et al., 2007). All included carriers participated in clinical and/or research studies at the host institutions after providing informed consent under the corresponding approved protocols. We used data from the genome-wide association study (GWAS) in the Collaborative Oncological Gene-environment Study Consortium (COGS) with the contribution of CIMBA. The GWAS design, quality controls, and statistical analyses have been described previously (Couch et al., 2013; Gaudet et al., 2013; Milne et al., 2017). For each variant, a per-allele odds ratio and standard error were obtained by logistic regression, including study and principal components as covariates (Antoniou et al., 2010).

2.11 | Haplotype analysis

The *BRCA1* haplotypes were inferred from the GWAS genotypes using Haploview v4.2 (Barrett et al., 2005). The study adopted values of accuracy of $r^2 > 0.3$ and of minor allele frequency > 0.005 . Phasing was performed for each germline pathogenic *BRCA1* variant relative to rs455055—in strong linkage disequilibrium (LD) with rs5820483—by counting genotypes within each founder group. Those containing homozygous carriers for the allele not associated with dupAGG were considered to be in trans. Founder groups were identified for each *BRCA1* pathogenic variant by hierarchical clustering analysis using the reciprocal of the pairwise shared haplotype length as the distance measure and Ward's linkage criterion. The shared haplotype length was defined as the combined distance from both sides of the pathogenic variant to the first homozygous discordant variant.

2.12 | Pull-down assays and iCLIP data

RNA pull-down assays were performed as previously described (Tammaro et al., 2014). Briefly, 1 nM of rs5820483 AGG or dupAGG RNA oligonucleotides (Table S1) was treated with 5 mM sodium m-periodate and bound to adipic acid dihydrazide-agarose beads (Sigma-Aldrich). The beads with bound RNA were incubated in 1× RNA buffer containing 600 μg of HeLa nuclear extract and 5 μg/μl of heparin for 30 min at room temperature in a final volume of 400 μl. The beads were then pelleted at 1000 rpm for 5 min and washed six times with 1 ml of 1× RNA buffer, before adding sodium dodecyl sulphate (SDS) sample buffer and loading onto a 12% SDS polyacrylamide gel electrophoresis (SDS-PAGE) gel in preparation for electrophoresis. Gels were subsequently electro-blotted onto a Hybond ECL membrane (GE Healthcare Systems). Proteins were then recognized in-house antibodies against hnRNP A1 and hnRNP H. The iCLIP results of hnRNP A1 genomic binding were taken from the GEO GSE34993 for HEK293 cells (bowtie format, mapped to hg18), and from the ENCODE repository for K562 and HepG2 cells (bigwig format, mapped to hg19).

2.13 | Data and code availability

Summary-level statistics are available from the public repository of CIMBA and requests for individual data should be addressed to the Data Access Coordination Committee (<http://cimba.ccge.medschl.cam.ac.uk/>). The custom code used to analyze the Geuvadis RNA-seq data and *BRCA1* haplotypes were uploaded to GitHub: <https://github.com/xa-lab/geuvadis> and https://github.com/larsmew/hap_phasing_BRCA1_rs5820483, respectively.

3 | RESULTS

3.1 | *loxP* insertion flanking *Brca1* exon 11 may alter mammary morphogenesis

To study mammary tumorigenesis arising from *Brca1* loss we obtained *Brca1* exon 11 floxed transgenic mice, from the 01XC8 line, against a C57BL/6J and 129/svJ background (National Cancer Institute), which we refer to hereafter as being with the *Brca1^{loxP}* allele. Unexpectedly, we found that, before exposure to Cre DNA recombinase, homozygous *Brca1^{loxP/loxP}* mice exhibit a prominent hyperplastic mammary phenotype at 4 and 6 months of age, as quantified by the diameter of primary and secondary ducts relative to those of heterozygous *Brca1^{loxP/+}* and C57BL/6J mice (Figures 1a and S1). The extent of mammary cell proliferation as determined by Ki67-positive cells was greater in *Brca1^{loxP/loxP}* than in *Brca1^{loxP/+}* and C57BL/6J 4-month-old mice (Figure 1b). At pubertal age (28 and 60 days old) *Brca1^{loxP/loxP}* animals also showed significantly greater cell proliferation relative to C57BL/6J mice (Figure 1c).

3.2 | *loxP* insertion in *Brca1* intron 10 alters splicing

Sequencing of the exon 11 flanking introns in *Brca1*^{*loxP/loxP*} mice confirmed that both *loxP* sites were present as reported (Xu et al., 1999), and there were no alterations in messenger RNA (mRNA) exon 11 junctions (Figure S2). The *loxP* sequences in *Brca1*^{*loxP/loxP*} mice were inserted in regions with no significant mammalian sequence conservation, but where there was overlap with substantial signals of histone H3 trimethylation at lysine 36 (H3K36me3; Figure 2a) that could alter splicing regulation (Luco et al., 2010). Using multiplex qRT-PCR (assays $n = 3$) with mRNA extracted from *Brca1*^{*loxP/loxP*} and C57BL/6J mammary fibroblasts, the $\Delta 11$ isoform was found to be underexpressed in *Brca1*^{*loxP/loxP*} cells: 36% relative to C57BL/6J cells using a housekeeping gene control (mouse ribosomal protein L32, *mL32*; equivalent results were obtained with control *Sparc*) and 50% when full-length *Brca1* was employed as a reference (Figure 2b).

An antibody that recognizes the N-terminus of mouse BRCA1 (Huber & Chodosh, 2005) was used to immunoblot protein lysates from primary mammary fibroblasts and epithelial cells, as well as whole mammary glands. Higher levels of full-length BRCA1 (p210) were detected in *Brca1*^{*loxP/loxP*} than in C57BL/6J cells and tissue (Figure 2c). Immunoprecipitation assays ($n = 3$) with control IgG or antibody to BRCA1 C-terminus and detection with N-terminus antibody, which recognized p210 and $\Delta 11$ products, showed that *Brca1*^{*loxP/loxP*} mammary fibroblasts have approximately 50% of the amount of $\Delta 11$ product relative to p210 (Figure 2d), which was consistent with the qRT-PCR results (Figure 2b). These results suggest that inserted *loxP* sequences alter regulation of *Brca1* exon 11 alternative splicing.

The *loxP* sites were found to be inserted with accompanying sequences that differed between *Brca1* introns 10 and 11 (Figure 2e). The *loxP* and accompanying sequences showed several binding sites of specific serine/arginine-rich splicing-enhancer factors (SRSFs), as predicted using ESE Finder version 3.0 (Cartegni et al., 2003) with relatively high threshold scores: ≥ 2.0 for SRSF1 and ≥ 3.0 for SRSF2, SRSF5, and SRSF6 (Figure 2e). To assess the effect of the *loxP* and accompanying sequences on splicing regulation, three constructs were produced using the splicing enhancer reporter minigene, pCDNA-DUP (Tournier et al., 2008): two constructs containing each the core *loxP* site and the corresponding accompanying sequence (45 and 66 bp), and a third construct containing solely the core *loxP* sequence (34 bp). Each construct was transfected into murine 3T3 cells, after which exon inclusion was determined by semi-qRT-PCR assays (assays $n = 2$). The *loxP* and accompanying sequence corresponding to the intron 10 insertion caused exon inclusion (Figure 2a), but neither of the other two constructs did so. The inclusion product accounted for 15% relative to the noninclusion product (Figure 2f).

SRSF2 is the only SRSF predicted to bind exclusively at the *loxP*-accompanying sequence in intron 10. This factor was found to be positively coexpressed with *Brca1* in mammary normal and tumor tissue, as indicated by the Pearson's correlation coefficients (PCCs) > 0.35 ($p < .05$) in three datasets (Figure S3) (Anderson et al., 2007; Bai et al., 2016; Meier-Abt et al., 2013). A previous study

showed that oncogenic SRSF2-P95H introduced into mouse leukemia cells alters recognition of exonic splicing enhancers (E. Kim et al., 2015). Analysis of this data set exposed significantly higher expression of *Brca1* exon 11 in cells harboring the P95H mutation (Figure 2g). Together, the data suggest that insertion of a *loxP*-containing sequence into the *Brca1* intron 10 causes partial alteration of exon 11 splicing, which might, in turn, lead to differences in mammary ductal morphogenesis.

3.3 | Common BRCA1 genetic variation is associated with breast cancer risk

The study of *Brca1*^{*loxP/loxP*} mice and *loxP*-containing sequences suggests a link between the splicing imbalance of the largest *Brca1* exon and mammary developmental abnormalities. It had previously been shown that common variation in the *BRCA1* wild-type allele modifies breast cancer risk in individuals with a pathogenic *BRCA1* variant (Cox et al., 2011). The proposed molecular mechanism underlying this observation was founded on genetic changes in the *BRCA1* promoter sequence that can lead to differences in the overall expression of this gene. Following on from these findings, we evaluated the association of common variants at the *BRCA1* locus with modification of cancer risk using the GWAS results generated by iCOGS with the participation of CIMBA (Couch et al., 2013; Gaudet et al., 2013; Milne et al., 2017). Genotyped ($n = 18$) and imputed ($n = 329$) variants were found to be associated (defined threshold $p < 10^{-3}$) with breast cancer risk in *BRCA1* pathogenic variant carriers, but not in *BRCA2* pathogenic variant carriers (Figure 3a and Table S2). No association was observed with ovarian cancer risk for *BRCA1* or *BRCA2* pathogenic variant carriers (Table S2). The relatively extended *BRCA1* region showing breast cancer risk associations is probably due in part to the existence of high LD throughout the entire locus (Cox et al., 2005). None of the genotyped variants found to be associated with breast cancer risk in *BRCA1* pathogenic variant carriers was shown to be nominally associated with breast cancer risk in the population-based study of the Breast Cancer Association Consortium (Michailidou et al., 2017), even after consideration of estrogen receptor α status (Table S3). Indeed, earlier studies did not detect associations between *BRCA1* common variation and breast cancer risk in case-control and population-based studies (Dunning et al., 1997; Freedman et al., 2005), which suggests that the potential modifying effects are specific to breast cancer risk in *BRCA1* pathogenic variant carriers.

Of the association signals detected, those corresponding to common variants located in the *BRCA1* intron 10 region analogous to the *loxP* insertion site in mouse intron 10 were explored in detail. The *loxP* insertion site that may modify splicing regulation in mice was located 243 bp from the intron 10 acceptor site. In humans, the rs5820483 variant corresponding to an AGG duplication (dupAGG, minor allele frequency of 0.33 in European populations) was identified at 245 bp from the intron 10 acceptor site, and this was also a region with substantial amounts of H3K36me3 in a variety of human cell types (Figure 3b). This imputed variant

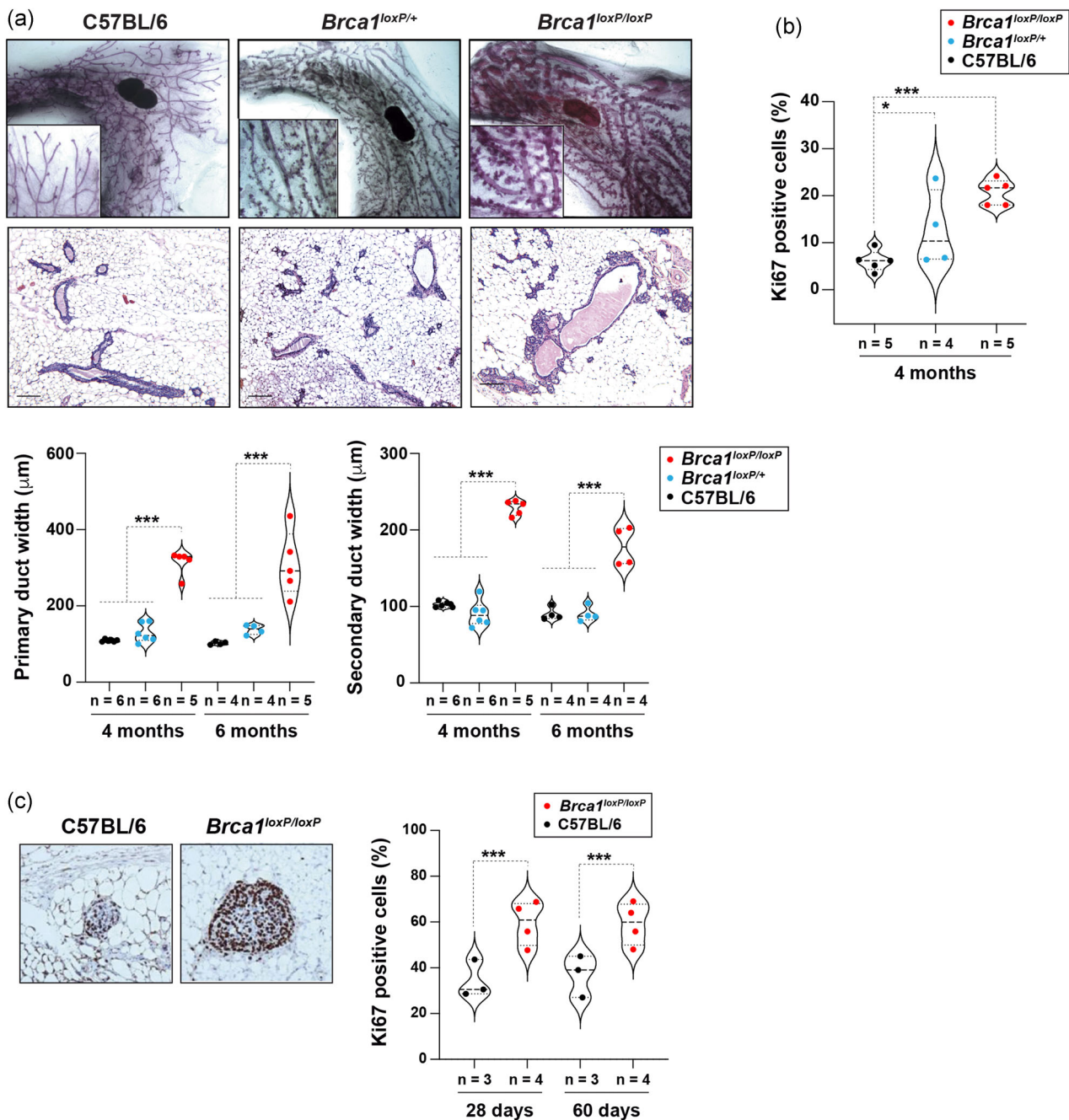


FIGURE 1 Aberrant phenotype of *Brca1*^{loxP/loxP} mammary glands. (a) Top panels, representative whole mounts (with higher-power inserts) of mammary glands from C57BL/6, *Brca1*^{loxP/+}, and *Brca1*^{loxP/loxP} animals, showing ductal dilatation and glandular hyperplasia, especially in the latter group. Middle panels, histology of the mammary glands shown above; note the cystic ducts in the *Brca1*^{loxP/loxP} tissue. Scale bar, 200 μm. Bottom panels, violin plots showing quantification of primary and secondary duct width: 9–26 ducts were measured in the mammary glands of 4–5 animals per genotype, and the data points indicate the average value of each animal in the study. The degree of significance of the two-tailed Student's *t* test is marked by asterisks (***p* < .001). (b) Quantification of Ki67-positive epithelial cells in 4-month mammary glands from C57BL/6, *Brca1*^{loxP/+}, and *Brca1*^{loxP/loxP} animals: 943–1538 cells were evaluated in 5–11 ducts of each of 4–5 animals studied per genotype, and the data points indicate the average value of each animal. The asterisks denote significant differences identified by two-tailed Student's *t*-tests (**p* < .05 and *p* < .001). (c) Quantification of the frequency of Ki67-positive epithelial cells in 28- and 60-day mammary glands from C57BL/6 and *Brca1*^{loxP/loxP} animals: 402–1108 cells were evaluated in 4–9 ducts of 3 (C57BL/6) and 4 (*Brca1*^{loxP/loxP}) animals, and the data points indicate the average value of each animal. The asterisks denote significant differences identified by two-tailed Student's *t* tests (***p* < .001). The left panels show representative images of Ki67-positive staining in 28-day mammary glands

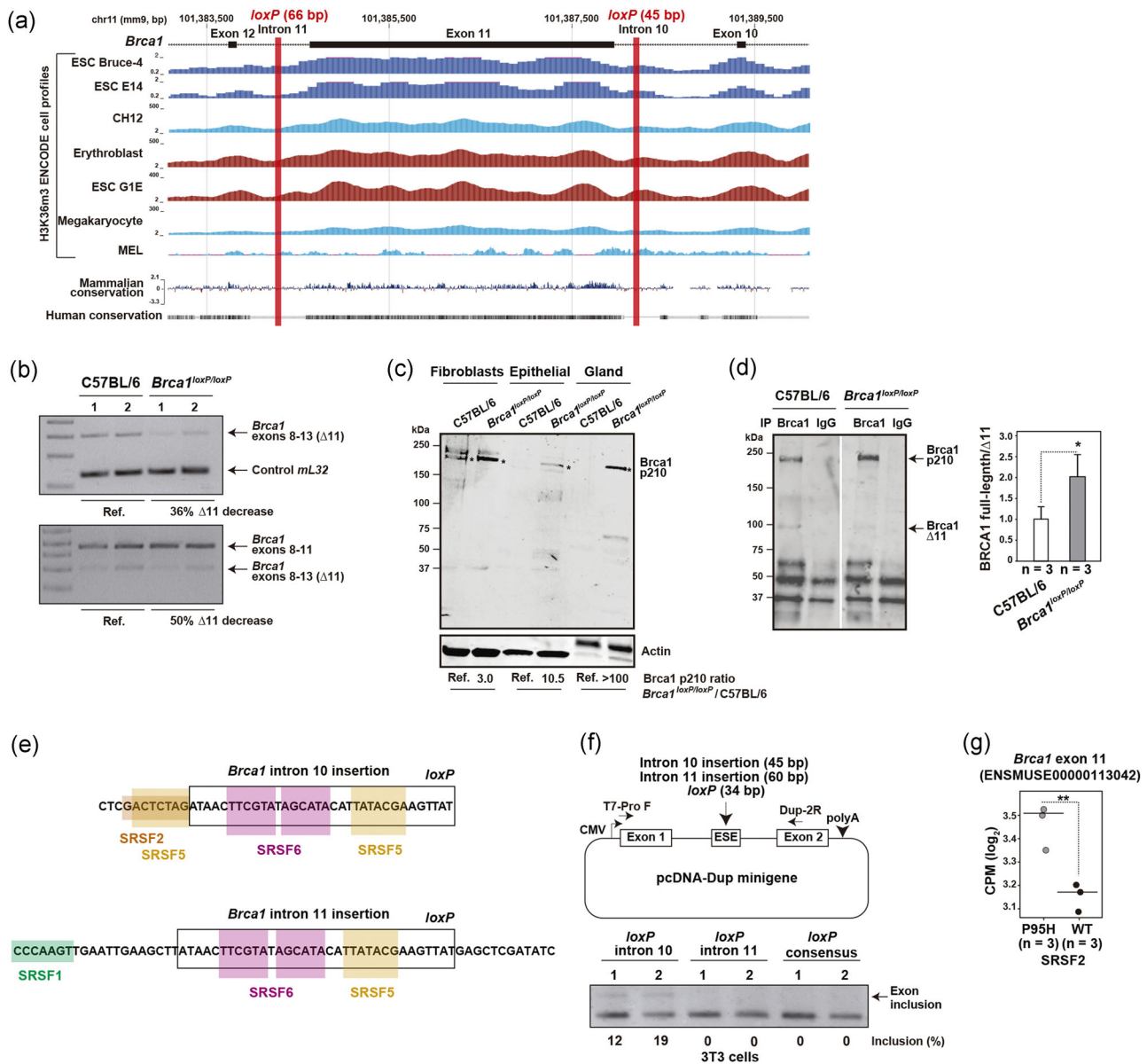


FIGURE 2 The insertion of *loxP* in *Brca1* intron 10 may alter exon 11 recognition during splicing. (a) Mouse *Brca1* genomic locus centered on exon 11 and showing the location of flanking *loxP* insertions. The graph also shows the values of H3K36m3 ENCODE cell profiles and the level of sequence conservation relative to human and mammal genomes. The ENCODE cell types are detailed in <http://genome.ucsc.edu/ENCODE/cellTypes.html>. (b) Multiplex semi-qRT-PCR results showing a 36% and 50% lower level of $\Delta 11$ isoform expression in *Brca1^{loxP/loxP}* relative to C57BL/6 mammary fibroblasts (reference [ref.] ratio) when using the expression of *mL32* (top panel) and *Brca1* including exon 11 (bottom panel), respectively. Two replicates are shown for each genotype. (c) Western blot showing a greater abundance of BRCA1 p210 (marked by asterisks) in extracts of primary fibroblasts, primary epithelial cells, and whole mammary glands of *Brca1^{loxP/loxP}* relative to C57BL/6 cells/tissue. Detection with antibody against the N-terminus of mouse BRCA1 (anti-BRCA1-N; a gift from Dr. Chodosh). The quantified BRCA1/actin ratios in *Brca1^{loxP/loxP}* relative to C57BL/6 (ref.) samples are detailed at the bottom. (d) Western blot showing a greater abundance of immunoprecipitated BRCA1 p210 relative to $\Delta 11$ -derived isoform in *Brca1^{loxP/loxP}* mammary fibroblast extracts (3 mg each). Detection of BRCA1 with R&D Systems antibody. The negative control IgG GFP immunoprecipitants are also shown. Quantification is shown in the right panel and significance based on two-tailed Student's *t* test (**p* < .05). (e) Sequences of *loxP* insertion in *Brca1* introns 10 and 11, and predicted binding sites of SRSFs. The box shows the sequence that is common to both *loxP*. (f) Top panel, pcDNA-Dup minigene construct, and *loxP* cloning for assessing exon inclusion/exclusion in murine 3T3 fibroblasts. Bottom panel, results of semi-qRT-PCR assays, and quantification of exon inclusion relative to the excluded product. This study comprised two independent transfections of each construct and, for each transfection, a multiplex qRT-PCR reaction (25 cycles) whose products were analyzed in an agarose gel. (g) *Brca1* exon 11 is significantly overexpressed (log₂ counts per million [CPM]) in primary MLL-AF9 mouse leukemia cells expressing SRSF2-P95H mutant relative to SRSF2 wild-type (WT; GEO GSE123774; E. Kim et al., 2015). The asterisks indicate a statistically significant outcome of a Wilcoxon signed-rank test (***p* < .01). IgG, immunoglobulin G; qRT-PCR, quantitative reverse-transcription polymerase chain reaction; SRSF, serine/arginine-rich splicing-enhancer factors

was among those associated with breast cancer risk in *BRCA1* pathogenic variant carriers ($n = 15,238$), with the minor allele reducing the risk, as indicated by the hazard ratio (HR) of 0.84 ($p = 7.5 \times 10^{-4}$).

Phasing the rs5820483 variant relative to the germline *BRCA1* pathogenic variant in each carrier revealed that a protective effect occurs when the dupAGG allele is in trans, that is, in the wild-type *BRCA1* allele: heterozygous dupAGG in trans, HR: 0.84, 95% confidence interval (CI) 0.79–0.88, $p < .0001$; heterozygous dupAGG in cis, HR: 0.99, 95% CI: 0.93–1.05, $p = .72$. This observation was in accordance with the originally described modifier effect(s) of the wild-type *BRCA1* allele (Cox et al., 2011).

3.4 | rs5820483 is associated with *BRCA1* exon 11 splicing imbalance

To examine the potential consequences of rs5820483 alleles for *BRCA1* exon 11 splicing, we used confirmed *BRCA1* wild-type lymphoblastoid cell lines from CIMBA with established rs5820483

iCOGS genotypes. *BRCA1* is ubiquitously expressed in normal proliferating cells, and genomic regulatory marks in this locus are relatively similar between lymphoblastoid cell lines and human mammary epithelial cells (Figure S4). Probe-based expression assays showed that lymphoblastoid cells with the dupAGG/dupAGG genotype had a significantly lower level of expression of full-length *BRCA1*, but higher levels of $\Delta 11$ and $\Delta 11q$ isoform expression than cells with no dupAGG allele (Mann–Whitney $p < .05$; logistic regression, recessive model: full-length *BRCA1* $p = .21$, $\Delta 11$ $p = .031$, and $\Delta 11q$ $p = .032$; assays $n = 3$; Figure 4a). Digital PCR (assays $n = 2$) of all samples pooled by the rs5820483 genotype confirmed the isoform imbalance: the levels of $\Delta 11$ and $\Delta 11q$ were 3.04 and 1.88 times higher in dupAGG homozygotes than in AGG homozygotes, respectively ($\Delta 11$ $1.43 \pm 0.41\%$ and $0.47 \pm 0.15\%$, and $\Delta 11q$ $19.1 \pm 1\%$ and $10.2 \pm 0.6\%$ in dupAGG and AGG homozygotes, respectively; Figure 4b). Conversely, the expression of the full-length isoform was 0.65 times lower in dupAGG homozygotes, although total *BRCA1* expression was not significantly different (Figure 4a,b).

To evaluate the previous results in an independent study, we analyzed the Geuvadis RNA-seq data of lymphoblastoid cell lines from unrelated healthy individuals (Lappalainen et al., 2013). With

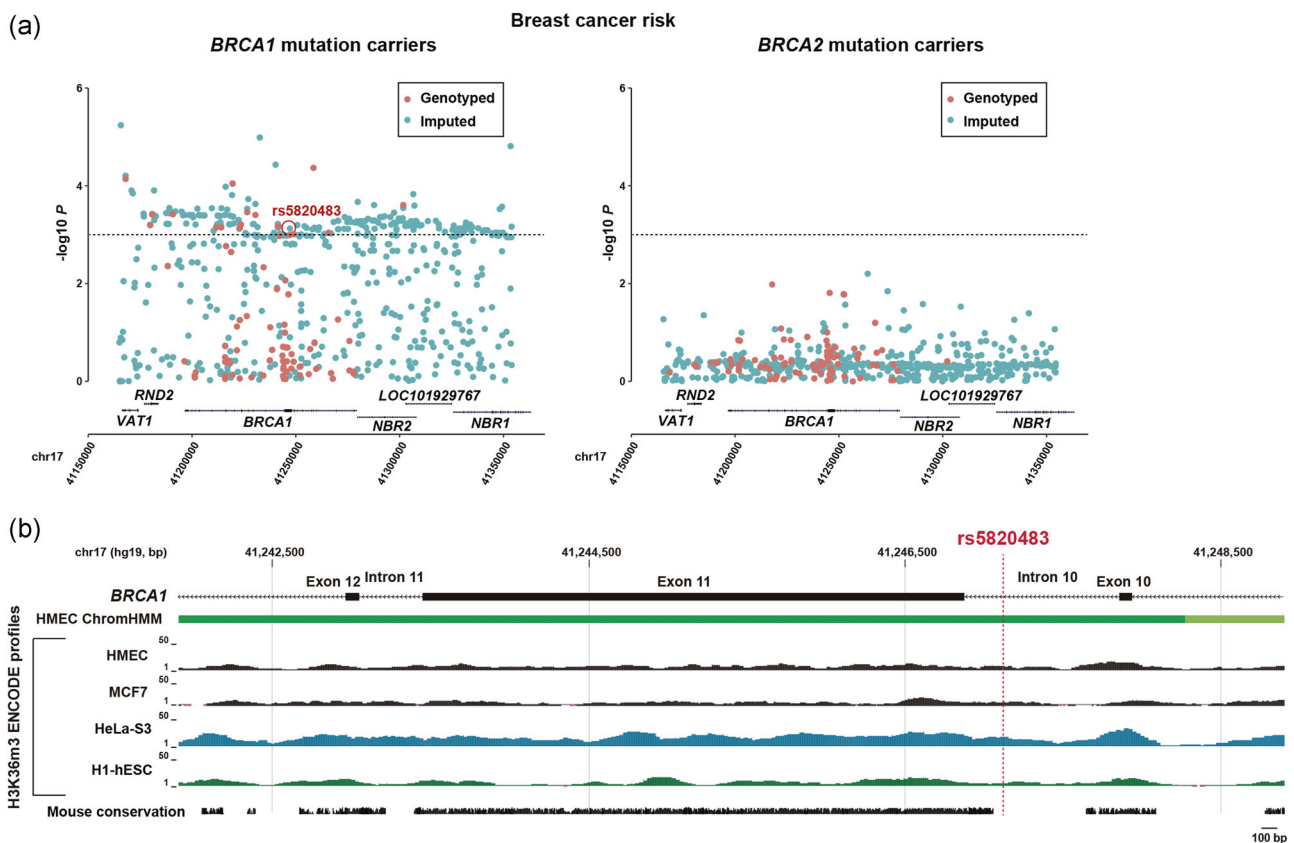


FIGURE 3 Common genetic variation in the *BRCA1* locus is associated with breast cancer risk in *BRCA1* pathogenic variant carriers. (a) Graphs showing the breast cancer association results ($-\log_{10}P$) in *BRCA1* (left panel) and *BRCA2* (right panel) pathogenic variant carriers for genotyped and imputed variants across the *BRCA1* locus. The genomic position (hg19) of genes is depicted on the X-axis, and the position of rs5820483 is highlighted. The association threshold of $p = 0.001$ is shown by a horizontal dashed line. (b) *BRCA1* genomic locus (hg19) centered on exon 11 and showing the location of rs5820483. The graph also shows the values of H3K36m3 ENCODE cell profiles and the level of sequence conservation relative to the mouse genome. The ENCODE cell types are detailed in <http://genome.ucsc.edu/ENCODE/cellTypes.html>

defined thresholds of $>60\times$ sequence coverage and >10 samples per genotype, we were able to compare isoforms in dupAGG/AGG heterozygotes and AGG homozygotes ($n = 16$ and 28 , respectively). The dupAGG heterozygous lymphoblastoid cell lines showed significant (Mann–Whitney $p < .05$) overexpression of $\Delta 11$ and $\Delta 11q$ isoforms, and underexpression of *BRCA1* full-length isoform (Figure 4c).

The above results indicate that the rs582048 dupAGG allele is associated with exon 11 splicing imbalance in a *BRCA1* wild-type background. Thus, modification of breast cancer risk by this allele

might be caused by relative overexpression of $\Delta 11$ and $\Delta 11q$ isoforms, underexpression of full-length *BRCA1*, and/or alteration of the balance between isoforms and full-length. To assess these possibilities, the association between rs5820483 and breast cancer risk was evaluated in carriers stratified by the position of the *BRCA1* pathogenic variant relative to exon 11q. Assuming there is no allelic expression imbalance, $\Delta 11q$ and $\Delta 11$ expression is expected to be unaffected by dupAGG in carriers of *BRCA1* pathogenic variants inside exon 11q (Figure S5). However, in these carriers, we found that

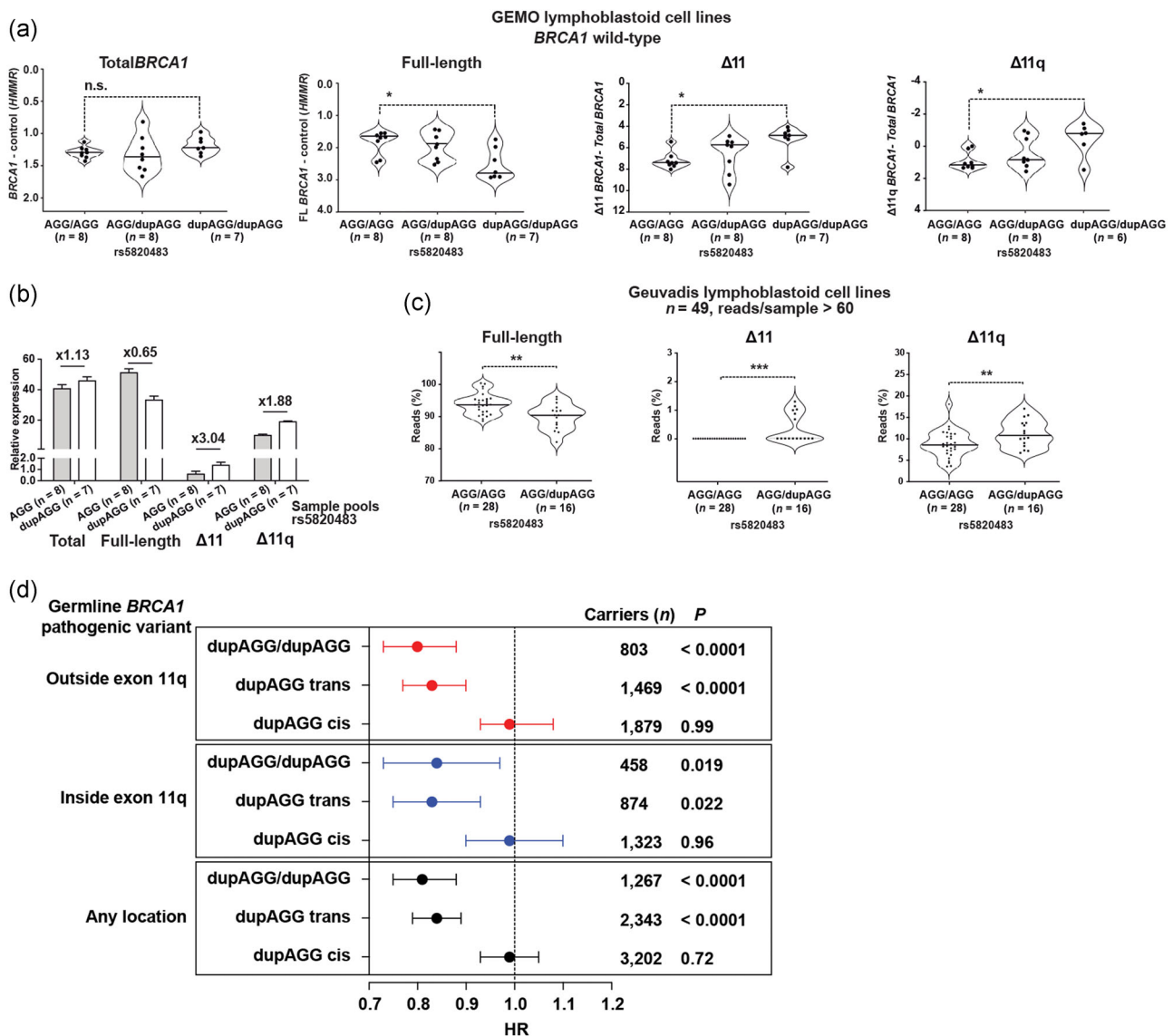


FIGURE 4 rs582043 is associated with exon 11 splicing imbalance. (a) Differences of expression of *BRCA1* full-length, $\Delta 11$, and $\Delta 11q$ isoforms between GEMO lymphoblastoid cell lines by rs582043 genotype (AGG/AGG, AGG/dupAGG, and dupAGG/dupAGG). The significance of the Mann–Whitney test is indicated and the number of samples assayed in each genotype is given. Nonsignificant differences were observed for total *BRCA1* expression. (b) Quantification of isoform differences using digital PCR assays with pooled samples of each genotype in GEMO lymphoblastoid cell lines. (c) Differences of exon 11 isoform expression between Geuvadis lymphoblastoid cell lines according to rs582043 genotypes (AGG/AGG; AGG/dupAGG). The significance of Mann–Whitney tests is indicated and the number of samples assayed for each genotype is shown. (d) Forest plot showing rs582043 HRs and 95% confidence intervals stratified by dupAGG in cis or trans, relative to the *BRCA1* germline pathogenic variant, and dupAGG homozygous, and considering the position of the mutation relative to exon 11q. The number (n) of dupAGG homozygous and heterozygous cis/trans, and the Cox regression likelihood p values are shown. The numbers of non-carriers of dupAGG were: 4388 pathological variants outside exon 11; 2663 pathological variants inside exon 11; and 7156 mutations in any location. HR, hazard ratio; PCR, polymerase chain reaction

dupAGG was still associated with breast cancer risk when located in trans, relative to the *BRCA1* pathogenic variants (Figure 4d). Therefore, relative overexpression of $\Delta 11$ and $\Delta 11q$ isoforms mediated by dupAGG is not the only cause of breast cancer risk modification in *BRCA1* pathogenic variant carriers.

Precomputed PSI values of TCGA breast cancer (Kahles et al., 2018) were analyzed to further evaluate the link between *BRCA1* exon 11 isoform imbalance and carcinogenesis. The tumors were categorized by the observed genotype of rs799916, located in intron 11 (rs5820483 LD $D = 0.99$ and $r^2 = .91$), and assessed for differences in exon 11 PSI. The $\Delta 11$ isoform was not detected in this data set, but there was an indication of an association between rs799916 and PSI of $\Delta 11q$ in basal-like tumors (Kruskal–Wallis test $p = .041$), but not in the other subtypes (Figure S6). A pan-cancer analysis did not show significant results ($n = 1605$, $p = .82$), but examination of tumors mutated in *BRCA1* suggested a similar association to that of basal-like cases, however, only nine tumors were included in this analysis (eight breast cancers and one melanoma; Kruskal–Wallis test $p = .028$; Figure S6).

3.5 | rs5820483 alters splicing in vitro and binds hnRNP A1

To assess the causal effect of rs5820483, we cloned the corresponding alleles within a *BRCA1* minigene construct encompassing exons 10 to 12 (Anczuków et al., 2008) (Figure 5a). Transient

transfection of these constructs in two cancer cell lines, MCF7 and HeLa, coupled to semi-qRT-PCR assays showed that the dupAGG allele significantly reduced inclusion of exon 11, with decreased levels of full-length *BRCA1* and increased levels of the $\Delta 11$ isoform (assays $n = 3$; Figure 5b).

Having corroborated the effect of rs5820483 on exon 11 splicing in vitro, we next evaluated the binding of predicted splicing silencers to the dupAGG allele. Applying predictions from SpliceAid 2 with standard parameters (splicing factors database 2013; Piva et al., 2012), this allele might generate sites for the splicing silencers hnRNP H1/2 and A1 (Figure 5c). In oligo-based pull-down assays, hnRNP H was found to bind similarly to both rs5820483 alleles; however, hnRNP A1 bound approximately twice as strongly to dupAGG (assays $n = 3$; Figure 5d). Analysis of public data from crosslinking and immunoprecipitation assays in three cancer cell lines identified hnRNP A1 binding sites close (<50 bp) to rs5820483 in all settings, and there was overlap between hnRNP A1 binding and rs5820483 location in K562 cells (ENCODE Project Consortium, 2012) (Figure S7).

4 | DISCUSSION

Exon 11 of *BRCA1* codes for signals mediating *BRCA1* nuclear localization (Huen et al., 2010). This exon shows differential splicing between normal and cancer cells, and between phases of the cell cycle (Orban & Olah, 2001). Results from mouse studies with targeted

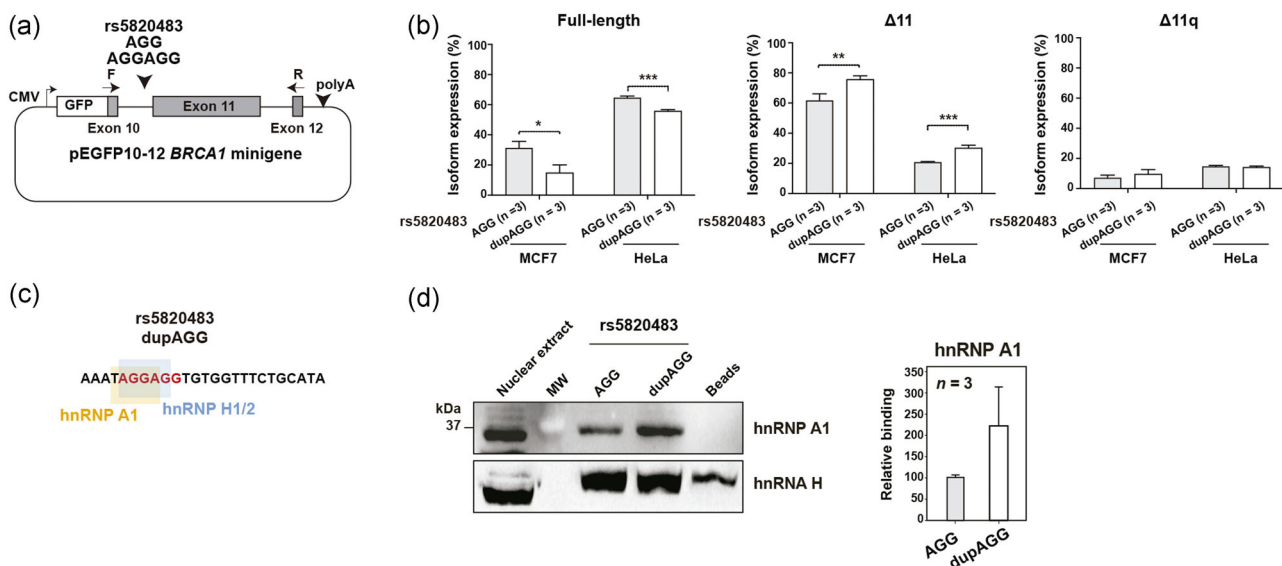


FIGURE 5 rs5820483 alters splicing in vitro and binds hnRNP A1. (a) pEGFP10-12 *BRCA1* minigene and rs582043 allele cloning. Three independent transfections, triplicate qRT-PCRs (reactions of 25 cycles) of each, and triplicate samples of each qRT-PCR were analyzed by capillary electrophoresis. (b) The presence of the dupAGG allele reduces inclusion of exon 11 in minigene assays using MCF7 and HeLa cells. Quantification of differences between the two cell models of expression of *BRCA1* including exon 11, $\Delta 11$, and $\Delta 11q$ following transfection of minigenes carrying the alleles. The significance of Mann–Whitney tests is indicated (* $p < .05$, ** $p < .01$, and *** $p < .001$). (c) rs582043 minor allele (dupAGG, red font) and flanking sequence, and predicted binding sites of exon splicing silencers. (d) Left panel, western blot results of hnRNP A1 and H binding to oligos for rs582043 alleles in pull-down assays. Right panel shows the quantification of three independent assays for hnRNP A1. qRT-PCR, quantitative reverse-transcription polymerase chain reaction

alteration of exon 11-centered isoforms demonstrate perturbations in key pathways and processes, with tumor development following $\Delta 11$ deletion in the context of *Trp53* deletion (S. S. Kim et al., 2006). The product of the $\Delta 11$ isoform can be translocated into the nucleus, but the DNA damage response is partially impaired if solely this isoform is expressed (Huber et al., 2001). However, it remains unclear whether altered levels of exon 11 splicing isoforms are associated with human cancer risk. Here we identify molecular and genetic factors that influence exon 11 splicing and are associated with modification of breast cancer risk in *BRCA1* pathogenic variant carriers.

We have identified a region in intron 10 that contributes to the regulation of exon 11-skipping in humans and mice. This region contains the rs5820483 variant in humans and is modified through a *loxP* insertion in *Brca1*^{*loxP/loxP*} mice. The presence of the minor allele of rs5820483, dupAGG, creates a binding site for the splicing silencer hnRNP A1 and, accordingly, is associated with higher levels of $\Delta 11$ and $\Delta 11q$ isoforms in human cell lines. In vitro assays using minigene constructs and two human cancer cell models confirm the dupAGG effect. However, these conclusions are limited by the cell models used in the assays. *BRCA1* is ubiquitously expressed in normal proliferating cells, and the genomic regulatory marks of the *BRCA1* locus appear to be similar among different cell types; however, we cannot rule out the possible existence of cancer cell-of-origin specificities and/or differential responses to defined conditions, such as exposure to hormones.

It is unlikely that the identified allele and regulatory region function in isolation and, therefore, unidentified variants and elements might cooperate to determine exon 11 inclusion/exclusion and/or isoforms. Indeed, the high LD throughout the entire *BRCA1* locus provides many other genetic variants as additional candidates influencing isoform expression and/or cancer risk. These could act as single modifier factors and/or interact with each other. Relevant variants could include changes in the promoter sequence, as previously reported (Cox et al., 2011), but we did not observe any substantial differences in overall *BRCA1* expression in relation to the rs5820483 genotypes. A comprehensive functional screen of all genetic variations in the *BRCA1* locus, similar to that performed for coding pathological variants (Findlay et al., 2018), may be warranted to accurately identify the impact of any difference in splicing regulation and/or cancer risk in *BRCA1* pathological variant carriers.

It has been previously shown that a region spanning approximately 250 bp from the beginning of human *BRCA1* exon 11 includes multiple splicing regulatory signals (Raponi et al., 2012). In this region, the synonymous c.693G>A variation reduced $\Delta 11q$ but increased $\Delta 11$, which showed allele-specific expression (Tammaro et al., 2012). In parallel, it was proposed that the splice site change c.IVS11+1G>A affected the expression of the *BRCA1* full-length isoform, but not of the $\Delta 11q$ isoform (Bonatti et al., 2006). It is plausible that the identified sequence in intron 10 interacts with several elements and/or variations to determine the outcome of exon 11 splicing. The effect on risk modification may be more complex than anticipated. The association between the dupAGG allele of rs582048 and reduced risk centers on the wild-type *BRCA1* allele; however, pondering the

position of *BRCA1* mutations inside or outside of the exon 11q region, the specific contribution of reduced full-length and/or increased $\Delta 11$ and/or increased $\Delta 11q$ isoform expression to risk modification remains to be quantified. Since the isoforms are regulatorily linked, precisely targeted studies may be needed to decipher specific isoform effects. Multifactorial regulation may also occur in mouse *Brca1* exon 11 splicing. Contrary to what occurs with rs5820483, the insertion of a *loxP* sequence in mouse intron 10 enhances exon 11 inclusion and thereby increases the expression of the full-length *Brca1* isoform. In parallel, the original strains of the *Brca1*^{*loxP/loxP*} mouse model, C57BL/6J and 129/svJ, display a genetic difference (rs49102316) in intron 10, 203 bp from the exon 11 acceptor site, that is predicted to disrupt an SRSF6/10 binding site (Figure S8). Thus, additional *Brca1* locus variants and the genetic background could also have influenced the observed phenotypic alterations.

It is not known how alteration of exon 11-centered isoforms could modify breast cancer risk specifically in *BRCA1* pathogenic variant carriers. Analysis of PSI $\Delta 11q$ data suggests a specific effect in basal-like tumors, which are frequently developed in *BRCA1* mutation carriers (Foulkes et al., 2003); however, this analysis was limited by the relatively low coverage of *BRCA1* isoform expression. Protein functions linked to exon 11 include homologous recombination-mediated DNA damage repair and mitotic control (Huen et al., 2010), which may be considered to be general tumor suppressor roles. Thus, the luminal progenitor cell population that has been linked to the origin of *BRCA1*-related breast cancers may depend differentially on these processes to maintain genomic integrity (Fu et al., 2020; Tharmapalan et al., 2019). Alternatively, or complementarily, isoform imbalance may expand the cancer-prone cell population. Previous observations in mouse models, including developmental abnormalities and cancer susceptibility with the $\Delta 11$ deletion (S. S. Kim et al., 2006), and our indications of mammary morphological differences, suggest that alterations in cell biology arise from the changes in the balance of *BRCA1* exon 11 isoforms and full-length expression. These alterations might precede loss of the *BRCA1* wild-type allele (i.e., second hit) according to a proposed model of events causing *BRCA1*-associated breast cancer (Martins et al., 2012). Further studies in mouse and human cells and tissue may be needed to assess the molecular and cellular impacts of the imbalance in exon 11 isoforms.

ACKNOWLEDGMENTS

The study of the *BRCA1* wild-type allele was first proposed and led by Dr. Olga Sinilnikova in CIMBA. We dedicate this study to her memory. The paper is also dedicated to Dr. David L. Kleinberg, who contributed to the study of the *BRCA1* mouse model. The authors acknowledge Dr. Weifeng Ruan for his contributions to early studies related to this project, the Experimental Pathology Core Laboratory at NYU School of Medicine for tissue sectioning and staining, the VA Medical Center Animal Facility, and Ms. Clara Montesino for help with the animal experiments. The results are partly based upon data generated by the TCGA Research Network (<https://www.cancer.gov/tcga>), and we express our gratitude to the TCGA consortium and coordinators for producing the data and clinical information used in

our study. The ICO-IDIBELL research was supported by the Generalitat de Catalunya (SGR 2017-449, 2017-899 and 2017-1282; PERIS MedPerCan and PFI-Salut SLT017-20-000076; URDCat; and CERCA programme for IDIBELL institutional support), Carlos III Institute of Health (ISCIII), funded by FEDER funds—a way to build Europe—(Ministry of Science, Innovation, and Universities; grants PI16/00563, PI18/01029, PI19/00553, and PI21/01306; CIBERONC and CIBER-BBN) and Ministry of Economy and Competitiveness-MINECO (grants RTI2018-095377-B-100 and RD16/0011/0024). G.R.G. was supported by an IDIBELL post-residency fellowship. The work at NYU was supported by DOD Breast Cancer Research Program funding to D.L.K. (W81XWH-11-1-0779) and M.H.B.-H. (W81XWH-11-1-780), and a Novartis Investigator-Initiated Grant to D.L.K. (CSOM230BUS03). The Baralle laboratory is supported by D.B.'s NIH Research Professorship (RP-2016-07-011). A.B.S. is supported by an NHMRC Investigator Fellowship (APP: 1177524) and P.R. is supported by funds from the Italian Association for Cancer Research (AIRC; IG number 22093).

CONFLICT OF INTERESTS

M.A.P. was the recipient of an unrestricted research grant from Roche Pharma for the development of the ProCURE ICO research program. X.S.P. is a cofounder of and has an ownership interest (including stock and patents) in DREAMgenics.

DATA AVAILABILITY STATEMENT

Summary-level statistics are available from <http://cimba.ccgce.medschl.cam.ac.uk/projects/>. Requests for individual data can be made to the Data Access Coordination Committee (DACC) of CIMBA.

WEB RESOURCES

cBioPortal: <https://www.cbioportal.org>

ENCODE cell types: <http://genome.ucsc.edu/ENCODE/cellTypes>

FeatureCounts algorithm: <http://bioinf.wehi.edu.au/featureCounts>

GWAS summary statistics: <http://cimba.ccgce.medschl.cam.ac.uk/projects>

STAR algorithm: <https://github.com/alexdobin/STAR>

TCGA genotype data: <https://gdc.cancer.gov/access-data>

TCGA PSI values: <https://gdc.cancer.gov/about-data/publications/>

[PanCanAtlas-Splicing-2018](#)

ORCID

Gorka Ruiz de Garibay  <http://orcid.org/0000-0001-9936-8419>

Sylvie Mazoyer  <http://orcid.org/0000-0002-2135-0160>

Amanda B. Spurdle  <http://orcid.org/0000-0003-1337-7897>

Miguel de la Hoya  <http://orcid.org/0000-0002-8113-1410>

Miquel A. Pujana  <http://orcid.org/0000-0003-3222-4044>

REFERENCES

Anczuków, O., Buisson, M., Salles, M.-J., Triboulet, S., Longy, M., Lidereau, R., Sinilnikova, O. M., & Mazoyer, S. (2008). Unclassified variants identified in BRCA1 exon 11: Consequences on splicing. *Genes Chromosomes Cancer*, 47, 418–426.

- Anderson, S. M., Rudolph, M. C., McManaman, J. L., & Neville, M. C. (2007). Key stages in mammary gland development. Secretory activation in the mammary gland: It's not just about milk protein synthesis! *Breast Cancer Research*, 9, 204.
- Antoniou, A. C., Sinilnikova, O. M., Simard, J., Léoné, M., Dumont, M., Neuhausen, S. L., Struwing, J. P., Stoppa-Lyonnet, D., Barjhoux, L., Hughes, D. J., Coupier, I., Belotti, M., Lasset, C., Bonadona, V., Bignon, Y. J., Genetic Modifiers of Cancer Risk in/Mutation Carriers Study, Rebbeck, T. R., Wagner, T., Lynch, H. T., ... Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) (2007). RAD51 135G->C modifies breast cancer risk among BRCA2 mutation carriers: Results from a combined analysis of 19 studies. *American Journal of Human Genetics*, 81, 1186–1200.
- Antoniou, A. C., Wang, X., Fredericksen, Z. S., McGuffog, L., Tarrell, R., Sinilnikova, O. M., Healey, S., Morrison, J., Kartsonaki, C., Lesnick, T., Ghousaini, M., Barrowdale, D., EMBRACE, Peock, S., Cook, M., Oliver, C., Frost, D., Eccles, D., Evans, D. G., ... Flesch-Janys, D. (2010). A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nature Genetics*, 42, 885–892.
- Bachelier, R., Vincent, A., Mathevet, P., Magdinier, F., Lenoir, G. M., & Frappart, L. (2002). Retroviral transduction of splice variant Brca1-Delta11 or mutant Brca1-W1777Stop causes mouse epithelial mammary atypical duct hyperplasia. *Virchows Archiv*, 440, 261–266.
- Bai, L., Yang, H. H., Hu, Y., Shukla, A., Ha, N. -H., Doran, A., Faraji, F., Goldberger, N., Lee, M. P., Keane, T., & Hunter, K. W. (2016). An integrated genome-wide systems genetics screen for breast cancer metastasis susceptibility genes. *PLOS Genetics*, 12, e1005989.
- Barnes, D. R., Rookus, M. A., McGuffog, L., Leslie, G., Mooij, T. M., Dennis, J., Mavaddat, N., Adlard, J., Ahmed, M., Aittomäki, K., Andrieu, N., Andrulis, I. L., Arnold, N., Arun, B. K., Azzollini, J., Balmaña, J., Barkardottir, R. B., Barrowdale, D., Benitez, J., ... Rantala, J. (2020). Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genetics in Medicine*, 22, 1653–1666.
- Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21, 263–265.
- Bonatti, F., Pepe, C., Tancredi, M., Lombardi, G., Aretini, P., Sensi, E., Falaschi, E., Cipollini, G., Bevilacqua, G., & Caligo, M. A. (2006). RNA-based analysis of BRCA1 and BRCA2 gene alterations. *Cancer Genetics and Cytogenetics*, 170, 93–101.
- Cao, L., Li, W., Kim, S., Brodie, S. G., & Deng, C. -X. (2003). Senescence, aging, and malignant transformation mediated by p53 in mice lacking the Brca1 full-length isoform. *Genes & Development*, 17, 201–213.
- Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q., & Krainer, A. R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Research*, 31, 3568–3571.
- Chenevix-Trench, G., Milne, R. L., Antoniou, A. C., Couch, F. J., Easton, D. F., & Goldgar, D. E., CIMBA. (2007). An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: The Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). *Breast Cancer Research*, 9, 104.
- Colombo, M., Blok, M. J., Whaley, P., Santamariña, M., Gutiérrez-Enríquez, S., Romero, A., Garre, P., Becker, A., Smith, L. D., De Vecchi, G., Brandão, R. D., Tserpelis, D., Brown, M., Blanco, A., Bonache, S., Menéndez, M., Houdayer, C., Foglia, C., Fackenthal, J. D., ... De La Hoya, M. (2014). Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: A report from the ENIGMA consortium. *Human Molecular Genetics*, 23, 3666–3680.
- Couch, F. J., Wang, X., McGuffog, L., Lee, A., Olsword, C., Kuchenbaecker, K. B., Soucy, P., Fredericksen, Z., Barrowdale, D.,

- Dennis, J., Gaudet, M. M., Dicks, E., Kosel, M., Healey, S., Sinilnikova, O. M., Lee, A., Bacot, F., Vincent, D., Hogervorst, F. B., ... Varesco, L. (2013). Genome-wide association study in *BRCA1* mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLOS Genetics*, *9*, e1003212.
- Cox, D. G., Kraft, P., Hankinson, S. E., & Hunter, D. J. (2005). Haplotype analysis of common variants in the *BRCA1* gene and risk of sporadic breast cancer. *Breast Cancer Research*, *7*, R171–R175.
- Cox, D. G., Simard, J., Sinnett, D., Hamdi, Y., Soucy, P., Ouimet, M., Barjhoux, L., Verny-Pierre, C., McGuffog, L., Healey, S., Szabo, C., Greene, M. H., Mai, P. L., Andrulis, I. L., Ontario Cancer Genetics, N., Thomassen, M., Gerdes, A. M., Caligo, M. A., Friedman, E., ... Consortium of Investigators of Modifiers of *BRCA1/2* (2011). Common variants of the *BRCA1* wild-type allele modify the risk of breast cancer in *BRCA1* mutation carriers. *Human Molecular Genetics*, *20*, 4732–4747.
- Dunning, A. M., Chiano, M., Smith, N. R., Dearden, J., Gore, M., Oakes, S., Wilson, C., Stratton, M., Peto, J., Easton, D., Clayton, D., & Ponder, B. A. (1997). Common *BRCA1* variants and susceptibility to breast and ovarian cancer in the general population. *Human Molecular Genetics*, *6*, 285–289.
- ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*, 57–74.
- Findlay, G. M., Daza, R. M., Martin, B., Zhang, M. D., Leith, A. P., Gasperini, M., Janizek, J. D., Huang, X., Starita, L. M., & Shendure, J. (2018). Accurate classification of *BRCA1* variants with saturation genome editing. *Nature*, *562*, 217–222.
- Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D. T., Weber, B., Lenoir, G., Chang-Claude, J., Sobol, H., Teare, M. D., Struwing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T. R., Tonin, P., Neuhausen, S., ... Zelada-Hedman, M. (1998). Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. The Breast Cancer Linkage Consortium. *American Journal of Human Genetics*, *62*, 676–689.
- Foulkes, W. D., Stefansson, I. M., Chappuis, P. O., Bégin, L. R., Goffin, J. R., Wong, N., Trudel, M., & Akslen, L. A. (2003). Germline *BRCA1* mutations and a basal epithelial phenotype in breast cancer. *Journal of the National Cancer Institute*, *95*, 1482–1485.
- Freedman, M. L., Penney, K. L., Stram, D. O., Riley, S., McKean-Cowdin, R., Le Marchand, L., Altshuler, D., & Haiman, C. A. (2005). A haplotype-based case-control study of *BRCA1* and sporadic breast cancer risk. *Cancer Research*, *65*, 7516–7522.
- Fu, N. Y., Nolan, E., Lindeman, G. J., & Visvader, J. E. (2020). Stem cells and the differentiation hierarchy in mammary gland development. *Physiological Reviews*, *100*, 489–523.
- Gaudet, M. M., Kuchenbaecker, K. B., Vijai, J., Klein, R. J., Kirchhoff, T., McGuffog, L., Barrowdale, D., Dunning, A. M., Lee, A., Dennis, J., Healey, S., Dicks, E., Soucy, P., Sinilnikova, O. M., Pankratz, V. S., Wang, X., Eldridge, R. C., Tessier, D. C., Vincent, D., ... Offit, K. (2013). Identification of a *BRCA2*-specific modifier locus at 6p24 related to breast cancer risk. *PLOS Genetics*, *9*, e1003173.
- Howell, A., Anderson, A. S., Clarke, R. B., Duffy, S. W., Evans, D. G., Garcia-Closas, M., Gescher, A. J., Key, T. J., Saxton, J. M., & Harvie, M. N. (2014). Risk determination and prevention of breast cancer. *Breast Cancer Research*, *16*, 446.
- de la Hoya, M., Soukariéh, O., López-Perolio, I., Vega, A., Walker, L. C., van Ierland, Y., Baralle, D., Santamariña, M., Lattimore, V., Wijnen, J., Whaley, P., Blanco, A., Raponi, M., Hauke, J., Wappenschmidt, B., Becker, A., Hansen, T. V., Behar, R., kConFab Investigators, ... Spurdle, A. B. (2016). Combined genetic and splicing analysis of *BRCA1* c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Human Molecular Genetics*, *25*, 2256–2268.
- Huber, L. J., & Chodosh, L. A. (2005). Dynamics of DNA repair suggested by the subcellular localization of *Brca1* and *Brca2* proteins. *Journal of Cellular Biochemistry*, *96*, 47–55.
- Huber, L. J., Yang, T. W., Sarkisian, C. J., Master, S. R., Deng, C. X., & Chodosh, L. A. (2001). Impaired DNA damage response in cells expressing an exon 11-deleted murine *Brca1* variant that localizes to nuclear foci. *Molecular and Cellular Biology*, *21*, 4005–4015.
- Huen, M. S. Y., Sy, S. M. H., & Chen, J. (2010). *BRCA1* and its toolbox for the maintenance of genome integrity. *Nature Reviews Molecular Cell Biology*, *11*, 138–148.
- Illa-Bochaca, I., Fernandez-Gonzalez, R., Shelton, D. N., Welm, B. E., Ortiz-de-Solorzano, C., & Barcellos-Hoff, M. H. (2010). Limiting-dilution transplantation assays in mammary stem cell studies. *Methods in Molecular Biology*, *621*, 29–47.
- de Jong, L. C., de Cree, S., Lattimore, V., Wiggins, G. A. R., Spurdle, A. B., kConFab Investigators, Miller, A., Kennedy, M. A., & Walker, L. C. (2017). Nanopore sequencing of full-length *BRCA1* mRNA transcripts reveals co-occurrence of known exon skipping events. *Breast Cancer Research*, *19*, 127.
- Kahles, A., Lehmann, K.-V., Toussaint, N. C., Hüser, M., Stark, S. G., Sachsenberg, T., Stegle, O., Kohlbacher, O., Sander, C., Cancer Genome Atlas Research Network, & Ratsch, G. (2018). Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell*, *34*, 211–224.
- Kim, E., Ilagan, J. O., Liang, Y., Daubner, G. M., Lee, S. C., Ramakrishnan, A., Li, Y., Chung, Y. R., Micol, J.-B., Murphy, M. E., Cho, H., Kim, M.-K., Zebari, A. S., Aumann, S., Park, C. Y., Buonamici, S., Smith, P. G., Deeg, H. J., Lobry, C., ... Abdel-Wahab, O. (2015). *SRSF2* mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell*, *27*, 617–630.
- Kim, S. S., Cao, L., Lim, S.-C., Li, C., Wang, R.-H., Xu, X., Bachelier, R., & Deng, C.-X. (2006). Hyperplasia and spontaneous tumor development in the gynecologic system in mice lacking the *BRCA1-Delta11* isoform. *Molecular and Cellular Biology*, *26*, 6983–6992.
- Kraya, A. A., Maxwell, K. N., Wubbenhorst, B., Wenz, B. M., Pluta, J., Rech, A. J., Dorfman, L. M., Lunceford, N., Barrett, A., Mitra, N., Morrisette, J. J. D., Feldman, M., Nayak, A., Domchek, S. M., Vonderheide, R. H., & Nathanson, K. L. (2019). Genomic signatures predict the immunogenicity of *BRCA*-deficient breast cancer. *Clinical Cancer Research*, *25*, 4363–4374.
- Kuchenbaecker, K. B., McGuffog, L., Barrowdale, D., Lee, A., Soucy, P., Dennis, J., Domchek, S. M., Robson, M., Spurdle, A. B., Ramus, S. J., Mavaddat, N., Terry, M. B., Neuhausen, S. L., Schmutzler, R. K., Simard, J., Pharoah, P., Offit, K., Couch, F. J., Chenevix-Trench, G., ... Antoniou, A. C. (2017). Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in *BRCA1* and *BRCA2* mutation carriers. *Journal of the National Cancer Institute*, *109*, djw302.
- Lappalainen, T., Sammeth, M., Friedländer, M. R., 't Hoen, P. A., Monlong, J., Rivas, M. A., González-Porta, M., Kurbatova, N., Griebel, T., Ferreira, P. G., Barann, M., Wieland, T., Greger, L., van Iterson, M., Almlöf, J., Ribeca, P., Pulyakhina, I., Esser, D., Giger, T., ... Dermitzakis, E. T. (2013). Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*, *501*, 506–511.
- Lesueur, F., Mebirouk, N., Mebirouk, N., Jiao, Y., Barjhoux, L., Belotti, M., Laurent, M., Léone, M., Houdayer, C., Bressac-de Paillerets, B., Vaur, D., Sobol, H., Nogués, C., Longy, M., Mortemousque, I., Fert-Ferrer, S., Mouret-Fourme, E., Pujol, P., Venat-Bouvet, L., Bignon, Y. J., ... GEMO Study, C. (2018). GEMO, a national resource to study genetic modifiers of breast and ovarian cancer risk in *BRCA1* and *BRCA2* pathogenic variant carriers. *Frontiers in Oncology*, *8*, 490.
- Li, D., Harlan-Williams, L. M., Kumaraswamy, E., & Jensen, R. A. (2019). *BRCA1* - No matter how you splice it. *Cancer Research*, *79*, 2091–2098.

- Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., & Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science*, 327, 996–1000.
- Martins, F. C., De, S., Almendro, V., Gönen, M., Park, S. Y., Blum, J. L., Herlihy, W., Ethington, G., Schnitt, S. J., Tung, N., Garber, J. E., Fettes, K., Michor, F., & Polyak, K. (2012). Evolutionary pathways in BRCA1-associated breast tumors. *Cancer Discovery*, 2, 503–511.
- Meier-Abt, F., Milani, E., Roloff, T., Brinkhaus, H., Duss, S., Meyer, D. S., Klebba, I., Balwiercz, P. J., Nimwegen, E., & van, Bentires-Alj, M. (2013). Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. *Breast Cancer Research*, 15, R36.
- Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemaçon, A., Soucy, P., Glubb, D., Rostamianfar, A., Bolla, M. K., Wang, Q., Tyrer, J., Dicks, E., Lee, A., Wang, Z., Allen, J., Keeman, R., Eilber, U., ... Humphreys, K. (2017). Association analysis identifies 65 new breast cancer risk loci. *Nature*, 551, 92–94.
- Milne, R. L., & Antoniou, A. C. (2016). Modifiers of breast and ovarian cancer risks for BRCA1 and BRCA2 mutation carriers. *Endocrine-Related Cancer*, 23, T69–T84.
- Milne, R. L., Kuchenbaecker, K. B., Michailidou, K., Beesley, J., Kar, S., Lindström, S., Hui, S., Lemaçon, A., Soucy, P., Dennis, J., Jiang, X., Rostamianfar, A., Finucane, H., Bolla, M. K., McGuffog, L., Wang, Q., Aalfs, C. M., ABCTB, I., Adams, M., ... Flyger, H. (2017). Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nature Genetics*, 49, 1767–1778.
- Orban, T. I., & Olah, E. (2001). Expression profiles of BRCA1 splice variants in asynchronous and in G1/S synchronized tumor cell lines. *Biochemical and Biophysical Research Communications*, 280, 32–38.
- Piva, F., Giulietti, M., Burini, A. B., & Principato, G. (2012). SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. *Human Mutation*, 33, 81–85.
- Raponi, M., Douglas, A. G. L., Tamaro, C., Wilson, D. I., & Baralle, D. (2012). Evolutionary constraint helps unmask a splicing regulatory region in BRCA1 exon 11. *PLOS One*, 7, e37255.
- Raponi, M., Smith, L. D., Silipo, M., Stuan, C., Buratti, E., & Baralle, D. (2014). BRCA1 exon 11 a model of long exon splicing regulation. *RNA Biology*, 11, 351–359.
- Rebbeck, T. R., Mitra, N., Wan, F., Sinilnikova, O. M., Healey, S., McGuffog, L., Mazoyer, S., Chenevix-Trench, G., Easton, D. F., Antoniou, A. C., Nathanson, K. L., Cimba, C., Laitman, Y., Kushnir, A., Paluch-Shimon, S., Berger, R., Zidan, J., Friedman, E., Ehrencrona, H., ... Muller, D. (2015). Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *JAMA*, 313, 1347–1361.
- Romero, A., García-García, F., López-Perolio, I., Ruiz de Garibay, G., García-Sáenz, J. A., Garre, P., Ayllón, P., Benito, E., Dopazo, J., Díaz-Rubio, E., Caldés, T., & Miguel de la, H. (2015). BRCA1 alternative splicing landscape in breast tissue samples. *BMC Cancer*, 15, 219.
- Tamaro, C., Raponi, M., Wilson, D. I., & Baralle, D. (2012). BRCA1 exon 11 alternative splicing, multiple functions and the association with cancer. *Biochemical Society Transactions*, 40, 768–772.
- Tamaro, C., Raponi, M., Wilson, D. I., & Baralle, D. (2014). BRCA1 EXON 11, a CERES (composite regulatory element of splicing) element involved in splice regulation. *International Journal of Molecular Sciences*, 15, 13045–13059.
- Tharmapalan, P., Mahendralingam, M., Berman, H. K., & Khokha, R. (2019). Mammary stem cells and progenitors: Targeting the roots of breast cancer for prevention. *EMBO Journal*, 38, e100852.
- Tournier, I., Vezain, M., Martins, A., Charbonnier, F., Baert-Desurmont, S., Olschwang, S., Wang, Q., Buisine, M. P., Soret, J., Tazi, J., Frébourg, T., & Tosi, M. (2008). A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. *Human Mutation*, 29, 1412–1424.
- Wang, R.-H., Yu, H., & Deng, C.-X. (2004). A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 17108–17113.
- Wilson, C. A., Payton, M. N., Elliott, G. S., Buaas, F. W., Cajulis, E. E., Grosshans, D., Ramos, L., Reese, D. M., Slamon, D. J., & Calzone, F. J. (1997). Differential subcellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-delta11b. *Oncogene*, 14, 1–16.
- Xu, X., Qiao, W., Linke, S. P., Cao, L., Li, W. M., Furth, P. A., Harris, C. C., & Deng, C. X. (2001). Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis. *Nature Genetics*, 28, 266–271.
- Xu, X., Wagner, K. U., Larson, D., Weaver, Z., Li, C., Ried, T., Hennighausen, L., Wynshaw-Boris, A., & Deng, C. X. (1999). Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. *Nature Genetics*, 22, 37–43.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ruiz de Garibay, G., Fernandez-García, I., Mazoyer, S., Leme de Calais, F., Ameri, P., Vijayakumar, S., Martinez-Ruiz, H., Damiola, F., Barjhoux, L., Thomassen, M., Andersen, L. v. B., Herranz, C., Mateo, F., Palomero, L., Espín, R., Gómez, A., García, N., Jimenez, D., Bonifaci, N., Extremera, A. I., ... Pujana, M. A. (2021). Altered regulation of BRCA1 exon 11 splicing is associated with breast cancer risk in carriers of BRCA1 pathogenic variants. *Human Mutation*, 42, 1488–1502.

<https://doi.org/10.1002/humu.24276>