IMPROVING THE RISK ASSESSMENT OF INHERITED BREAST AND OVARIAN CANCER: CLINICAL SIGNIFICANCE OF *BRCA1/2* VARIANTS AND RISK MODIFIERS

Rita Brandão

IMPROVING THE RISK ASSESSMENT OF INHERITED BREAST AND OVARIAN CANCER: CLINICAL SIGNIFICANCE OF *BRCA1/2* VARIANTS AND RISK MODIFIERS

Rita Dias Brandão

Improving the risk assessment of inherited breast and ovarian cancer: clinical significance of *BRCA1/2* variants and risk modifiers
ISBN: 978-9-08-891486-7

Cover design: Oksana Grivina/Mariana Madeira
Lay-out: Rita Dias Brandão
Printed by: Proefschriftmaken.nl || Uitgeverij BOXPress

IMPROVING THE RISK ASSESSMENT OF INHERITED BREAST AND OVARIAN CANCER: CLINICAL SIGNIFICANCE OF *BRCA1/2* VARIANTS AND RISK MODIFIERS

DISSERTATION

To obtain the degree of Doctor
at the Maastricht University,
on the authority of the Rector Magnificus,
Prof dr. L.L.G. Soete
in accordance with the decision of the Board of Deans,
to be defended in public

on Wednesday 31st October 2012, at 12 hours
by

Rita Dias Brandão

Born on 15th November 1980, Lisbon (Portugal)

Supervisor:
Prof. dr. Hubert J. Smeets

Co-supervisors:
Dr. ir. Marinus J. Blok
Dr. Encarna B. Goméz García

Assessment Committee:
Prof. dr. Manon van Engeland, chair
Dr. Kathleen Claes, Ghent University Hospital, Belgium
Prof. dr. Peter Devilee, Leiden University Medical Centre, the Netherlands
Dr. Theo de Kok
Prof. dr. Roy F.P.M. Kruitwagen

Rita D. Brandão was supported by Portuguese fellowships: Fundação Calouste Gulbenkian (79117), and Fundação para a Ciência e Tecnologia (SFRH/BD/32386/2006), co-financed by the POPH and European Community Fund. Printing costs were in part defrayed by a stipend from MRC-Holland. Other financial support received for the studies herein are acknowledged in chapters 2 through 6.

Table of Contents

List of Abbreviations

a.a.	amino acid
ANKLE1	ankyrin repeat and LEM domain containing gene 1
AR	androgen receptor gene
AS-PCR	allele-specific polymerase chain reaction
ASS	acceptor splice site
ATM	ataxia telangiectasia mutated gene
ATR	ataxia telangiectasia and Rad3 related gene
BABAM1	BRISC and BRCA1 A complex member gene 1
BAP1	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase) gene
BARD1	BRCA1 associated RING domain 1
BAX	BCL2-associated X gene
BC	breast cancer
BER	base excision repair
BIC	breast cancer information core
BMI	body mass index
BNC2	basonuclin gene 2
BRAF35	BRCA2-associated factor 35
BRCA1	breast and ovarian cancer susceptibility gene 1
BRCA2	breast and ovarian cancer susceptibility gene 2
BRCT	BRCA C-terminus
BRIP1	BRCA1 interacting protein C-terminal helicase gene 1
CASP8	caspase-8 gene
CDH1	cadherin-1 gene
cDNA	complementary DNA
CHEK2	CHK2 checkpoint homolog (S. pombe) gene
CIMBA	The Consortium of Investigators of Modifiers of BRCA1/2
COGS	Collaborative Oncological Gene-environment Study
DDB2	damage-specific DNA binding gene 2
DHX9	DEAH (Asp-Glu-Ala-His) box polypeptide 9
DMC1	dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (yeast)
DNA	Deoxyribonucleic acid

DROSHA	drosha, ribonuclease type III
DSB	double strand break
DSS	donor splice site
DSS1	Deleted in split hand/split foot protein 1
EBV	Epstein-Barr virus
EJC	exon junction complex
ENIGMA	evidence-based network for the interpretation of germline alleles
ER	estrogen receptor
ESE	Exonic Splice Enhancer
ESR	estrogen receptor 1 gene
FANCD2	Fanconi anemia, complementation group D2
FGFR2	fibroblast growth factor receptor 2
GADD45	growth arrest and DNA-damage-inducible
GWAS	Genome-wide Association Studies
H2A	Histone 2A
H2AX	H2A histone family, member X
HBOC	Hereditary breast/ovarian cancer
HER2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
HR	Homologous recombination
HRAS1	v-Ha-ras Harvey rat sarcoma viral oncogene homolog gene
HWE	Hardy-Weinberg equilibrium
IL2	interleukin 2
LCLs	Lymphoblastoid cell lines
LIFR	Leukemia inhibitory factor receptor
LINE	long interspersed nuclear element
STK11	serine/threonine kinase 11 gene, *LKB1*
LOH	loss of heterozygosity
LSP1	lymphocyte-specific protein 1 gene
MAP3K1	mitogen-activated protein kinase kinase kinase 1 gene
MLPA	Multiplex Ligation-dependent Probe Amplification
MMR	mismatch repair
MNT	Max binding gene

MRI	magnetic-resonance imaging
mRNA	messenger RNA
NBS1	Nijmegen breakage syndrome 1 (nibrin)
NER	Nucleotide Excision Repair
NGS	next-generation sequencing
NHEJ	non-homologous end joining
NMD	nonsense-mediated mRNA decay
OC	ivarian cancer
p53	tumor protein p53, TP53
PALB2	partner and localizer of BRCA2
PARP	poly (ADP-ribose) polymerase
PBL	peripheral blood lymphocytes
PBMC	peripheral blood mononuclear cells
PCR	polymerase chain reaction
PHA	Phytohaemagglutinin
PLK1	polo-like kinase 1
PR	progesterone receptor
PTC	premature termination codon
PTEN	phosphatase and tensin homolog gene
PTHLH	parathyroid hormone-like hormone
RAD50	RAD50 homolog (S. cerevisiae)
RAD51	RAD51 homolog (S. cerevisiae)
RAD51C	RAD51 homolog C (S. cerevisiae)
RAD51D	RAD51 homolog D (S. cerevisiae)
RAD51L1	RAD51-like 1 (S. cerevisiae)
RNA	ribonucleic acid
RRM	risk-reducing mastectomy
RRSO	risk-reducing salpingo-oophorectomy
SMAD3	SMAD family member 3
SRE	splice regulatory element
SSB	single strand break
TGFB1	transforming growth factor, beta 1 gene
TNRC9	TOX high mobility group box family member 3 gene, *TOX3*

TP53	tumor protein p53
TR2	nuclear receptor subfamily 2, group C, member 1
VUS	variant of undetermined significance
WT	wild-type
XIST	X (inactive)-specific transcript (non-protein coding)
XPC	xeroderma pigmentosum, complementation group C gene
XRCC2	X-ray repair cross-complementing 2 gene
ZN365	zinc finger protein 365 gene

CHAPTER 1

General Introduction

The first reports on hereditary breast cancer were published by the French surgeons Henri François le Dran and Paul Broca in 1757 and 1866, respectively [1, 2]. Henri F. le Dran describes the history of a nun diagnosed with breast cancer (BC) at age 19, whose grandmother and great-uncle had also developed breast cancer. Initially, she refused operation in the belief that it would not be a solution, since the disease was caused by something bad in her families' blood, i.e. that it was hereditary [1]. Paul Broca described a family with 10 cases of breast cancer in different generations in his book as an example of a hereditary influence to cancer predisposition [2]. Since then many families with hereditary breast/ovarian cancer (HBOC) syndrome were identified. More recently it became clear that hereditary cases account for around 5-10% of all breast and ovarian cancer cases [3, 4]. A major breakthrough came in the 1990's with the identification of the early-onset breast cancer susceptibility genes *BRCA1* and *BRCA2* [5, 6]. This chapter introduces hereditary breast and ovarian cancer due mutations in the *BRCA1* and *BRCA2* genes, the characterization of variants of undetermined significance (VUS) in these genes, with an emphasis in those that might affect splicing, and genetic variants in genes other than *BRCA1* and *BRCA2* that modify the cancer risk in carriers. As indicated in the outline, these topics have been the focus of the work described in this thesis.

Hereditary breast/ovarian cancer syndrome

HBOC is generally characterized by multiple BC and/or OC cases in first- and second- and third-degree relatives at a young age of onset (age < 50 years). It may also include male breast cancer cases, especially associated with *BRCA2*-mutations [7]. Women with a breast cancer under the age of 40 may also be indicated for *BRCA1/2* genetic screening, even if they are the only case in the family. Both *BRCA1*- and *BRCA2*-mutation carriers are predisposed to early-onset of breast and/or ovarian cancer. However, mutations in these genes show variable penetrance. The results of a meta-analysis indicated that at age 70 the risk of BC is 55% (95% CI, 50% to 59%) for *BRCA1* and 47% (95% CI, 42 to 51%) for *BRCA2* mutation carriers. Regarding the risk of OC, it is 39% (95% CI, 34 to 45%) for *BRCA1*, whereas for *BRCA2* mutation carriers it is 17% (95% CI, 13 to 21%) [8]. Moreover, *BRCA2* is more pleiotropic than *BRCA1*, i.e., besides breast and ovarian cancer, the spectrum of malignancies associated with *BRCA2* mutations includes prostate cancer, pancreatic cancer and melanoma [9, 10]. Mutations in the *BRCA1/2* genes account for about sixteen per cent of the families with familial breast/ovarian cancer syndrome, whereas mutations in the other known genes account for a smaller percentage (<10%) of families. Other, less penetrant, genes that account for hereditary breast/ovarian cancer susceptibility include *PALB2*, *BRIP1*, *ATM*, *RAD51C*, *CHEK2*, *RAD50*, and *RAD51D* [9, 11, 12]. The vast majority of the HBOC families probably have mutations in yet unknown genes [9].

Inherited breast cancer syndromes

In addition to HBOC caused by *BRCA1* and *BRCA2* mutations, there are several other inherited cancer syndromes with an increased risk of BC caused by mutations in a variety

of genes, such as Cowden and Bannayan-Riley-Rivalcaba (*PTEN* mutations), Li-Fraumeni (*TP53* mutations), lobular breast cancer with diffuse gastric cancer (*CDH1* mutations), Peutz-Jeghers (*STK11*/*LKB1*), and Nijmegen breakage syndrome (*NBS1*) [13, 14]. All these syndromes have autosomal dominant inheritance with variable penetrance and expressivity. They can be recognized by specific accompanying clinical features. For instance, macrocephaly and/or learning disability/autism in Cowden syndrome; paediatric cancer, sarcomas and brain tumours in Li-Fraumeni syndrome; and "freckles" on the lips, Sertoli cell ovarian tumours and/or hamartomatous polyps in Peutz-Jeghers syndrome [13, 14].

***BRCA1/2* genes, protein structure and functions**

The *BRCA1* gene was identified in 1994 [5] and *BRCA2* in 1995 [6] by positional cloning, using families with multiple cases. *BRCA1* mapped to chromosome 17q21. Its genomic sequence is distributed over 81,188bp (NG_005905.2) and encodes a protein of 1863 amino acids (NP_009225.1). *BRCA1* is a large gene with 24 exons, 22 of which are coding. *BRCA2* is located on chromosome 13q12.3 with 27 exons, 26 of which are coding. The *BRCA2*-genomic sequence is spread over 91,193 bp, encoding a protein of 3418 amino acids (NP_000050.2). Although the *BRCA1* and *BRCA2* genes are not homologous, both have an unusual large exon 11 and a translational start site in exon 2 [5, 15, 16]. Additionally, both genes have a high percentage of repetitive elements. The *BRCA1* region consists of 42% Alu sequences, while in the *BRCA2* region these account for 20%. *BRCA2* also has 27% LINE and MER repetitive DNA sequences [15, 17]. BRCA1 and BRCA2 proteins are normally located in the nucleus and contain phosphorylated residues (reviewed in [17]). The functional domains of the proteins and some of the interacting protein sites are shown in Figure 1.

BRCA1 is more abundant in the cell during the S and G2 phases than in other cell cycle phases [18]. It interacts with several regulatory proteins in the cell nucleus and participates in different biologic processes such as: DNA damage repair, regulation of gene expression, cell cycle control during the S and G2/M checkpoints, chromatin remodelling and ubiquitylation. BRCA1 is phosphorylated by ATM and CHK2 in response to ionizing radiation [19] and by ATR in response to ultraviolet irradiation [20]. ATM phosphorylation of BRCA1 at Ser1423 is essential for the role of BRCA1 in G2/M checkpoint control. Phosphorylated BRCA1 protein is also required for DNA double-strand breaks (DSB) repair by homologous-recombination and contributes to the regulation of expression of p53-responsive elements, such as *p21*, *Bax*, *XPC*, *DDB2*, and *GADD45* genes [21-24]. Additional evidence that BRCA1 is involved in transcriptional activation comes from the observation that BRCA1 associates with the RNA polymerase II holoenzyme complex via RNA helicase A [25]. However, BRCA1 is dispensable for basal transcription. BRCA1 is also involved in meiotic sex chromosome inactivation as it was observed that BRCA1-deficient cells lose ATR localisation to the XY body and H2AX phosphorylation. It was initially suggested that BRCA1 would be necessary for maintenance of XIST RNA on the inactive X chromosome [26], yet another study reported weak evidence for this role of

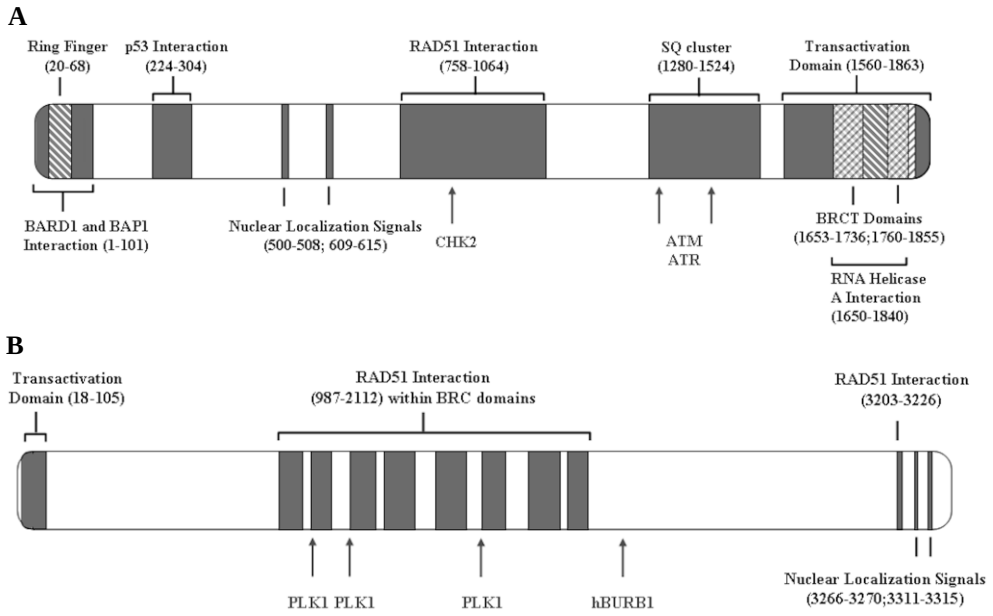


Figure 1. Features of the BRCA proteins. Functional domains and some interacting proteins sites are shown. A) BRCA1 (220kDa) and B) BRCA2 (384kDa). Arrows indicate phosphorylation sites.

BRCA1 [27]. Recently, it was shown that BRCA1 is involved in regulation of miRNAs biogenesis [28], accelerating the processing of primary miRNA primary transcripts, via the DROSHA microprocessor complex and SMAD3/p53/DHX9.

The RING domain of BRCA1 has E3 ubiquitin-ligase activity, which is enhanced when it forms a heterodimer with BARD1 [29]. The ubiquitylation process involves targeting proteins for degradation, alterations of cellular localization, and enzymatic activity changes. Morris and Solomon [30] observed that ubiquitylation events occur rapidly at stalled replication forks in S-phase cells after treating cells with hydroxyurea or irradiation and accumulate at DSB repair sites in vivo, and that depletion of BRCA1 or BARD1 abolished the ubiquitylation events. The *in vivo* targets of the BRCA1:BARD1 complex remain unknown, but are likely to be proteins involved in DNA repair, checkpoint signalling and/or regulation of chromatin dynamics [31, 32]. Moreover, it was reported that BRCA1 associates with the centrosome during mitosis and that its hypophosphorylated form binds to γ-tubulin, which is responsible for microtubules and mitotic spindle formation [33]. Regarding the role of BRCA1 in chromatin remodelling, it was recently reported that BRCA1 keeps the centromeric regions of heterochromatin in closed state via ubiquitylation of H2A and loss of BRCA1 leads to increased expression of non-coding DNA, i.e. satellite DNA [34]. BRCA1 also interacts with ER-α and PR, regulating their activities [35-37] and deregulation of this activity in $BRCA1^{+/-}$ cells may explain, at least in part, the tissue-specificity of the tumours arising in BRCA1 mutation carriers.

The BRCA2 protein accumulates in the nucleus mostly during G1/S phase [38]. BRCA2 is phosphorylated by Polo-like kinase (Plk1) in the BRC repeat region, but not at the repeats themselves [39], and at the Ser3291 residue [40]. Plk1-dependent phosphorylation of BRCA2 was shown to be enhanced during mitotic progression and inhibited by DNA damage, suggesting that this post-transcriptional modification coordinates the role of BRCA2 in these biological processes.

BRCA2 participates in DNA DSB repair by homologous-recombination, regulating the nuclear localization and function of RAD51. The interactions between BRCA2 and RAD51 are fundamental for the maintenance of cell division and chromosome structure. Studies have shown that in *BRCA2*-deficient cells, nuclear transport of RAD51 is impaired [41]. Additionally, *BRCA2*-deficient cells accumulate chromosomal breaks and aberrant mitotic exchanges during culture, whereas cell cycle checkpoint control and apoptotic responses to DNA damage remain intact [42]. Interaction of BRCA2 with RAD51 is mediated by six of the eight BRC repeats, located in the region defined by amino acids (a.a.) 987-2112 [43, 44] and an unrelated domain, named TR2, located at the C-terminal binding domain (a.a. 3203-3226) [40]. Another BRCA2-interacting protein is DSS1. The interaction occurs via BRCA2 a.a 2472-2957 and these proteins act together to efficiently target RAD51 to sites of DSB [45]. There is evidence that BRCA2 also mediates G2/M-phase control by interacting with BRCA2-associated factor 35 (BRAF35), which binds to branched DNA structures [46]. The interaction with BRAF35 occurs through the BRCA2 a.a. 1648-2190 [47]. Besides RAD51 and BRAF35, BRCA2 interacts with many other proteins, through different domains. For instance, BRCA2 interacts via a.a. 22-106 with EMSY, which is recruited to DNA-damage sites and is amplified in breast and ovarian cancers [48]. BRCA2 is involved in the Fanconi anaemia pathway and binds directly to some of its components: PALB2 via a.a. 21-39 [49] and FANCD2 via a.a. 2118-2566 [50]. BRCA2 a.a. 2386-2411 also interact with DMC1, necessary for meiotic homologous recombination [51].

Both *BRCA1* and *BRCA2* follow the Knudson's classic "two-hit" model, which describes the inactivation of tumor suppressor genes. In this respect, it was observed that the majority of *BRCA1/2*-related tumours arise after loss-of-heterozygosity (LOH) of the *BRCA1/2* genes [52]. LOH is the loss of the normal allele of a gene (second hit), when the other allele is already inactivated (first hit). In the case of *BRCA1/2*-mutation carriers one of the alleles is already inactivated by germline mutations and acquiring a somatic mutation is the second hit. Therefore, both *BRCA1* and *BRCA2* are tumour-suppressor genes.

***BRCA1/2* mutation spectrum**

Identification of the *BRCA1* and *BRCA2* genes allowed the genetic screening of the complete coding sequence of these genes in individuals from breast and ovarian cancer syndrome families. Mutations in the *BRCA1* and *BRCA2* genes are found throughout the whole coding sequence of the genes and there are no “mutation hot spots”. In general, evident pathogenic mutations lead to premature stop codons and, consequently, non-

functional or absent, due to degradation, proteins. The majority of the pathogenic mutations are small frameshift deletions or insertions and nonsense mutations, followed by splice variants that cause exon skipping or intron retention leading to a premature stop codon. Several large rearrangements were also reported in the literature, 80 in *BRCA1* and 17 in *BRCA2* [53].

The Breast Cancer Information Core (BIC) Database [54] lists more than 911 and 835 different pathogenic mutations in the *BRCA1* and *BRCA2* genes, respectively (data from March 2012). However, these only account for 48.6% and 41.0% for *BRCA1* and *BRCA2*, respectively, of all the different entries in the BIC Database (data from March 2012). In the LOVD (Leiden Open Variation Database) [55], a database for *BRCA1/2* variants reported in literature and results from functional assays, only 178 out of 528 variants from *BRCA1* and 107 out of 493 variants of *BRCA2* gene are described "as mutant control", "predicted deleterious", or other equivalent terms suggesting pathogenicity under the column "assays results" (data from March 2012). This indicates that for the majority of sequence variants it is not possible to determine the clinical relevance conclusively, also after assessment of its possible effect on protein function. In fact, during genetic screening, besides clearly pathogenic mutations and polymorphisms (generally considered neutral variants in terms of disease risk), several variants of undetermined clinical significance (VUS), also called unclassified variants (UVs), are identified.

Variants of undetermined clinical significance

VUS pose problems to the genetic counselling, since it is not possible to offer predictive genetic analysis with certainty to relatives at risk and the presence of a VUS is a complicating factor in the decision to perform risk reducing surgeries. In the BIC database 696 and 1116 variants are reported with unknown clinical relevance for BRCA1 and BRCA2, respectively. Most of the VUS, for which it is not possible to confirm or exclude a pathogenic effect, are missense variants with an unknown effect on folding or function of the protein. Additionally, VUS may also affect splicing, as discussed below. To deal with this problem, several bioinformatic tools have been developed to predict the consequence of an a.a. substitution based on nucleotide and a.a. conservation, its effect on the protein structure, RNA splicing, possible interaction site, phosphorylation, etc, which indirectly indicate whether the folding or function of the protein is affected. In Table 1 some of these online available tools and the methods used are described.

Table 1. Online programs to predict the impact of missense substitutions

Program	Website	Algorithm method
PolyPhen-2 [56]	http://genetics.bwh.harvard.edu/pph2/	uses sequence conservation, structure and SWISS-PROT annotation
SIFT [57, 58]	http://sift.jcvi.org/	sequence homology-based; input is multiple sequence alignment
Align-GVGD [59]	http://agvgd.iarc.fr/agvgd_input.php	uses a combination of Grantham Variation (GV), which is the amino acid (a.a.) evolutionary variation, and Grantham Deviation (GD), which is the difference in the biophysical properties between the WT a.a. and the newly encoded a.a.
Xvar [60]	http://mutationassessor.org/	Produces alignments based on the Uniprot ID, takes into account the annotated functional regions and protein-protein interactions.

While these *in silico* tools provide some evidence of the impact of the amino acid substitution in the protein, especially if their predictions are in agreement [61], conclusions should not be made based on these data alone [61, 62]. It is noteworthy that the sensitivity and specificity of these programs are different when different alignments of the same gene are provided [61]. Therefore, additional information such as family history analysis, the presence of the mutation in unrelated healthy controls, and/or functional tests also have to be taken into account in the decision process.

Multifactorial prediction models have been developed to ascertain the pathogenicity of variants [63-67]. The outcome of these models is in the form of likelihood ratios, which are based on the probability of the observed data being pathogenic against being neutral. Plon et al. proposed a 5-class system for the discrimination of different types of variants, depending on the probability score, together with clinical and research recommendations for each class [68]. Multifactorial models integrate data from clinical parameters, such as personal and family cancer history, co-segregation and co-occurrence data, *in silico* predictions and some models also allow inclusion of functional test results. In co-segregation studies the segregation of the unclassified variant through the family with the phenotype is assessed. Reports for co-occurrence of the variant with another clearly pathogenic mutation *in trans* also provide useful information. This can exclude pathogenicity of variants, since it is known that *BRCA2* compound heterozygotes have Fanconi anaemia, type D1 [69] and biallellic inactivation of *BRCA1* is thought to be embryonically lethal, as observed in mice [70]. Additionally, no individuals with homozygous *BRCA1* mutations were reported. Not even among the Ashkenazi Jewish

population, who have a very high incidence of two specific *BRCA1* and one *BRCA2* mutation.

Several functional assays have been designed and may be used to evaluate the effect of VUS on certain functions of the BRCA1/2 proteins. The basic goal/principles of these assays are briefly described in Table 2. The functional assays have, however, a number of limitations. They are restricted to existing knowledge on the protein functions; they do no test for all the *BRCA1/2* functions, rely on cDNA expression vectors to deliver part of the mutant proteins and are mostly performed by transfecting cancer cell lines, which already have genomic instability. Moreover, the assays described are extremely laborious and are not easy to implement in a diagnostic laboratory. An indirect assay is gene expression profiling using microarrays, which allows testing the changes in expression of genes and/or pathways in the presence of defective BRCA protein compared to WT protein. A few studies have tried to use this agnostic approach to identify gene classifiers in fibroblast cultures [87] and lymphoblastoid cell-lines (LCLs) [88, 89] from BRCA1/2-carriers and controls. These gene classifiers still require improvement and validation, as it is recognized that gene signatures obtained with the classifiers may be too specific to the samples used [90, 91]. An illustration of this is the fact that there is little overlap of the gene signatures found in different studies. Alternatively, gene signatures for the classification of VUS could be improved by including information on relevant biological processes.

Table 2. Functional assays used for evaluation of BRCA1/2 unclassified variants

Functional assay	**Brief description of the goal/principle**
Subcellular localization [64, 71]	BRCA1/2 contain nuclear localization signals and most of the known functions take place in the nucleus. Plasmids with the full-length BRCA1/2 proteins tagged to a immunofluorescent protein were transfected into cell lines. Predominant localization of the proteins in the cytoplasm suggests that either localization sites are disrupted or the protein folding is affected.
Centrosome amplification [64, 71]	Centrosome amplification is a marker of cell cycle disruption and chromosomal instability. BRCA1 and BRCA2 wild-type and mutant constructs are transfected in cell-lines. Centrosome amplification was measured for each construct, using immunofluorescence.
Transcription activation [71-75]	The activation of transcription is analyzed *in vitro* using a plasmid that encodes for fusion proteins containing the BRCA1 C-terminus (wt or with the VUS). The expression of reporter genes correlates with the transcription activation.
Small colony phenotype [76-78]	The expression of a fusion protein containing the C-terminal of BRCA1 in yeast results in small colony phenotypes. This phenotype is abolished in the presence of frameshift, nonsense and dysfunctional missense mutations in the C-terminal region of BRCA1.
Ubiquitin ligase activity [79]	BRCA1 N-terminal forms a stable heterodimer with BARD1 and the resulting complex has enhanced E3 ubiquitin-ligase activity. Missense variants in this region may affect the ubiquitin ligase activity. The yeast two-hybrid system is used to test the physical interactions in addition to the *in vitro* ubiquitin ligase assay. Variants that cause reductions in both tests are presumably deleterious.

Proteolytic sensitivity [71, 75, 80]	Expression of the BRCT-region (1646-1863) was carried out *in vitro*. BRCT-folding defects resulting from BRCA1 truncation mutations and missense substitutions are degraded by trypsin, whereas full-length BRCT is resistant to cleavage. Protein digested products are analysed after separation by electrophoresis.
Phosphopeptide binding [75, 81]	The tandem BRCT repeats (within 1646-1863) of BRCA1 function as phosphopeptide-binding modules and have high affinity to a specific motif as found in BRIP1 (pSXXF). Following the expression of the BRCT-region, the suspension is added to a bead-immobilized peptide affinity resin containing the phosphopeptide library (pSXXF) or with the corresponding dephosphorylated peptide (SXXF). Mutations affect the binding affinity.
DNA damage sensitivity [64, 82]	After, γ-irradiation and mitomycin C treatment, cells with mutant BRCA1/2 proteins will repair DNA DSBs by NHEJ, which is error-prone and will accumulate chromosomal instability and undergo apoptosis. The sensitivity to the treatment is reduced if normal *BRCA1* or *BRCA2* is introduced in Brca1- and BRCA2-deficient cells, respectively.
Homology-directed repair [64, 83, 84]	Cells transfected with *BRCA1/2* variants that affect the HR function will not express GFP, from another construct, which is quantified under the microscope.
Embryonic stem cell-based [85, 86]	BRCA1 and BRCA2 are required for ES cell viability. Human *BRCA1* and *BRCA2* BAC DNAs are used to rescue the lethality of Brca1- or Brca2- deficient ES cells, respectively.

Pre-mRNA Splicing

Splicing defects are a less well-recognized group of pathogenic mutations in the BRCA1 and BRCA2 genes. Many VUS with a putative effect on splicing remain unclassified as experimental tools and approaches are lacking. This partly, because the normal splice pattern and variation is unknown for both genes. Studies in this thesis aim at analysing the effect of VUS on pre-mRNA splicing.

Pre-mRNA splicing is an important step during the protein biosynthesis process, which occurs between transcription and translation. The presence of variants in the pre-mRNA molecule may disrupt important sequence elements and affect the pre-mRNA splicing.

Human genes contain relatively short exons and usually much larger introns. The latter account for at least 90% of the primary transcript and must be removed from the pre-mRNA molecule during the process of pre-mRNA splicing to generate stable protein-coding transcripts. To accomplish this, it is necessary that the splicing machinery recognizes the exons or introns and their boundaries. It is thought that in mammals there is predominantly "exon definition", whereas "intron definition" occurs in most other metazoans [92]. During the pre-mRNA splicing in humans, "exon definition" involves initial interaction across the exon between factors recognizing the donor splice site (DSS, 5'-SS) and the upstream acceptor splice site (ASS, 3'-SS). Subsequent to exon definition,

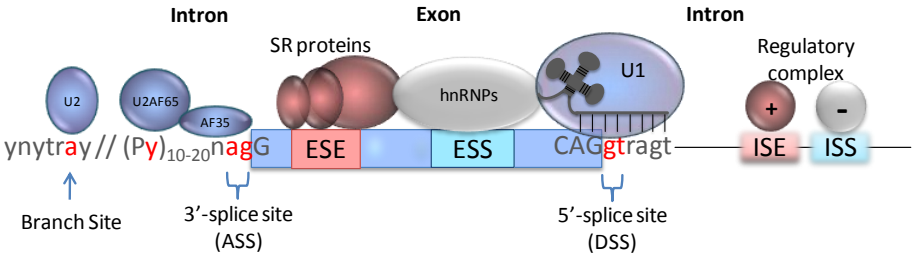


Figure 2 . Splicing regulatory elements of the pre-mRNA. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics [93], copyright (2004). The AG and GU consensus dinucleotides that directly flank the exon and the branch-point adenosine are conserved elements. There is also a polypyrimidine tract of variable length (y represents a pyrimidine base, cytosine or thymine) upstream of the ASS site. Additionally, there are exonic splice enhancers, exonic splice inhibitors, intronic splice enhancers, and intronic splice inhibitors, which play a role in splicing regulation. U1, U2, 35, U2AF65 refer to small nuclear ribonucleoproteins (snRNP) U1 and U2, U2 snRNP auxiliary factor 35 and U2 snRNP auxiliary factor 65, respectively, which are part of the spliceosome complex.

coordinated structural rearrangements occur to activate the spliceosome, an assembly of ribonucleoproteins that recognizes the splice sites and branch site. The branch site lies 20-50 bases upstream of the ASS and the nucleotide "A" of the consensus sequence is conserved in all genes. In addition to the donor splice site, acceptor splice site and branch site, other splicing regulatory elements (SREs) exist in the pre-mRNA molecule that serve as suppressors or promoters of splicing. Figure 2 depicts many of the important splice regulatory regions and interacting proteins. Depending on their function and location, SREs are called exonic splice enhancers (ESEs) or silencers (ESSs) and intronic splice enhancers (ISEs) or silencers (ISS). Most ESEs recruit members of the SR protein (proteins rich in serine and arginine) family, whereas ESSs bind to splicing repressors of the heterogeneous nuclear ribonucleoproteins (hnRNP) class.

Splicing and nonsense-mediated mRNA decay

After transcription, pre-mRNA molecules acquire a cap at the 5' ends and poly(A) tail at 3' ends. Then, through the process of splicing, which can be tissue specific, the intronic regions are excised and exons are joined to form mature mRNA molecules [94]. These events take place in the nucleus, before the mRNAs are transported to the cytosol, where translation occurs [94]. During splicing, an exon-junction complex (EJC) is deposited in each exon-exon junction [95].

Transcripts containing a premature termination codon (PTC) may undergo degradation due to a surveillance mechanism of the cells, named nonsense-mediated mRNA decay (NMD), which is thought to prevent the synthesis of truncated proteins due to splicing errors, which could have toxic effects, e.g. dominant negative interactions. NMD targets newly synthesised PTC-containing mRNA, when it is still bound by the RNA cap-binding protein heterodimer, during the translation process. A PTC is distinguished from a canonical stop codon, due to the presence of a downstream EJC. However, if the PTC lies

in the region starting 50-55 bases upstream of the last exon-exon junction (i.e. 25-30nt of an EJC), up to and including the last exon, it is not recognized and the transcript will not undergo NMD (reviewed in [96-98]). NMD does not usually downregulate completely the expression of aberrant *BRCA1/2* transcripts [99-101], however its effect depends on the variant. Therefore, NMD should be considered when analysing the effect of VUS on splicing. It is possible to inhibit the process, which will allow enriching for the mRNA derived from the variant allele, which otherwise may be degraded at some extent, hence ensuring that PTC-containing mRNAs are detectable.

Splice prediction software

Any nucleotide change either affecting or creating splice site sequences and SREs may affect the correct splicing of the transcript and thereby lead to production of non-functional protein (reviewed in [93, 102]). The putative effect of a variant on splicing may be predicted *in silico* using several available algorithms [103-106]. Although several algorithms were developed independently, nowadays there is also the possibility of using an integrated program Alamut (Interactive Biosoftware), which is commercial software that allows analyzing simultaneously seven of the online algorithms. Another advantage is the visualization of the results through graphics, which facilitate interpretation. For determination of splice sites, Alamut contains the SpliceSiteFinder-like, MaxEntScan, NNSplice, GeneSplicer and Human Splice Finder (HSF) algorithms. These algorithms compare the splice site signals of the wild-type sequence to mutant sequences using different approaches. The SpliceSiteFinder-like algorithm uses position weight matrices based on a set of human constitutive exon/intron junctions found in more than 10.000 mRNA sequences [103]. The DSS matrices were built using 3 bases from the exon and 6 bases from the intron, whereas the ASS matrices use 14 bases from the intron, which includes the polypyrimidine sequence, and 3 bases from the exon. The MaxEntScan is based on the maximum entropy principle and takes into account the surrounding nucleotides: for DSS it analyzes 3 bases in the exon and 6 bases in the intron and for ASS it analyzes 20 bases from the intron and 3 bases from the exon [104]. The NNSplice is based on neural networks trained to recognize GT donor and AG acceptor sites of dataset from 269 genes [105]. The DSS datasets contain 7 bases of the exon and 8 bases of the intron, starting with GT, whereas the ASS datasets contain 70 bases of intron and 20 bases of the exon. The GeneSplicer combines different splice prediction techniques to detect standard GT/AG splice sites found in 1115 genes [106]. The HSF was designed with information of all the human exons and introns and the DSS and ASS weight matrices were derived from Shapiro and Senapathy [103], containing 9 and 14 bases for the DSS and ASS, respectively [107]. For the branch site prediction, HSF uses the human consensus sequence YNYCRAY [108] and to select for strong candidates the software uses exclusion criteria, which include, for instance, the distance to the ASS [107]. In Alamut the detection of ESE sites is performed with ESEFinder and RESCUE-ESE. The ESEFinder algorithm detects ESE binding sites of four SR proteins (SF2, SC35, SRp40 and SRp55) as determined *in vitro*

[109], whereas the RESCUE-ESE is based on a set of 238 putative ESE, which are hexamer motifs found to be overrepresented in exons with weak splice sites versus introns and exons with strong splice sites [110]. Despite the fact that these algorithms present an estimation of the effect of variants on splicing, it remains indispensable to confirm the results experimentally, in particular when it concerns positions beyond the highly conserved splice donor (GT) and acceptor motifs (AG). This confirmation also contributes to the interpretation and improvement of the algorithms.

Genetic cancer-risk modifiers

It has been observed that in HBOC families both the penetrance of the disease and the phenotypic expression varies from family to family and even among individuals with the same mutation within the same family. Individualized advice about the most suitable risk-reducing strategy has been hampered by this heterogeneity of risks observed among *BRCA1/2* carriers. Comprehension of the individual genetic cancer site-specific risk among the *BRCA1/2* women would greatly improve the genetic counselling of these patients, by providing them a personalised clinical advice. It is believed that, besides environmental factors, variants in other genes act as genetic risk modifiers [111].

Several variants have been identified as breast cancer risk modifiers using candidate gene approaches and genome-wide association studies (GWAS) both in the general population and in *BRCA1/2* patients and families. Candidate gene approaches focus on variants in specific genes, with a known biological function, i.e., genes from specific pathways. Some variants affecting the risk of breast cancer, identified by this approach, are in progesterone receptor (*PR*) [112-114], caspase-8 (*CASP8*), and transforming growth factor beta (*TGFB1*) [115]. Also variants in the androgen receptor (*AR*) [116] and rs1801320 (*RAD51*) [117] were found to modify the age of breast cancer diagnosis in *BRCA1* mutation carriers and to reduce the risk of ovarian cancer among *BRCA1/2* mutation carriers, respectively. Candidate gene approaches, which allow to identify associations between traits or disease status and rare variants, are limited to the knowledge about the pathways and genes involved, therefore limiting the number of candidate genes [118]. Unlike the candidate gene approach, GWAS allow to identify variants throughout the entire genome which may be associated with specific traits in an unbiased way [119]. The drawback of the GWAS is that the identified variants are not necessarily the causative variants, but they might be in strong linkage disequilibrium with the functional variant. Another disadvantage of these studies is that they will only allow identification of common variants (minor allele frequency > 5%) and, therefore, they underestimate rare variants [120]. It is also noteworthy that the effect of variants reported in studies using this approach have small magnitudes (median OR of 1.3) [121]. GWAS have identified variants that affect breast cancer risk in the general population in the genes rs2981582 (*FGFR2*), rs3803662 (*TOX3*, also named *TNRC9*), rs889312 (*MAP3K1*), rs13281615 (8q24), rs3817198 (*LSP1*), rs13387042 (2q35), rs4973768 (3p24), rs6504950 (17q23), rs11249433 (1p11.2), rs999737 (*RAD51L1*), rs2046210 (*ESR1*), rs865686 (9q31.2), rs3734805 and

rs9383938 (*ESR1*) [122-129]. It was postulated that these variants could also affect the risk of *BRCA1/2* mutation carriers [130, 131]. Indeed, several studies have shown this effect, at least for some of the variants [132-136]. The polymorphism in *FGFR2* shows the strongest association with increased risk of breast cancer within *BRCA2* mutation carriers [132], although the variant rs1801320 (RAD51) has a higher hazardous rate [137]. A GWAS among *BRCA2* carriers showed that the FGFR2 rs2981575 was strongly associated with a modest increase of BC [138], whereas rs16917302 (ZN365) and rs311499 (20q13.3) had a protective effect.Within *BRCA1* mutation carriers the variant that showed the strongest association with breast cancer risk in GWAS lies in the region of *C19orf61/ANKLE1* (rs8170) [139], whereas decreased risk of breast cancer has been associated with the polymorphism *CASP8* D302H [133], with a polymorphism in the promoter region of the wild-type allele of *BRCA1* (rs16942) [140], and with rs10771399 (near *PTHLH*) [136].

For ovarian cancer, the first gene modifier in *BRCA1/2* carriers was found by Phelan and colleagues [141]. The authors reported that the rare *HRAS1* allele, previously reported to increase the risk of breast cancer in the general population, doubled the risk of ovarian cancer among *BRCA1* carriers. Breast cancer susceptibility genes *FGFR2, TNRC9* and *CASP8* D302H revealed no association within *BRCA1/2* ovarian cancer cases [142]. However this study included only 54 OC cases. Since 2009, GWAS have identified loci associated with sporadic ovarian cancer cases. The rs3814113 (*BNC2*), rs10088218 (8q24), rs8170 (*BABAM1*), and rs2363956 (19p13) were protective, whereas rs207590 (2q31), rs2665390 (3q25), and rs9303542 (17q21) were risk factors [143-145]. It is noteworthy that the locus 19p13 had also been previously found to be associated with breast cancer risk among *BRCA1* female carriers. Several of these loci also modify ovarian cancer risk among *BRCA1/2* carriers [146-148]. A recent study associated variants in the genes *IL1A* and *ALOX5* with a protective effect of ovarian cancer within a panel of 27 inflammatory genes [149]. The results from previous studies suggest that also these variants might influence the ovarian cancer risk among BRCA1/2 carriers.

Surveillance and risk-reducing strategies

Individuals at high risk of breast and ovarian cancer can benefit from adequate risk-managing strategies when a pathogenic *BRCA1/2* mutation has been found or a VUS can be classified as pathogenic. *BRCA1/2*-mutation carriers are offered two main strategies: periodic surveillance or risk-reducing surgery. Periodic surveillance consists of extensive regular screening by means of physical exams, MRI and mammography, pelvic physical and ecographic examination and measurement of serum cancer antigen 125 (CA 125) [150]. However, most BRCA-related ovarian tumours are high-grade serous carcinomas and there is no reliable test for early detection [151, 152].

Risk-reducing surgery includes bilateral salpingo-oophorectomy (RRSO) and prophylactic bilateral mastectomy (RRM). RRSO consists of surgical removal of the ovaries and fallopian tubes. In general, after RRSO, the remaining risk of extra ovarian or fallopian tube cancer is 3 - 9% in the peritoneum, which is the lining of the abdomen, and it

also reduces the breast cancer risk by 50%, if performed before the menopause [153]. However, RRSO is not an option for women who wish future maternity and has adverse effects, such as sexual dysfunction and increased risk of non-gynaecological morbidity and mortality, such as psychological and coronary diseases [154-156]. RRM is the surgical removal of the breasts, though it is not possible to remove the breast tissue entirely. This procedure reduced the risk of breast cancer by 90% and 95% in women that did not undergo RRSO and those that did, respectively [157, 158].

In addition, chemoprevention may also be a possibility for *BRCA1/2*-mutation carriers. Oral contraceptive pills are protective for ovarian cancer, however, since they contain an oestrogen, they increase the risk of breast cancer [159]. Therefore it is not a risk-free prevention strategy for *BRCA1/2* patients. Alternatively, these patients may opt for other chemoprevention options to decrease the risk of breast cancer, including anti-oestrogens such as tamoxifen or raloxifene. Anti-oestrogen drugs act by blocking the oestrogen receptor. Tamoxifen, known to reduce the risk of ER-positive breast tumours, was shown to reduce breast cancer incidence of cancer-free *BRCA2* mutation carriers but not BRCA1 mutation carriers, which is most likely associated with the fact that *BRCA1*-associated tumours are more frequently ER-negative [160]. Effects similar to tamoxifen in reducing the risk of invasive breast cancer in cancer-free post-menopausal women at high-risk of breast cancer were reported with raloxifene, with the additional advantage of having fewer adverse effects [161]. Additionally, tamoxifen also reduces the incidence of a second breast cancer both in pre- and post-menopausal *BRCA1/2* carriers [162, 163]. As an alternative to anti-oestrogens, there are drugs that inhibit the production of oestrogens. Aromatase inhibitors, such as anastrazole and exemestane, block the enzyme aromatase, which converts the hormone androgen into oestrogen. Consequently, less oestrogen is available to stimulate oestrogen-receptor positive cancers. Currently, an International Breast Cancer Intervention (IBIS II) clinical trial investigates the potential benefit of anastrazole regarding the breast cancer prevention of healthy post-menopausal *BRCA1/2*-mutation carriers [164]. Another potential drug that could be used as chemoprevention in women wishing future maternity is deslorelin. It inhibits the production of oestrogens in the ovaries and is a good candidate to substitute RRSO intervention, since its effect is reversible. In a preliminary study performed among premenopausal *BRCA1*-mutation carriers, deslorelin, given together with low-dose sex steroids, was well tolerated and it had minimal side effects [165].

Therapeutic-specific options

The classification of VUS is also very important for the treatment choice, since *BRCA1/2*-mutation carriers are eligible for specific treatment options. Double-strand and single strand breaks (DSB and SSB) are repaired by different mechanisms. DSB are repaired through highly accurate homologous recombination (HR) and error-prone non-homologous end-joining (NHEJ) mechanisms. SSBs are repaired by DNA base excision repair (BER), nucleotide excision repair (NER), and mismatch repair (MMR). BRCA1 plays important

roles in both HR and NHEJ, whereas BRCA2 is only involved in HR mechanism. Synthetic lethality is the process by which combined nonlethal cellular defects results in cell death [166]. This concept is the basis of the PARP-inhibitors therapeutics for BRCA-associated tumours. The rational is that inhibition of PARP1, which is involved in BER, in BRCAness tumours leads to cell death. Triple-negative breast carcinomas (ER, PR, and HER2 negative) are mostly frequent in BRCA1-carriers, but may also occur sporadically. Approximately 60% of the sporadic triple-negative (TN) breast cancers are also HR-deficient [167, 168]. Therefore, targeting PARP1 in these tumours was also investigated in clinical trials. However, the results of a Phase II and a Phase III trials were not promising [169, 170] and further studies are warranted. In contrast, clinical trials in hereditary *BRCA1/2*-associated tumours have shown promising results [171, 172], although a number of BRCA-associated tumours seems to be resistant to PARP-inhibitors. Co-targeting the P13K pathway was shown to potentiate the effect of PARP-inhibitors [173] and might be promising for a number of cases. Other cases, with low levels of PARP protein, will require different therapeutic approaches [174].

AIMS AND OUTLINE OF THIS THESIS

Genetic risk assessment in families with breast and ovarian cancer is currently based on genetic testing for a mutation in the *BRCA1* and *BRCA2* genes. If a mutation is found, an advice is given for surveillance and risk reducing strategies following national guidelines. Although this process looks quite straightforward, it has a number of limitations. One of them is the interpretation of the clinical relevance of the genetic variant found and consequently, the personal advice that can be given. In addition, since penetrance and cancer risk site are variable, individualized cancer risk is difficult to assess and therefore, currently, not possible in clinical practice. The major common aim of the studies described in this thesis is the improvement of the cancer risk assessment of the individuals at high risk of developing breast and ovarian cancer. This has been achieved by analysis of *BRCA1/2* genetic variants and genetic risk modifiers to determine their contribution to breast/ovarian cancer susceptibility and by setting-up tools for the characterization of genetic variants in *BRCA1/2* with respect to pathogenicity. Different strategies, described below for each of the chapters, were used to achieve this goal.

Outline of the thesis

In this thesis, an introduction to hereditary breast cancer, *BRCA1/2* genes, options to prevent cancer onset, and *BRCA1/2*-specific treatment scenarios is given in **chapter 1**. Special emphasis is given to *BRCA1/2* variants of undetermined clinical significance (VUS) and to genetic risk factors that modify the cancer risk for carriers.

.The identification and characterisation of variants in the *BRCA1/2* genes in Portuguese families, including the description of a Portuguese *BRCA2* founder mutation, and the observed associated site-specific cancer risks are presented in **chapter 2**. For the

screening of the Portuguese founder mutation an easier and faster three-step PCR method (patent nr. PT103726) has been developed.

As the cancer risk is not only associated with the specific *BRCA1/2* mutation, but can also be modified by mutations or variants in other genes, we chose a candidate-SNP approach to identify functionally relevant modifiers. The cancer site-specific risks associated with known functional polymorphisms involved in steroid hormone metabolism: two polymorphisms of the progesterone receptor (*PR*) gene, and the fibroblast growth factor receptor 2 (*FGFR2*) SNP rs2981582, as well as a three polymorphisms previously shown to modify sporadic breast cancer risk, were assessed in our local population (**chapter 3.1**). In **chapter 3.2** the risk of *FGFR2* SNP rs2981582 for ovarian cancer among an international cohort consisting of more than 20,000 *BRCA1/2* female mutation carriers was assessed.

The screening for mutations in the *BRCA1/2* genes yields a considerable number of VUS, which causes considerable anxiety and uncertainty in the families involved, since accurate genetic counselling of the members at risk is not possible. Therefore, we developed and optimized tools to study and classify these variants more accurately. In **chapter 4** the effect of VUS with a putative effect on splicing, as predicted by *in silico* algorithms, was analyzed at RNA level. The techniques used require the design of new primer sets for the analysis of each variant and are time consuming. In order to overcome this, the potential use of MLPA to assess *BRCA1* exon skipping events at RNA level was evaluated (**chapter 5**). Since only part of the VUS affects splicing, we developed a more general strategy to determine the significance of a VUS using gene expression profiling. In **chapter 6** we described the transcriptional response associated with *BRCA1* mutations. Since BRCA1 plays a role in DNA damage repair and cell cycle arrest, a differential gene-expression response is expected between irradiated cells from *BRCA1*-mutation carriers and controls. These differences were explored using pathways and network analyses. The aim was to construct a robust genetic signature to classify *BRCA1* VUS as pathogenic or neutral based on a selection of relevant genes.

The main findings of the studies presented in this thesis are presented and discussed in **chapter 7** and future perspectives are given.

REFERENCES

1. leDran, H.F., *Mémoire avec un précis de plusieurs observations sur le cancer.* Mémoires de l'Académie royale de chirurgie, 1757. **3**: p. 1-54.
2. Broca, P., *Traité des tumeurs*. Vol. 1. 1866, Paris: P. Asselin.
3. Lynch, H.T., et al., *Genetics, biomarkers, and control of breast cancer: A review.* Cancer Genet Cytogenet, 1984. **13**(1): p. 43-92.
4. Ford, D. and D.F. Easton, *The genetics of breast and ovarian cancer.* Br J Cancer, 1995. **72**(4): p. 805-812.
5. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
6. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-792.
7. The Breast Cancer Linkage, C., *Cancer risks in BRCA2 mutation carriers.* J Natl Cancer Inst, 1999. **91**(15): p. 1310-1316.
8. Chen, S. and G. Parmigiani, *Meta-Analysis of BRCA1 and BRCA2 Penetrance.* J Clin Oncol, 2007. **25**(11): p. 1329-1333.
9. van der Groep, P., E. van der Wall, and P. van Diest, *Pathology of hereditary breast cancer.* Cell Oncol, 2011. **34**(2): p. 71-88.
10. van Asperen, C.J., et al., *Cancer risks in BRCA2 families: estimates for sites other than breast and ovary.* J Med Genet, 2005. **42**(9): p. 711-719.
11. Loveday, C., et al., *Germline mutations in RAD51D confer susceptibility to ovarian cancer.* Nat Genet, 2011. **43**(9): p. 879-882.
12. Vargas, A., J. Reis-Filho, and S. Lakhani, *Phenotype-genotype correlation in familial breast cancer.* J Mammary Gland Biol Neoplasia, 2011. **16**(1): p. 27-40.
13. Foulkes, W.D., *Inherited susceptibility to common cancers.* N Engl J Med, 2008. **359**(20): p. 2143-2153.
14. Bennett, R.L., *Cancer genetics in the clinic: the chalenges and responsabilities of counselling and treating women at risk*, in *The role of genetics in breast and reproductive cancers*, P. Welcsh, Editor. 2009, Springer: New York. p. 3-20.
15. Smith, T.M., et al., *Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1.* Genome Res, 1996. **6**: p. 1029-1049
16. Tavtigian, S.V., *The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds.* Nature Genet, 1996. **12**: p. 333-337.
17. Welcsh, P.L. and M.C. King, *BRCA1 and BRCA2 and the genetics of breast and ovarian cancer.* Hum Mol Genet, 2001. **10**: p. 705 - 713.
18. MacLachlan, T.K. and W. El-Deiry, *Functional Interactions Between BRCA1 and the Cell Cycle*. 2005, Austin, USA: Landes Bioscience.
19. Scully, R., *Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage.* Cell, 1997. **90**: p. 425-435.
20. Tibbetts, R.S., et al., *Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress.* Genes Dev, 2000. **14**(23): p. 2989-3002.
21. Ouchi, T., et al., *BRCA1 regulates p53-dependent gene expression.* Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2302-2306.
22. Zhang, H., et al., *BRCA1 physically associates with p53 and stimulates its transcriptional activity.* Oncogene, 1998. **16**: p. 1713-1721.
23. Lu, M. and B.A. Arrick, *Transactivation of the p21 promoter by BRCA1 splice variants in mammary epithelial cells: evidence for both common and distinct activities of wildtype and mutant forms.* Oncogene, 2000. **19**(54): p. 6351-6360.
24. Hartman, A.-R. and J.M. Ford, *BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair.* Nat Genet, 2002. **32**(1): p. 180-184.
25. Anderson, S.F., et al., *BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A.* Nat Genet, 1998. **19**(3): p. 254-256.

26. Ganesan, S., et al., *BRCA1 supports XIST RNA concentration on the inactive X chromosome.* Cell, 2002. **111**(3): p. 393-405.
27. Xiao, C., et al., *The XIST noncoding RNA functions independently of BRCA1 in X inactivation.* Cell, 2007. **128**(5): p. 977-989.
28. Kawai, S. and A. Amano, *BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex.* J Cell Biol, 2012. **197**(2): p. 201-208.
29. Hashizume, R., et al., *The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation.* J Biol Chem, 2001. **276**: p. 14537-14540.
30. Morris, J.R. and E. Solomon, *BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair.* Hum Mol Genet, 2004. **13**(8): p. 807-817.
31. Joukov, V., et al., *The BRCA1/BARD1 heterodimer modulates Ran-dependent mitotic spindle assembly.* Cell, 2006. **127**(3): p. 539-552.
32. Boulton, S.J., *Cellular functions of the BRCA tumour-suppressor proteins.* Biochem Soc Trans, 2006. **34**: p. 633 - 645.
33. Hsu, L.-C. and R.L. White, *BRCA1 is associated with the centrosome during mitosis.* Proc Natl Acad Sci USA, 1998. **95**: p. 12983–12988.
34. Zhu, Q., et al., *BRCA1 tumour suppression occurs via heterochromatin-mediated silencing.* Nature, 2011. **477**(7363): p. 179-184.
35. Zheng, L., et al., *BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor.* Proc Natl Acad Sci U S A, 2001. **98**(17): p. 9587-9592.
36. Ma, Y., et al., *The breast cancer susceptibility gene BRCA1 regulates progesterone receptor signaling in mammary epithelial cells.* Mol Endocrinol, 2006. **20**(1): p. 14-34.
37. Calvo, V. and M. Beato, *BRCA1 counteracts progesterone action by ubiquitination leading to progesterone receptor degradation and epigenetic silencing of target promoters.* Cancer Res, 2011. **71**(9): p. 3422-3431.
38. Vaughn, J.P., et al., *Cell cycle control of BRCA2.* Cancer Research, 1996. **56**(20): p. 4590-4594.
39. Lee, M., M.J. Daniels, and A.R. Venkitaraman, *Phosphorylation of BRCA2 by the Polo-like kinase Plk1 is regulated by DNA damage and mitotic progression.* Oncogene, 2004. **23**(4): p. 865-872.
40. Esashi, F., et al., *Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2.* Nat Struct Mol Biol, 2007. **14**(6): p. 468-474.
41. Davies, A.A., et al., *Role of BRCA2 in control of the RAD51 recombination and DNA repair protein.* Mol Cell, 2001. **7**(2): p. 273-282.
42. Yoshida, K. and Y. Miki, *Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage.* Cancer Sci, 2004. **95**(11): p. 866-871.
43. Chen, P.L., et al., *The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment.* Proc Natl Acad Sci U S A, 1998. **95**: p. 5287-5292.
44. Wong, A.K.C., et al., *RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene BRCA2.* J Biol Chem, 1997. **272**: p. 31941−31944.
45. Marston, N.J., et al., *Interaction between the Product of the Breast Cancer Susceptibility Gene BRCA2 and DSS1, a Protein Functionally Conserved from Yeast to Mammals.* Mol Cell Biol, 1999. **19**(7): p. 4633-4642.
46. Marmorstein, L.Y., et al., *A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression.* Cell, 2001. **104**(2): p. 247-257.
47. Shiekhattar, R., *BRAF35 protein and BRCA2/BRAF35 complex and methods of use*, U.S. Patent, Editor. 2010, The Wistar Institute: USA.
48. Hughes-Davies, L., et al., *EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer.* Cell, 2003. **115**: p. 523 - 535.
49. Oliver, A.W., et al., *Structural basis for recruitment of BRCA2 by PALB2.* EMBO Rep, 2009. **10**(9): p. 990-996.
50. Hussain, S., et al., *Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways.* Hum Mol Genet, 2004. **13**(12): p. 1241-1248.

51. Thorslund, T., F. Esashi, and S.C. West, *Interactions between human BRCA2 protein and the meiosis-specific recombinase DMC1*. EMBO J, 2007. **26**(12): p. 2915-2922.
52. Osorio, A., et al., *Loss of heterozygosity analysis at the BRCA loci in tumor samples from patients with familial breast cancer*. Int J Cancer, 2002. **99**: p. 305 - 309.
53. Sluiter, M. and E. van Rensburg, *Large genomic rearrangements of the BRCA1 and BRCA2 genes: review of the literature and report of a novel BRCA1 mutation.* Breast Cancer Res Treat. **125**(2): p. 325-349.
54. Logan, C.W., et al., *Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity.* Hum Mutat, 2010. **31**(6): p. 1484-1505.
55. *LOVD - Leiden Open Variation Database*.
56. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Meth, 2010. **7**(4): p. 248-249.
57. Ng, P.C. and S. Henikoff, *SIFT: predicting amino acid changes that affect protein function.* Nucleic Acids Res, 2003. **31**(13): p. 3812-3814.
58. Kumar, P., S. Henikoff, and P.C. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protocols, 2009. **4**(8): p. 1073-1081.
59. Mathe, E., *Computational approaches for predicting the biological effect of p53 missense mutation: a comparison of three sequence analysis based methods.* Nucleic Acids Res, 2006. **34**: p. 1317-1325.
60. Reva, B., Y. Antipin, and C. Sander, *Determinants of protein function revealed by combinatorial entropy optimization.* Genome Biol, 2007. **8**(11): p. R232.
61. Hicks, S., et al., *Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed.* Hum Mutat, 2011. **32**(6): p. 661-668.
62. Dorfman, R., et al., *Do common in silico tools predict the clinical consequences of amino-acid substitutions in the CFTR gene?* Clin Genet, 2010. **77**(5): p. 464-473.
63. Goldgar, D.E., et al., *Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2.* Am J Hum Genet, 2004. **75**: p. 535 - 544.
64. Wu, K., et al., *Functional Evaluation and Cancer Risk Assessment of BRCA2 Unclassified Variants.* Cancer Res, 2005. **65**(2): p. 417-426.
65. Easton, D.F., et al., *A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes.* Am J Hum Genet, 2007. **81**(5): p. 873 - 883.
66. Tavtigian, S.V., et al., *Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications.* Hum Mutat, 2008. **29**(11): p. 1342-1354.
67. Gomez Garcia, E., et al., *A method to assess the clinical significance of unclassified variants in the BRCA1 and BRCA2 genes based on cancer family history.* Breast Cancer Res, 2009. **11**(1): p. R8.
68. Plon, S.E., et al., *Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results.* Human Mutation, 2008. **29**(11): p. 1282-1291.
69. Howlett, N.G., et al., *Biallelic inactivation of BRCA2 in Fanconi anemia.* Science, 2002. **297**(5581): p. 606-609.
70. Hakem, R., et al., *The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse.* Cell, 1996. **85**(7): p. 1009-1023.
71. Lovelock, P.K., et al., *Genetic, functional, and histopathological evaluation of two C-terminal BRCA1 missense variants.* J Med Genet, 2006. **43**: p. 74 - 83.
72. Vallon-Christersson, J., et al., *Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families.* Hum Mol Genet, 2001. **10**: p. 353 - 360.
73. Phelan, C.M., et al., *Classification of BRCA1 missense variants of unknown clinical significance.* J Med Genet, 2005. **42**: p. 138 - 146.
74. Carvalho, M.A., et al., *Determination of cancer risk associated with germ line BRCA1 missense variants by functional analysis.* Cancer Res, 2007. **67**: p. 1494-1501.
75. Lee, M.S., et al., *Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays.* Cancer Res, 2010. **70**(12): p. 4880-4890.

76. Humphrey, J.S., et al., *Human BRCA1 inhibits growth in yeast: potential use in diagnostic testing.* Proc Natl Acad Sci U S A, 1997. **94**(11): p. 5820-5825.
77. Coyne, R.S., et al., *Functional characterization of BRCA1 sequence variants using a yeast small colony phenotype assay.* Cancer Biol Ther, 2004. **3**: p. 453 - 457.
78. Caligo, M.A., et al., *A yeast recombination assay to characterize human BRCA1 missense variants of unknown pathological significance.* Hum Mutat, 2009. **30**(1): p. 123-133.
79. Morris, J.R., et al., *Genetic analysis of BRCA1 ubiquitin ligase activity and its relationship to breast cancer susceptibility.* Hum Mol Genet, 2006. **15**(4): p. 599-606.
80. Williams, R.S., et al., *Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations.* J Biol Chem, 2003. **278**: p. 53007 - 53016.
81. Williams, R.S., et al., *Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1.* Nat Struct Mol Biol, 2004. **11**(6): p. 519-525.
82. Moynahan, M.E., T.Y. Cui, and M. Jasin, *Homology-directed DNA repair, mitomycin-C resistance, and chromosome stability is restored with correction of a Brca1 mutation.* Cancer Res, 2001. **61**(12): p. 4842-4850.
83. Ransburgh, D.J., et al., *Identification of breast tumor mutations in BRCA1 that abolish its function in homologous DNA recombination.* Cancer Res, 2010. **70**(3): p. 988-95.
84. Farrugia, D.J., et al., *Functional assays for classification of BRCA2 variants of uncertain significance.* Cancer Res, 2008. **68**(9): p. 3523-3531.
85. Chang, S., et al., *Expression of human BRCA1 variants in mouse ES cells allows functional analysis of BRCA1 mutations.* JClin Invest, 2009. **119**(10): p. 3160-3171.
86. Kuznetsov, S.G., P. Liu, and S.K. Sharan, *Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2.* Nat Med, 2008. **14**(8): p. 875-881.
87. Kote-Jarai, Z., et al., *Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage.* Clin Cancer Res, 2006. **12**(13): p. 3896-3901.
88. Waddell, N., et al., *BRCA1 and BRCA2 missense variants of high and low clinical significance influence lymphoblastoid cell line post-irradiation gene expression.* PLoS Genet, 2008. **4**(5): p. e1000080.
89. Walker, L.C., et al., *Use of DNA-damaging agents and RNA pooling to assess expression profiles associated with BRCA1 and BRCA2 mutation status in familial breast cancer patients.* PLoS Genet, 2010. **6**(2): p. e1000850.
90. Michiels, S., S. Koscielny, and C. Hill, *Prediction of cancer outcome with microarrays: a multiple random validation strategy.* Lancet, 2005. **365**(9458): p. 488-492.
91. Ein-Dor, L., O. Zuk, and E. Domany, *Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer.* Proc Natl Acad Sci USA, 2006. **103**(15): p. 5923-5928.
92. Xiao, X., et al., *Coevolutionary networks of splicing cis-regulatory elements.* PNAS, 2007. **104**(47): p. 18583-18588.
93. Pagani, F. and F.E. Baralle, *Genomic variants in exons and introns: identifying the splicing spoilers.* Nat Rev Genet, 2004. **5**(5): p. 389-396.
94. Berg, J.M., J.L. Tympczko, and L. Stryer, *Biochemistry*. 5th ed. 2002, New York: W H Freeman.
95. Le Hir, H., et al., *The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions.* EMBO J, 2000. **19**(24): p. 6860-6869.
96. Isken, O. and L.E. Maquat, *The multiple lives of NMD factors: balancing roles in gene and genome regulation.* Nat Rev Genet, 2008. **9**(9): p. 699-712.
97. Brogna, S. and J. Wen, *Nonsense-mediated mRNA decay (NMD) mechanisms.* Nat Struct Mol Biol, 2009. **16**(2): p. 107-113.
98. Mühlemann, O., et al., *Recognition and elimination of nonsense mRNA.* Biochim Biophys Acta, 2008. **1779**(9): p. 538-549.
99. Perrin-Vidoz, L., et al., *The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons.* Hum Mol Genet, 2002. **11**(23): p. 2805-2814.
100. Anczuków, O., et al., *Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins?* Hum Mutat, 2008. **29**(1): p. 65-73.

101. Ware, M.D., et al., *Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene?* Oncogene, 2006. **25**(2): p. 323-328.
102. Douglas, A.G.L. and M.J.A. Wood, *RNA splicing: disease and therapy.* Brief Funct Genomics, 2011. **10**(3): p. 151-164.
103. Shapiro, M.B. and P. Senapathy, *RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression.* Nucleic Acids Res, 1987. **15**(17): p. 7155-7174.
104. Yeo, G. and C.B. Burge, *Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals.* J Comput Biol, 2004. **11**(2-3): p. 377-394.
105. Reese, M., et al., *Improved splice site detection in Genie.* J Comput Biol, 1997. **4**(3): p. 311-23.
106. Pertea, M., X. Lin, and S.L. Salzberg, *GeneSplicer: a new computational method for splice site prediction.* Nucleic Acids Res, 2001. **29**(5): p. 1185-1190.
107. Desmet, F.-O., et al. (2009) *Human Splicing Finder: an online bioinformatics tool to predict splicing signals*. Nucl Acids Res, 1-14 DOI: 10.1093/nar/gkp215.
108. Green, M.R., *Biochemical mechanisms of constitutive and regulated pre-mRNA splicing.* Annu Rev Cell Biol, 1991. **7**(1): p. 559-599.
109. Cartegni, L., et al., *ESEfinder: a web resource to identify exonic splicing enhancers.* Nucleic Acids Res, 2003. **31**(13): p. 3568-3571.
110. Fairbrother, W.G., et al. (2004) *RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons*. Nucleic Acids Res **32**, W187–190 DOI: 10.1093/nar/gkh393.
111. Antoniou, A.C., et al., *Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study.* Genetic Epidemiology, 2001. **21**(1): p. 1-18.
112. Dunning, A.M., et al., *A systematic review of genetic polymorphisms and breast cancer risk.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(10): p. 843-854.
113. Vivo, I.D., et al., *A functional polymorphism in the progesterone receptor gene is associated with an increase in breast cancer risk.* Cancer Res, 2003. **63**(17): p. 5236-5238.
114. Romano, A., et al., *Impact of two functional progesterone receptor polymorphisms (PRP): +331G/A and PROGINS on the cancer risks in familial breast/ovarian cancer.* The Open Cancer Journal, 2007. **1**: p. 1-8.
115. Cox, A., et al., *A common coding variant in CASP8 is associated with breast cancer risk.* Nat Genet, 2007. **39**(3): p. 352-358.
116. Rebbeck, T.R., et al., *Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat.* Am J Hum Genet, 1999. **64**: p. 1371 - 1377.
117. Wang, W.W., et al., *A single nucleotide polymorphism in the 5' untranslated region of RAD51 and risk of cancer among BRCA1/2 mutation carriers.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(9): p. 955-960.
118. Zhu, M. and S. Zhao, *Candidate gene identification approach: progress and challenges.* Int J Biol Sci, 2007. **3**: p. 420-427.
119. Witte, J.S., *Genome-wide association studies and beyond.* Ann Rev Public Health, 2010. **31**(1): p. 9-20.
120. Iles, M.M., *What can genome-wide association studies tell us about the genetics of common disease?* PLoS Genet, 2008. **4**(2): p. e33.
121. Hindorff, L.A., et al., *Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.* Proc Natl Acad Sci USA, 2009. **106**(23): p. 9362-9367.
122. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**: p. 1087-1093.
123. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nat Genet, 2007. **39**(7): p. 870-874.
124. Stacey, S.N., et al., *Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.* Nat Genet, 2007. **39**: p. 865 - 869.
125. Stacey, S.N., et al., *Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.* Nat Genet, 2008. **40**(6): p. 703-706.
126. Zheng, W., et al., *Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1.* Nat Genet, 2009. **41**(3): p. 324-328.

127. Ahmed, S., et al., *Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2.* Nat Genet, 2009. **41**(5): p. 585-590.
128. Fletcher, O., et al., *Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study.* J Natl Cancer Inst 2011.
129. Ghoussaini, M., et al., *Genome-wide association analysis identifies three new breast cancer susceptibility loci.* Nat Genet, 2012. **44**(3): p. 312-318.
130. Antoniou, A., et al., *A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes.* Br J Cancer, 2002. **86**: p. 76-83.
131. Antoniou, A.C. and D.F. Easton, *Models of genetic susceptibility to breast cancer.* Oncogene, 2006. **25**: p. 5898-5905.
132. Antoniou, A.C., et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. **82**: p. 937-948.
133. Engel, C., et al., *Association of the variants CASP8 D302H and CASP10 V410I with breast and ovarian cancer risk in BRCA1 and BRCA2 mutation carriers.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(11): p. 2859-2868.
134. Antoniou, A.C., et al., *Common alleles at 6q25.1 and 1p11.2 are associated with breast cancer risk for BRCA1 and BRCA2 mutation carriers.* Hum Mol Genet, 2011.
135. Antoniou, A.C., et al., *Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers.* Hum Mol Genet, 2009. **18**(22): p. 4442-4456.
136. Antoniou, A.C., et al. (2012) *Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers*. Breast Cancer Res **14**, R33 DOI: doi:10.1186/bcr3121.
137. Antoniou, A.C., et al., *RAD51 135G-->C modifies breast cancer risk among BRCA2 mutation carriers: results from a combined analysis of 19 Studies.* Am J Hum Genet, 2007. **81**(6): p. 1186-1200.
138. Gaudet, M.M., et al., *Common genetic variants and modification of penetrance of BRCA2-associated breast cancer.* PLoS Genet, 2010. **6**(10): p. e1001183.
139. Antoniou, A.C., et al., *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.* Nat Genet, 2010. **42**(10): p. 885-892.
140. Cox, D.G., et al., *Common variants of the BRCA1 wild-type allele modify the risk of breast cancer in BRCA1 mutation carriers.* Hum Mol Genet, 2011. **20**(23): p. 4732-4747.
141. Phelan, C.M., et al., *Ovarian cancer risk in BRCA1 carriers is modified by the HRAS1 variable number of tandem repeat (VNTR) locus.* Nature Genet, 1996. **12**: p. 309-311.
142. Latif, A., et al. (2010) *Breast cancer susceptibility variants alter risk in familial ovarian cancer*. Fam Cancer, DOI: doi:10.1007/s10689-010-9349-2.
143. Song, H., et al., *A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2.* Nat Genet, 2009. **41**(9): p. 996-1000.
144. Goode, E.L., et al., *A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24.* Nat Genet, 2010. **42**(10): p. 874-879.
145. Bolton, K.L., et al., *Common variants at 19p13 are associated with susceptibility to ovarian cancer.* Nat Genet, 2010. **42**(10): p. 880-884.
146. Ramus, S.J., et al., *Genetic Variation at 9p22.2 and Ovarian Cancer Risk for BRCA1 and BRCA2 Mutation Carriers.* J Natl Cancer Inst, 2011. **103**(2): p. 105-116.
147. Couch, F.J., et al., *Common Variants at the 19p13.1 and ZNF365 Loci Are Associated with ER Subtypes of Breast Cancer and Ovarian Cancer Risk in BRCA1 and BRCA2 Mutation Carriers.* Cancer Epidemiol Biomarkers Prev, 2012.
148. Ramus, S.J., et al. (2012) *Ovarian cancer susceptibility alleles and risk of ovarian cancer in BRCA1 and BRCA2 mutation carriers*. Hum Mutat, n/a-n/a DOI: 10.1002/humu.22025.
149. White, K.L., et al., *Ovarian cancer risk associated with inherited inflammation-related variants.* Cancer Res, 2012. **72**(5): p. 1064-1069.

150. Robson, M. and K. Offit, *Management of an inherited predisposition to breast cancer.* N Engl J Med, 2007. **357**(2): p. 154-162.
151. Hermsen, B.B.J., et al., *No efficacy of annual gynaecological screening in BRCA1//2 mutation carriers; an observational follow-up study*. Br J Cancer, 2007. **96**(9): p. 1335-1342.
152. van der Velde, N.M., et al., *Time to stop ovarian cancer screening in BRCA1/2 mutation carriers?* Int J Cancer, 2009. **124**(4): p. 919-923.
153. Rebbeck, T.R., N.D. Kauff, and S.M. Domchek, *Meta-analysis of risk reduction estimates associated with risk-reducing salpingo-oophorectomy in BRCA1 or BRCA2 mutation carriers.* J Natl Cancer Inst, 2009. **101**(2): p. 80-87.
154. Robson, M., et al., *Quality of life in women at risk for ovarian cancer who have undergone risk-reducing oophorectomy*. Gynecol Oncol, 2003. **89**(2): p. 281-287.
155. Rocca, W.A., et al., *Long-term risk of depressive and anxiety symptoms after early bilateral oophorectomy.* Menopause, 2008. **15**(6): p. 1050-1059.
156. Rivera, C.M., et al., *Increased cardiovascular mortality after early bilateral oophorectomy*. Menopause, 2009. **16**(1): p. 15-23.
157. Meijers-Heijboer, H., et al., *Breast cancer after prophylactic bilateral mastectomy in women with a BRCA1 or BRCA2 mutation.* N Engl J Med, 2001. **345**(3): p. 159-164.
158. Rebbeck, T.R., et al., *Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE study group.* J Clin Oncol, 2004. **22**(6): p. 1055-1062.
159. Brohet, R.M., et al., *Oral contraceptives and breast cancer risk in the international BRCA1/2 carrier cohort study: a report from EMBRACE, GENEPSO, GEO-HEBON, and the IBCCS collaborating group.* J Clin Oncol, 2007. **25**(25): p. 3831-3836.
160. King, M.-C., et al., *Tamoxifen and breast cancer incidence among women with inherited mutations in BRCA1 and BRCA2*. JAMA, 2001. **286**(18): p. 2251-2256.
161. Vogel, V.G., et al., *Update of the national surgical adjuvant breast and bowel project study of tamoxifen and raloxifene (STAR) P-2 trial: preventing breast cancer.* Cancer Prev Res, 2010. **3**(6): p. 696-706.
162. Metcalfe, K., et al., *Contralateral breast cancer in BRCA1 and BRCA2 mutation carriers.* J Clin Oncol, 2004. **22**(12): p. 2328-2335.
163. Gronwald, J., et al., *Tamoxifen and contralateral breast cancer in BRCA1 and BRCA2 carriers: An update.* Int J Cancer, 2006. **118**(9): p. 2281-2284.
164. Cuzick, J., *IBIS II: a breast cancer prevention trial in postmenopausal women using the aromatase inhibitor anastrozole.* Expert Rev Anticancer Ther, 2008. **8**(9): p. 1377-85.
165. Weitzel, J.N., et al., *Reduced mammographic density with use of a gonadotropin-releasing hormone agonist-based chemoprevention regimen in BRCA1 carriers.* Clin Cancer Res, 2007. **13**(2): p. 654-658.
166. Kaelin, W.G., *The concept of synthetic lethality in the context of anticancer therapy.* Nat Rev Cancer, 2005. **5**(9): p. 689-698.
167. Miyoshi, Y., K. Murase, and K. Oh, *Basal-like subtype and BRCA1 dysfunction in breast cancers.* Int J Clin Oncol, 2008. **13**(5): p. 395-400.
168. Turner, N.C., et al., *BRCA1 dysfunction in sporadic basal-like breast cancer*. Oncogene, 2006. **26**(14): p. 2126-2132.
169. Gelmon, K.A., et al., *Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study*. Lancet Oncol, 2011. **12**(9): p. 852-861.
170. O'Shaughnessy, J., et al., *Iniparib plus chemotherapy in metastatic triple-negative breast cancer.* N Engl J Med, 2011. **364**(3): p. 205-214.
171. Fong, P.C., et al., *Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers.* N Engl J Med, 2009. **361**(2): p. 123-134.
172. Tutt, A., et al., *Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial.* Lancet, 2010. **376**(9737): p. 235-244.
173. Kimbung, S., et al., *Co-targeting of the PI3K pathway improves the response of BRCA1 deficient breast cancer cells to PARP1 inhibition.* Cancer Lett, 2012. **319**(2): p. 232-241.

174. Domagala, P., et al., *PARP-1 expression in breast cancer including BRCA1-associated, triple negative and basal-like tumors: possible implications for PARP-1 inhibitor therapy.* Breast Cancer Res Treat, 2011. **127**(3): p. 861-869.

CHAPTER 2

Screening for a *BRCA2* rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes

Patrícia M. Machado*, Rita D. Brandão*, Branca M. Cavaco, Joana Eugénio, Sandra Bento, Mónica Nave, Paula Rodrigues, Aires Fernandes, Fátima Vaz
*contributed equally to this study

Journal of Clinical Oncology, 2007, 25(15):2027-34

ABSTRACT

BRCA2 rearrangements are rare genetic events. A large *BRCA2* genomic insertion was recurrently observed in our participants, and we sought to characterize it at the molecular and phenotypic level.

We studied 210 high-risk breast/ovarian cancer families. Fifty-three probands were fully screened for *BRCA1/2* mutations, and three of 53 had a large insertion in exon 3 of *BRCA2*. This finding was analyzed by polymerase chain reaction (PCR), reverse transcriptase PCR (RT-PCR), and sequencing. An additional 157 consecutive families were screened for this mutation by a three-step PCR method. Phenotype and haplotype analysis was also performed.

Sixteen *BRCA* mutations were observed in 19 of 53 patients (36% detection rate). A recurrent Alu motif insertion in position c.156_157 was observed after sequencing of an abnormal fragment obtained after the amplification of *BRCA2* exon 3. RT-PCR revealed exon 3 skipping. Screening of this rearrangement identified 14 additional families (out of 157). In total, 17 (8%) of 210 high-risk families ascertained in our clinic were positive for this mutation. Segregation of a common haplotype (from D13S260 to D13S1695) confirmed a common origin, estimated to have occurred 2,400 to 2,600 years ago. The following four cancer phenotypes were observed in the 17 positive families: female breast (n = 9), male breast (n = 4), breast/ovarian (n = 2), and heterogeneous (n = 2). Male breast cancer was more frequently observed in c.156_157insAlu–positive families compared with negative families (23% *v* 12%, respectively), and 33% of all male breast cancer families with an identified *BRCA* mutation were c.156_157insAlu positive.

c.156_157insAlu is a founder mutation of Portuguese origin and is the most frequent *BRCA2* rearrangement described to date.

INTRODUCTION

Genomic rearrangements, or the wholesale movement of sequences from one position to another in genomic DNA [1], correspond to 8% to 40% of all mutations in the *BRCA1* gene,2 but only seven rearrangements have been described in *BRCA2*-positive families [2-7]. These include large deletions [3, 4], duplications [4], deletions/insertions [5], and Alu insertions [6, 7]. Rarity of *BRCA*2 rearrangements may be a result of the fact that sensitive assays for rearrangement detection have only recently been added to the routine molecular diagnosis of breast/ovarian cancer predisposition [8, 9]. Furthermore, it has been reported that families negative for *BRCA2* rearrangements were breast/ovarian cancer families or female breast cancer–only families. Male breast cancer families, which are mainly associated with *BRCA2* mutations, could have a higher frequency of rearrangements in this gene [2]. In fact, three of the described *BRCA2* rearrangements were observed in male breast cancer families negative for *BRCA1/2* point mutations [4].

Although no founder effect has yet been described for *BRCA2* rearrangements, *BRCA1* genomic deletions have been found to represent major founder mutations in the Dutch population [10]. Founder mutations are identified only in specific countries or ethnic groups, suggesting that they have spread from a single ancestor. Founder *BRCA1/2* mutations have been described in the Ashkenazi Jewish population [11] and in other countries such as Finland, the Netherlands, Norway, Russia, Latvia [12], and Spain [13]. Identification of founder mutations and their ethnic and geographic origins allows a more rational and faster approach to mutational screening and genetic counseling in defined subpopulations because *BRCA1/2* are large genes (each one spanning > 100 kb of genomic DNA) and do not have hot spots. Unless a fast screening possibility exists, full screening of all the exonic and exon-intron boundary sequences is necessary, which makes this diagnosis expensive and time consuming [14].

The spectrum of *BRCA1/2* mutations in Portugal includes few recurrent mutations, probably because the Portuguese genetic background is heterogeneous. In fact, peoples of different origins invaded the Iberian Peninsula, and Portuguese sailors and emigrants have been in contact with several peoples, in all continents, since the 15th century.

During full gene *BRCA1/2* screening of high-risk breast/ovarian cancer families, a large insertion in exon 3 of *BRCA2* was recurrently observed. Characterization of this event revealed an Alu insertion in nucleotide c.156_157 of *BRCA*2, which was previously observed in a Portuguese family [7]. Because this large insertion was likely to represent a founder mutation, we optimized a three-step polymerase chain reaction (PCR) method to screen all new families ascertained in our clinic for*BRCA1/2* genetic testing, and the founder effect hypothesis was further explored by haplotype analysis. Phenotypic characterization of positive families was carried out.

PATIENTS AND METHODS

Participants

Selection criteria for full *BRCA1/2* screening were 25% of *BRCA1/2* combined probability of a mutation [15, 16], family history of male breast cancer, or diagnosis of breast cancer at less than 30 years of age. Pedigrees included at least three family generations, and all patients underwent genetic counseling and signed an informed consent form, according to procedures approved by the ethics committee of our institute. In patients found to be positive for the most frequent *BRCA2* mutation, permission was obtained to disclose the results to relatives at risk and invite them for genetic screening and for haplotype analysis with polymorphic *BRCA2*-linked markers.

The initial group of 53 patients reported in this study was the first to be analyzed for *BRCA1/2* mutations in the context of a multidisciplinary group, and counseling occurred between July 2000 and July 2002. After complete *BRCA1/2* screening and clarification of recurrent mutations of the first group, all consecutive index nonrelated patients counseled between September 2002 and March 2006 (a total of 157 patients) were prescreened for the two recurrent mutations observed. All samples from patients negative for the recurrent mutations were later included in the general *BRCA1/*2 mutation analysis.

General mutation analysis

DNA was extracted from whole blood using the Puregene Genomic DNA purification kit (Gentra System, Minneapolis, MN). DNA was amplified by PCR using primers [17] specific for the coding sequence and exon-intron boundaries of *BRCA1/2*. Mutation screening was performed by conformation-sensitive gel electrophoresis (CSGE) [18].

Samples negative for *BRCA1*/2 mutations were tested for *BRCA1* rearrangements using the multiplex ligation-dependent probe amplification assay (MRCHolland, Amsterdam, the Netherlands), following the manufacturer's protocol. Amplification products were analyzed with an ABI Prism 310 automatic sequencer using the Genescan software (Applied Biosystems, Foster City, CA).

Identification of the c.156_157insAlu recurrent mutation in exon 3 of *BRCA2*

DNA analysis. The Alu insertion in exon 3 (c.156_157insAlu) of *BRCA2* was identified by PCR with the following primers: 3F: 5'-GATCTTTAACTGTTCTGGGT CACA-3' and 3R: 5'-CCCAGCATGACACAATTAATGA-3'. The PCR product was visualized by agarose gel electrophoresis, and the expected 425–base pair (bp) DNA fragment and an extra approximately 800-bp fragment were identified in positive patients.

To specifically amplify the allele with the c.156_157insAlu mutation, we performed a first PCR with primers 3F and 3R, followed by a nested PCR with the following primers: 3AluF: 5'CGGATCACGAGGTCAGGA-3'; and 3AluR: 5'-GGTTTGGTTCGTAATTGTTGTTT-3'. Primer 3AluF was designed to recognize a

sequence in the Alu insertion, and primer 3AluR binds to exon 3. The nested PCR product (approximately 300 bp) was analyzed by agarose gel electrophoresis and detected only in the DNA of patients with the Alu insertion. PCR conditions are available upon request.

RNA analysis. Total RNA was extracted from peripheral-blood leukocytes using TRIzol reagent (Invitrogen, Paisley, United Kingdom) according to the manufacturer's protocol. cDNA was synthesized using random primers (Roche, Mannheim, Germany) and Superscript II reverse transcriptase (Invitrogen) according to the manufacturers' protocols. Primers (1FcDNA and 10RcDNA)3 were used to amplify the *BRCA2* coding region from exon 1 through exon 10.

Purification and sequencing analysis. PCR products were isolated and purified using QIAquick Gel extraction kit (Qiagen, Hilden, Germany). DNA sequencing was performed using the same primers for PCR and the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) in an automated sequencer ABI Prism 310 (Applied Biosystems).

Screening of *BRCA2* c.156_157insAlu

A routine three-step PCR procedure was optimized. The three steps were as follows: 1) PCR with primers 3F/3R; 2) nested PCR with primers 3AluF/3AluR; and 3) reverse transcriptase (RT) PCR and cDNA sequencing. RT-PCR was performed only for patients with the extra DNA fragment observed in step 1 and the fragment resulting from the nested PCR.

Haplotype analysis

Index cases and relatives were genotyped with microsatellite polymorphic markers flanking the *BRCA2* gene. As controls, thirty unrelated Portuguese individuals (from different areas of the country) were also genotyped, and allele frequencies were estimated.

The nine microsatellite markers used were localized in a 5.36-cM (2.9-Mb) region encompassing *BRCA2* (locus order: cen-D13S1246-D13S1229-D13S260-D13S1699-D13S1698-D13S1701-D13S171-D13S1695-D13S1493-tel) on chromosome 13q12.3-13q13.2 [19, 20]. Fluorescently labeled primers were used to amplify the microsatellite polymorphic regions (PCR conditions are available upon request). PCR products were analyzed in an automated sequencer ABI Prism 310 using the 310 Genescan 3.1.2 software (Applied Biosystems). Allele sizes are given as the size of the PCR amplicons containing the microsatellites. One internal *BRCA2* polymorphism (H372N) was also screened through direct sequencing.

Estimation of founder mutation age

The age of the c.156_157insAlu mutation in generations (G) was calculated using the following equation: $G = \log\delta/\log(1-\theta)$. The linkage disequilibrium measure (δ) between the mutation and each of the closest recombinant microsatellite markers, D13S1698 and D13S1701, was calculated as $\delta = (Pd - Pn)/(1 - Pn)$, with Pd being the frequency of the

ancestral microsatellite allele among the chromosomes carrying the mutated *BRCA2* and Pn being the frequency of that microsatellite allele on chromosomes not carrying the mutation. The symbol θ represents the recombination fraction between a marker and the gene [21]. The genetic distances were inferred from the GDB [20] and Ensembl databases [19].

RESULTS

Probands

Characteristics of all probands submitted to genetic screening are listed in Table 1. In the first group, five different *BRCA1* and 11 *BRCA2* mutations were diagnosed in 19 of 53 patients, corresponding to a 35.9% detection rate. The following two*BRCA2* mutations were recurrent: c.156_157insAlu (three nonrelated families) and c.7208_7211del4 (two nonrelated families; Table 2). The second mutation was only detected after CSGE analysis, but c.156_157insAlu was immediately identified by agarose gel electrophoresis (Fig 1). The 157 additional consecutive families were then screened for the recurring mutations observed in the first set.

Identification of the *BRCA2* c.156_157insAlu recurrent mutation

In probands with the c.156_157insAlu *BRCA2* mutation, PCR amplification of exon 3 originated, besides the expected 425-bp product, an aberrant fragment with approximately 800 bp (Fig 1A). Sequencing of this fragment with primers 3F and 3R (Figs 1B and 1C) revealed an unknown sequence starting in nucleotide 156 of*BRCA2* cDNA. Primers 3AluF and 3AluR were designed to sequence the whole fragment that, after a Basic Local Alignment Search Tool search [24], revealed an Alu motif insertion [25, 26], subtype Ya5.

Amplification of cDNAs from two c.156_157insAlu–positive probands with primers 1FcDNA and 10RcDNA revealed the expected 1,300-bp product and an abnormal 1,100-bp band (Fig 1D). Sequencing of this abnormal 1,100-bp product revealed the in frame deletion of exon 3, which resulted in the fusion of exon 2 with exon 4 (Fig 1E). Exon 3 of *BRCA2* encodes a transcriptional activation domain [24], and its relevance in the tumor suppression function of *BRCA2* has been previously described [3].

Screening for the *BRCA2* c.156_157insAlu mutation

One hundred fifty-seven consecutive nonrelated high-risk individuals were screened for *BRCA2* c.156_157insAlu, which was observed in 14 additional individuals (one of whom was a woman with breast cancer who belonged to the family described by Teugels et al [7]). The three-step PCR procedure described was used for this screening.

Table 1. Phenotypic characterization of 210 probands

Characteristic	First Set of Probands (n = 53)	Second Set of Probands (n = 157)	Total (N = 210)
Sex, No.			
Male	15	8	23
Female	38	149	187
Age, years			
Male			
Mean	64	62	63
Range	42-83	46-73	42-83
Female			
Mean	49	48	48
Range	34-77	20-82	20-82
Female breast cancer, No.	42	152	194
Age at diagnosis, years			
Mean	44	42	43
Range	28-74	18-69	18-74
Ovarian cancer, No.	7	11	18
Age at diagnosis, years			
Mean	47	50	49
Range	14-61	27-82	14-82
Male breast cancer, No.	16	7	23
Age at diagnosis, years			
Mean	60	60	60
Range	39-74	45-72	39-74
Prostate cancer, No.	1	1	2
Age at diagnosis, years			
Mean	66	59	62
Range	—	—	59-66
Family phenotypes, %			
Female breast cancer	47	69	64
Male breast cancer	34	6	13
Breast/ovarian cancer	13	13	13
Heterogeneous*	4	10	9
Ovarian cancer specific	2	1	1

Family phenotype was classified as heterogeneous if no more than two cases of breast cancers were observed and other cancers known to be associated with*BRCA2* were present, such as gastric cancers, multiple myeloma, and melanoma.

Table 2. *BRCA* mutations identified in 53 Portuguese breast cancer families

Nucleotide Change (designation in BIC Database)	Mutation Type	Mutation Status*	Proband No.	Proband Phenotype	Age at Diagnosis (years)	Family Phenotype
BRCA1						
c.211A>G (R71G)	S	BIC:20 Founder [13] (Galician)	43	BC	37	BCF
c.536delA	F	Novel	4	BBC OC	28, 35 63	BC/OCF
g.Ex11_Ex15del	F	Novel	34	OC	57	BC/OCF
g.Ex13ins6kb (exon13 ins6kb)	F	BIC:10 Founder [22] (unknown origin)	12	BC	41	BC/OCF
c.5263_5264insC (5382insC)	F	BIC:837 Founder [11] (Ashkenazi)	50	BBC	37	BCF
BRCA2						
c.156_157insAlu (384insAlu)	IFD	BIC:1 Founder (Portuguese)	24	PC; MBC	66; 74	MBCF
			40	BBC	35, 38	BCF
			52	MBBC	52, 54	MBCF
c.658_659del2 (886delGT)	F	BIC: 25	9	BC	38	BCF
c.1310_1311del2	F	Novel	49	BBC	47, 53	BCF
c.1369_1370ins2	F	Novel	11	BBC	33, 35	BCF
c.1423G>T	N	Novel	41	MBC	73	MBCF
c.1786G>C (D596H)	M	BIC:32	27	MBC	65	MBCF
c.2808_2811del4 (3036del4)	F	BIC:78	93	MBC	66	MBCF
c.6037A>T (K2013X)	N	BIC:10	35	BC; OC	54; 58	BC/OCF
c.6468_6469del2 (6696delTC)	F	BIC:17	7	MBC	65	MBCF
c.7208_7211del4 (7436del4)	F	BIC:1	16	BC	41	BCF
			92	BBC	52, 54	MBCF
c.9098_9099insA (9326insA)	F	BIC:18	32	MBC	56	MBCF

NOTE. The mutations observed more than once (c.156_157insAlu and c.7208_7211del4) were screened in 157 consecutive families. Mutation c.7208_7211del4, screened by conformation-sensitive gel electrophoresis, was not further observed. Mutation nomenclature is according to last revision [23].
Abbreviations: BIC, Breast Cancer Information Core; S, splice; BC, breast cancer; BCF, female breast cancer family; F, frameshift; BBC, bilateral breast cancer; OC, ovarian cancer; BC/OCF, breast and ovarian cancer family; IFD, in frame deletion; PC, prostate cancer; MBC, male breast cancer; MBCF, male breast cancer family; MBBC, male bilateral breast cancer; N, nonsense; M, missense.
* Number of times described in the BIC Database.

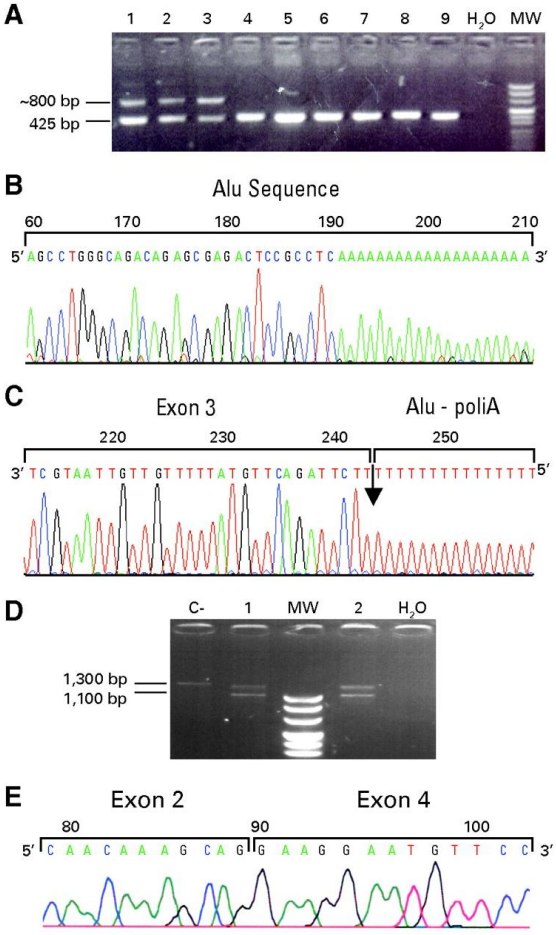


Figure 1. Polymerase chain reaction (PCR) of *BRCA2* exon 3 and reverse transcription (RT) PCR of exons 2 to 9; the approximately (A) 800–base pair (bp) fragment observed in samples 1 to 3 was sequenced with primers (B) 3AluF and (C) 3AluR. (D) The negative control (C–) has the expected RT-PCR product (1,300 bp), but samples 1 to 2 (c.156_157insAlu positive) display an additional band (approximately 1,100 bp) the sequencing of which (E) revealed exon 3 skipping.

Considering the three of 53 c.156_157insAlu–positive families initially observed and these additional 14 families (of 157), 8% of all families tested were positive for the*BRCA2* rearrangement. The following four different phenotypes could be distinguished: female breast (nine families, 53%), male breast (four families, 24%), breast/ovarian (two families, 12%) and heterogeneous (two families, 12%). These latter families had no more than two female breast cancers and no ovarian or male breast cancer, but other neoplasias like gastric cancer (two members in family 108), head and neck cancer (one member in family 108 and two members in family 129), and multiple myeloma (one member in family 108) were observed. Examples of representative phenotypes are shown in Figure 2, and

phenotypic characteristics of all index patients and their families are shown in Appendix Figure A1.

Phenotypes of c.156_157insAlu–positive and –negative families are listed in Table 3. Among the 210 families, 28 had male breast cancer, and 12 were diagnosed with a *BRCA2* mutation (four families were positive for the c.156_157insAlu mutation, two were positive for the c.9098_9099insA mutation, and the remaining six families showed six different *BRCA2* mutations: c.1423G>T; c.1786G>C; c.2808_2811del4; c.5063_5066del4; c.6468_6469del2; and c.7208_7211del4). These data demonstrate that 33% of all male breast cancer families with a *BRCA* mutation identified are c.156_157insAlu positive.

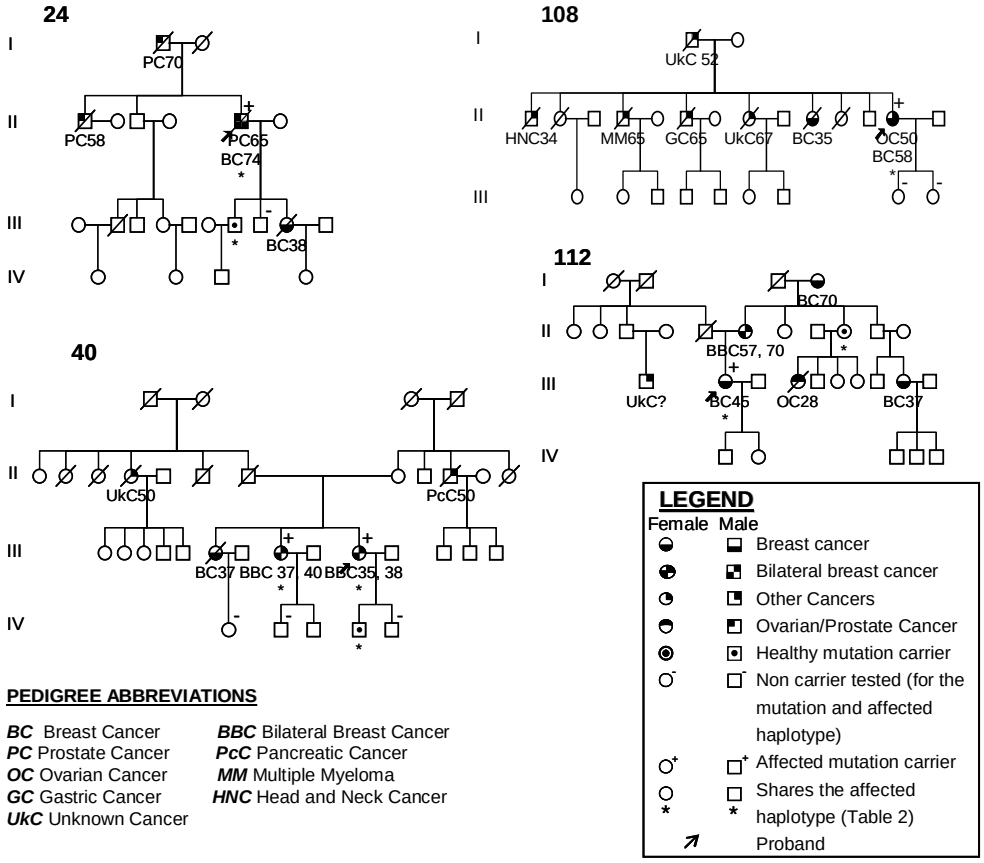


Figure 2. Representative pedigrees of c.156_157insAlu families. One example of each phenotype observed in these families is shown, including male breast cancer (family 24), breast cancer (family 40), heterogeneous (family 108), and breast/ovarian cancer (family 112). Results of c.156_157insAlu screening and segregation of the common haplotype are also shown.

Table 3. Phenotypic characterization of families positive and negative for c.156_157insAlu

	c.156_157insAlu Positive		c.156_157insAlu Negative		Total	
Family Phenotype	No. of Families	%	No. of Families	%	No. of Families	%
Breast cancer	9	53	125	65	134	64
Breast/ovarian cancer	2	12	25	13	27	13
Male breast cancer	4	23	24	12	28	13
Heterogeneous	2	12	17	9	19	9
Ovarian cancer	0	0	2	1	2	1
Total	17	100	193	100	210	100

Haplotype analysis and estimation of mutation age

Most of the families positive for the c.156_157insAlu mutation come from Central Portugal (Appendix Fig A2), and haplotype analysis was performed in 37 individuals (14 index patients and 23 of their relatives). The results of c.156_157insAlu screening in relatives of representative families are shown in Figure 2.

We observed a conserved haplotype cosegregating with the mutation (Table 4) and absent in noncarriers of these families (Fig 2). Encompassing *BRCA2*, it represents a 1.09-Mb interval from D13S260 to D13S1695, and its random population frequency is approximately one in 8,600. In two of 14 families, recombination events, either centromeric (marker D13S1698) or telomeric (marker D13S1701) to *BRCA2*, reduced the shared haplotype region to 0.63 Mb and 0.54 Mb, respectively (Table 4). Demonstration of a shared haplotype between all c.156_157insAlu families provides evidence for a common ancestry among these families.

Table 4. Haplotypes of probands with the *BRCA2* c.156_157 insAlu mutation

		Proband No.													
Marker	Position (Mb)	24	40	52	55	59	69	108	112	129	158	184	200	207	209
D13S1246	−1.78	197/199	199/207	195/201	203/205	197/203	197/201	203/205	201/203	199/201	201/205	201	205/207	199/203	199/205
D13S1229	−1.41	135/143	133/143	133/135	131/133	131/135	—	135	135	133	133/135	135	133	133/143	131
D13S260	−0.45	**160**/168	**160**	**160**/162	**160**/166	158/**160**	**160**/166	156/166	**160**	**160**	**160**/168	**160**	**160**	**160**/166	**160**
D13S1699	−0.28	**153**	**153**	**153**	150	**153**	**153**	150/**153**	**153**	**153**	**153**	**153**/159	**153**	150/**153**	150/**153**
D13S1698	−0.185	**156**	**156**	**156**/160	**156**/160	**156**	**156**/160	166	**156**	**156**	**156**	**156**/166	**156**/174	**156**	160
BRCA2	—	+	+	+	+	+	+	+	+	+	+	+	+	+	+
H372N	—	**C**	**C**	**C**	**C**/A	**C**/A	**C**	**C**/A	**C**/A	**C**/A	**C**/A	**C**/A	**C**/A	**C**/A	**C**/A
D13S1701	+0.171	287/**299**	**299**	295/**299**	291/**299**	291/**299**	**299**/303	**299**	287	295/**299**	283/**299**	291/303	295/**299**	291/**299**	291/**299**
D13S171	+0.28	**226**/236	**226**/236	**226**/236	**226**/236	222/**226**	**226**	**226**	222/238	**226**/236	**226**/228	**226**/236	**226**/236	222/**226**	**226**/236
D13S1695	+0.55	**238**/256	**238**/246	**238**/246	**238**/246	**238**/246	**238**/246	**238**/246	250/254	**238**	**238**	**238**	**238**/246	**238**/246	**238**/246
D13S1493	+1.04	231	231/235	231	223/227	235	227	223/239	235/239	231/239	223/231	227	219/235	223/231	223/227

NOTE. On the basis of estimated allele frequencies for each of the six shared markers (D13S260: allele 160; D13S1699: allele 153; D13S1698: allele 156; D13S1701: allele 299; D13S171: allele 226; D13S1695: allele 238) and the polymorphism (H372N: allele C), calculated from 30 unrelated controls, the random population frequency of this haplotype would be expected to be approximately one in 8,600. Shared alleles are indicated in bold.

The linkage disequilibrium measures (δ) between the *BRCA2* c.156_157insAlu mutation and each of the recombinant microsatellite markers D13S1698 and D13S1701 were calculated as $\delta = 0.7857$ and $\delta = 0.8137$, respectively. The recombination fraction θ was determined from the physical distances between markers D13S1698 and D13S1701 and the *BRCA2* gene. The distance for D13S1698 is 0.185 Mb (approximately 0.185 cM), and the distance for D13S1701 is 0.171 Mb (approximately 0.171 cM),19,29 assuming that 1 cM is equivalent to approximately 1 Mb [25, 26]. Given these genetic distances, we estimated that the c.156_157insAlu mutation occurred between 120 and 130 generations ago, that is, approximately 2,400 to 2,600 years ago, assuming 20 years per generation.

DISCUSSION

In this study, we demonstrate that the insertion of an Alu fragment in position 156 of *BRCA2* cDNA is a founder mutation of Portuguese origin. To date, only seven *BRCA2* rearrangements have been detected (including c.156_157insAlu) in unique families [2-7]. Therefore, besides the relevant implications in the genetic screening of breast/ovarian cancer families of Portuguese ancestry, our finding demonstrates that the c.156_157insAlu mutation is the most frequent *BRCA2* rearrangement described to date.

In our target population, including all of South Portugal, c.156_157insAlu has been observed in 8% of screened families, and taking into account families that have already been fully screened for *BRCA1/2*, it corresponds, approximately, to one in every six *BRCA*1/2 mutations identified. It is interesting to note that, in our families most extensively studied, not only was this rearrangement the most frequent genetic event observed, but also other rearrangements were diagnosed in negative families by CSGE. This observation reinforces the need to search for rearrangements in high-risk families negative for point mutations.

The contribution of Alu insertions as disease-causing mutations in humans has been estimated as approximately one mutation in 600 [27]. Before the c.156_157insAlu mutation, which was first reported in one family of Portuguese origin [7], only one *BRCA1/2* Alu insertion was described [6]. Our initial finding of c.156_157insAlu in three of 53 families raised the hypothesis of a founder effect in our population, and we implemented the prescreen of all high-risk consecutive families ascertained in our clinic for this mutation. Although the first description of this mutation, in a Portuguese family, was obtained by Southern blotting [7] after negative PCR-based screening, we optimized a three-step PCR for a quick and effective way to screen and confirm the presence of this rearrangement. A simple PCR reaction is enough to detect a positive individual. The Alu fragment is sequenced, and the pathogenic effect of the mutation (exon 3 skipping) is confirmed by RT-PCR. The clinical relevance of this strategy is that 8% of our families obtain a quick and less expensive result, without the need to wait for full screening. All c.156_157insAlu–negative families must be fully screened for *BRCA1/2* genes.

Using the three-step PCR, 157 consecutive families were screened for the c.156_157insAlu mutation, and an additional 14 nonrelated positive families were observed. Fifteen of these families came from the central part of Portugal, further supporting the possibility of a common ancestor. This hypothesis was confirmed through the observation of a conserved haplotype surrounding the *BRCA2* locus that was found to segregate with the mutation in 12 of 14 index patients. In the remaining two families, recombination events reduced the shared haplotype region.

Estimation of the age of the mutation suggested that the founder event occurred 2,400 to 2,600 years ago, that is, before the invasion of the Iberian Peninsula by the Romans and the Germanic warriors. At that time, the Lusitanians [28] inhabited the territory that now includes the districts of origin of most of these families. All migrations that involved Portuguese people in the following centuries were all outward of this territory, either to the western part of the country or abroad. This may explain the higher prevalence of the c.156_157insAlu mutation in that area and also allows us to speculate that the two families from the north were also originated from the same founder. This historical hypothesis is still acceptable even if the age of the mutation is being overestimated, either because of the fact that mutation rates of the microsatellite markers were not taken into account or because recombination events in two families were considered.

Breast and prostate cancer were the malignancies most frequently observed in c.156_157insAlu families, and the most frequent phenotype was female breast cancer, with a mean age at diagnosis of 48 years. Four male breast cancer families were also registered, and it is remarkable that 33% of all our male breast cancer families with *BRCA2* mutations harbor the founder mutation. This may reflect a particular phenotype associated to this mutation, or it may be the result of the high frequency of this genetic event in our *BRCA2* families.

Prostate cancer was the most frequent cancer diagnosis in men belonging to c.156_157insAlu families. Two of the male breast cancer probands had also been diagnosed with prostate cancer, and three of the eight prostate cancer patients registered were observed in only one family, with a mean age at diagnosis of 65 years (family 24). The association of *BRCA2* mutations with prostate cancer is well known [29, 30], not only in relatives of women with breast and ovarian cancer, but also in men unselected for family history but with early-onset disease [31]. The clustering of early-onset prostate cancer in families may be a result not only of the *BRCA2*mutation described but also of unknown modifier factors that may affect its penetrance in men.

Two families had heterogeneous phenotypes, with only two cases of breast cancers each but with two cases of head and neck and gastric cancers in one of the families and a case of multiple myeloma and head and neck cancer in the other family. Because the probands included in this study were the only affected patients alive, we cannot conclude that all of these tumors are related to the mutation, even though gastric cancer [32] and multiple myeloma [33] have been associated with *BRCA2*. Only two women were diagnosed with ovarian cancer, although they were diagnosed at an early age (23 and 38

years old). This observation may be explained by the fact that c.156_157insAlu is at the 5' region of the *BRCA2* gene [34]. Further follow-up of these families will clarify ovarian cancer incidence and age at diagnosis in positive women.

In conclusion, the results of this study suggest that screening of the c.156_157insAlu mutation should be extended to all high-risk breast/ovarian cancer families with Portuguese ancestry. Because most patients were from the central and southern part of the country, we cannot rule out the possibility that this mutation can be found in other regions of the country and also in other areas of the world, where Portuguese sailors, traders, and emigrants have settled since the 15th century to the present. Targeting the initial screening to the founder mutation in these families will clarify the global incidence of this mutation in a fast and inexpensive manner. More importantly, it will help to clarify the risk for several cancers in high-risk individuals and, through vigilance, preventive attitudes, or inclusion in clinical studies, help to modify the incidence and mortality by cancer in these families.

Author's Note

During the revision of this article, the c.156_157insAlu rearrangement was identified in three additional apparently nonrelated families.

ACKNOWLEDGMENTS

We are grateful to all families for their cooperation and to Valeriano Leite, MD, PhD, and Sylvie Mazoyer, PhD, for their comments on this article. We thank the support that Jorge Soares, MD, PhD, gave us on the implementation of the genetic screening of *BRCA1/2* mutations in our center, as well as Barbara L. Weber, MD, and Arupa Ganguly, PhD, whose advice was crucial for the development of the Breast Cancer Risk Evaluation Clinic and the methodologies of mutation screening.

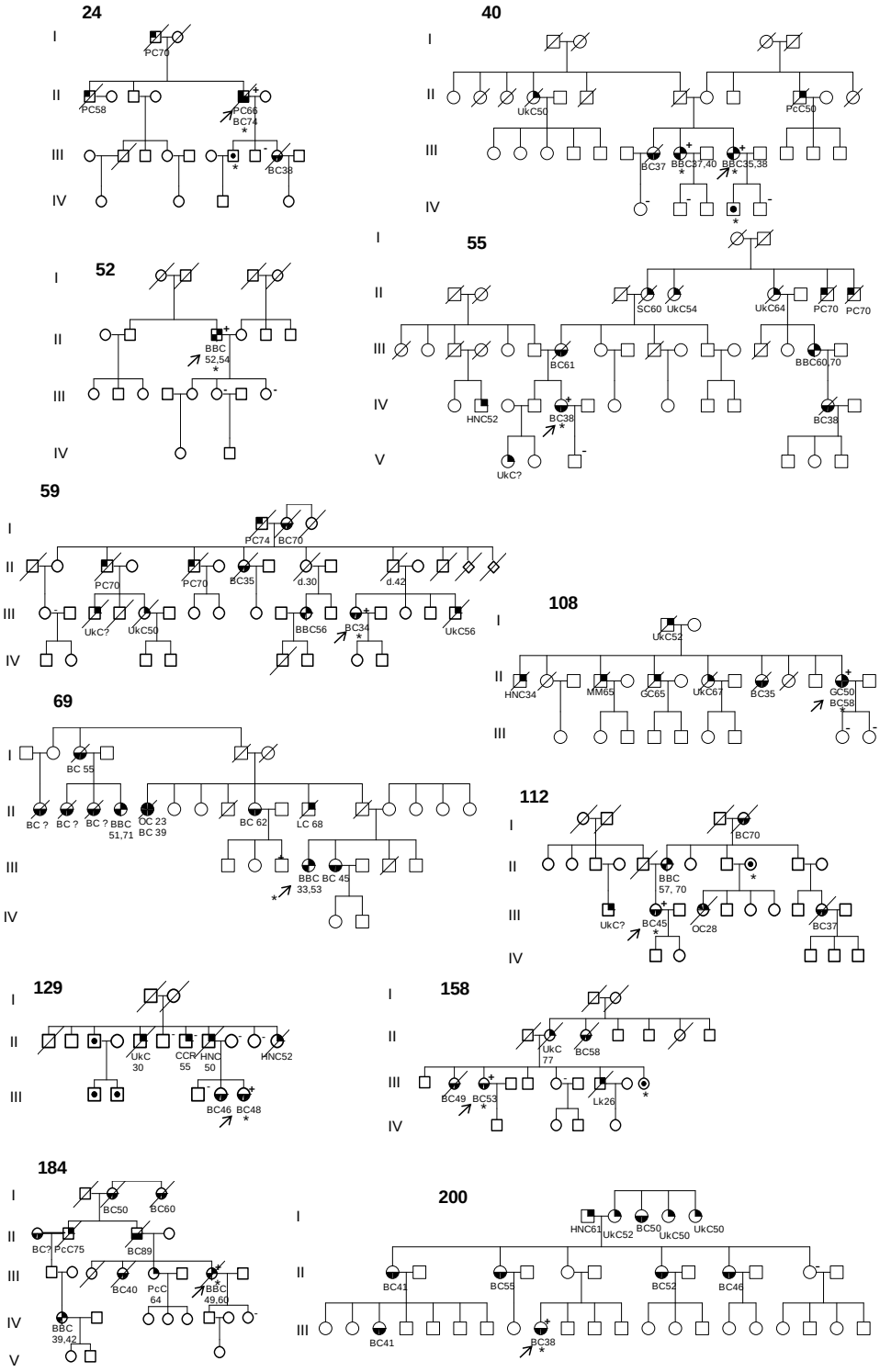
Supported by Grant No. 47320 from Serviço de Saúde e Desenvolvimento da Fundação Calouste Gulbenkian, Lisboa, and by grants (2000/2001 and 2004/2006) from Terry Fox, Núcleo Regional Sul da Liga Portuguesa Contra o Cancro, Lisboa, Portugal.

REFERENCES

1. Melcher, U. *Molecular genetics by Ulrich Melcher at Oklahoma State University: DNA rearrangements*. Available from: http://opbs.okstate.edu/~melcher/MG/MGW3/MG32.html.
2. Mazoyer, S., *Genomic rearrangements in the BRCA1 and BRCA2 genes.* Hum Mutat, 2005. **25(**5): p. 415-422.
3. Nordling, M., P. Karlsson, J. Wahlström, et al., *A large deletion disrupts the exon 3 transcription activation domain of the BRCA2 gene in a breast/ovarian cancer family.* Cancer Res, 1998. **58**(7): p. 1372-1375.
4. Tournier, I., B.B.-d. Paillerets, H. Sobol, et al., *Significant contribution of germline BRCA2 rearrangements in male breast cancer families.* Cancer Res, 2004. **64**(22): p. 8143-8147.
5. Wang, T., I. Lerer, Z. Gueta, et al., *A deletion/insertion mutation in the BRCA2 gene in a breast cancer family: a possible role of the Alu-polyA tail in the evolution of the deletion.* Genes Chromosomes Cancer, 2001. **31**(1): p. 91-5.
6. Miki, Y., T. Katagiri, F. Kasumi, T. Yoshimoto, and Y. Nakamura, *Mutation analysis in the BRCA2 gene in primary breast cancers.* Nat Genet, 1996. **13**(2): p. 245 - 247.
7. Teugels, E., S. De Brakeleer, G. Goelen, et al., *De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes.* Hum Mutat, 2005. **26**(3): p. 284-284.
8. Casilli, F., Z.C. Di Rocco, S. Gad, et al., *Rapid detection of novel BRCA1 rearrangements in high-risk breast-ovarian cancer families using multiplex PCR of short fluorescent fragments.* Hum Mutat, 2002. **20**(3): p. 218-226.
9. Hogervorst, F.B.L., P.M. Nederlof, J.J.P. Gille, et al., *Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method.* Cancer Res, 2003. **63**(7): p. 1449-1453.
10. Petrij-Bosch, A., T. Peelen, M.v. Vliet, et al., *BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients.* Nat Genet, 1997. **17**(3): p. 341-5.
11. Levy-Lahad, E., R. Catane, S. Eisenberg, et al., *Founder BRCA1 and BRCA2 mutations in Ashkenazi Jews in Israel: frequency and differential penetrance in ovarian cancer and in breast-ovarian cancer families.* Am J Hum Genet, 1997. **60**(5): p. 1059-67.
12. Neuhausen, S.L., *Founder populations and their uses for breast cancer genetics.* Breast Cancer Res, 2000. **2**(2): p. 77 - 81.
13. Vega, A., B. Campos, B. Bressac-de-Paillerets, et al., *The R71G BRCA1 is a founder Spanish mutation and leads to aberrant splicing of the transcript.* Hum Mutat, 2001. **17**(6): p. 520-521.
14. Welcsh, P.L. and M.-C. King, *BRCA1 and BRCA2 and the genetics of breast and ovarian cancer.* Hum Mol Genet, 2001. **10(**7): p. 705-713.
15. Parmigiani, G., D.A. Berry, and O. Aguilar, *Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2.* Am J Hum Genet, 1998. **62**(1): p. 145-158.
16. Frank, T.S., S.A. Manley, O.I. Olopade, et al., *Sequence analysis of BRCA1 and BRCA2: correlation of mutations with family history and ovarian cancer risk.* J Clin Oncol, 1998. **16(**7): p. 2417-25.
17. Rahman, N. and M.R. Stratton, *The genetics of breast cancer susceptibility.* Ann Rev Genet, 1998. **32**(1): p. 95-121.
18. Ganguly, A., M.J. Rock, and D.J. Prockop, *Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes.* Proc Natl Acad Sci USA, 1993. **90**(21): p. 10325-10329.
19. Peto, J., D.F. Easton, F.E. Matthews, D. Ford, and A.J. Swerdlow, *Cancer mortality in relatives of women with breast cancer: the OPCS study.* Int J Cancer, 1996. **65**: p. 275 - 283.
20. Hemminki, K., P. Vaittinen, and P. Kyyronen, *Age-specific familial risks in common cancers of the offspring.* Int J Cancer, 1998. **78**: p. 172 - 175.
21. Risch, N., D. de Leon, L. Ozelius, et al., *Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population.* Nat Genet, 1995. **9**(2): p. 152-9.

22. Mazoyer, S. and The BRCA1 Exon 13 Duplication Screening Group, *The exon 13 duplication in the BRCA1 gene is a founder mutation present in geographically diverse populations.* Am J Hum Genet, 2000. **67**(1): p. 207-212.
23. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.
24. Milner, J., B. Ponder, L. Hughes-Davies, M. Seltmann, and T. Kouzarides, *Transcriptional activation functions in BRCA2.* Nature, 1997. **386**(6627): p. 772-773.
25. Kerangueven, F., L. Essioux, and A. Dib, *Loss of heterozygosity and linkage analysis in breast-carcinoma: indication for a putative 3rd susceptibility gene on the short arm of chromosome 8.* Oncogene, 1995. **10**: p. 1023 - 1026.
26. Wajnrajch, M.P., J.M. Gertner, A.S. Sokoloff, et al., *Haplotype analysis of the growth hormone releasing hormone receptor locus in three apparently unrelated kindreds from the indian subcontinent with the identical mutation in the GHRH receptor.* Am J Med Genet A, 2003. **120A**(1): p. 77-83.
27. Kazazian, H.J., *An estimated frequency of endogenous insertional mutations in humans.* Nat Genet, 1999. **22**(2): p. 130.
28. de Motes, J., *Western Europe before the colonisations: The Iberian Peninsula and Southern Galia*, in *Historia Universal - O Mundo Mediterrânico e a Europa Continental da Pré-História ao Século III*, J. Saraiva, Editor. 1985, Publicações Alfa: Lisboa. p. 215-226.
29. The Breast Cancer Linkage Consortium, *Cancer Risks in BRCA2 Mutation Carriers.* J Natl Cancer Inst, 1999. **91**(15): p. 1310-1316.
30. Liede, A., B.Y. Karlan, and S.A. Narod, *Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature.* J Clin Oncol, 2004. **22**(4): p. 735-742.
31. Edwards, S.M., Z. Kote-Jarai, J. Meitz, et al., *Two percent of men with early-onset prostate cancer harbor germline mutations in the BRCA2 gene.* Am J Hum Genet, 2003. **72**(1): p. 1-12.
32. Jakubowska, A., K. Nej, T. Huzarski, R.J. Scott, and J. Lubinski, *BRCA2 gene mutations in families with aggregations of breast and stomach cancers.* Br J Cancer, 2002. **87**(8): p. 888-891.
33. Struewing, J.P., P. Hartge, S. Wacholder, et al., *The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi jews.* N Engl J Med, 1997. **336**(20): p. 1401-1408.
34. Thompson, D. and D. Easton, *Variation in cancer risks, by mutation position, in BRCA2 mutation carriers.* Am J Hum Genet, 2001. **68**(2): p. 410-419.

SUPPLEMENTARY DATA


24
I
II
III
IV
PC70
PC58
PC66
BC74
BC38
40
I
II
III
IV
UkC50
PCC50
BC37
BBC37,40
BBC35,38
52
I
II
III
IV
BBC
52,54
55
I
II
III
IV
V
SC60
UkC54
UKC64
PC70
PC70
BC61
BBC60,70
HNC52
BC38
BC38
UkC?
59
I
II
III
IV
PC74
BC70
PC70
PC70
BC35
d.30
d.42
UkC?
UkC50
BBC56
BC34
UkC56
108
I
II
III
UKC52
HNC34
MM65
GC65
UKC67
BC35
GC50
BC58
69
I
II
III
IV
BC 55
BC ?
BC ?
BC ?
BBC
51,71
OC 23
BC 39
BC 62
LC 68
BBC BC 45
33,53
112
I
II
III
IV
BC70
BBC
57, 70
UkC?
BC45
OC28
BC37
129
I
II
III
UkC
30
CCR
55
HNC
50
HNC52
BC46
BC48
158
I
II
III
IV
UkC
77
BC58
BC49
BC53
Lk26
184
I
II
III
IV
V
BC50
BC60
BC?
PCC75
BC89
BC40
PcC
64
BBC
49,60
BBC
39,42
200
I
II
III
HNC61
UkC52
BC50
UkC50
UkC50
BC41
BC55
BC52
BC46
BC41
BC38

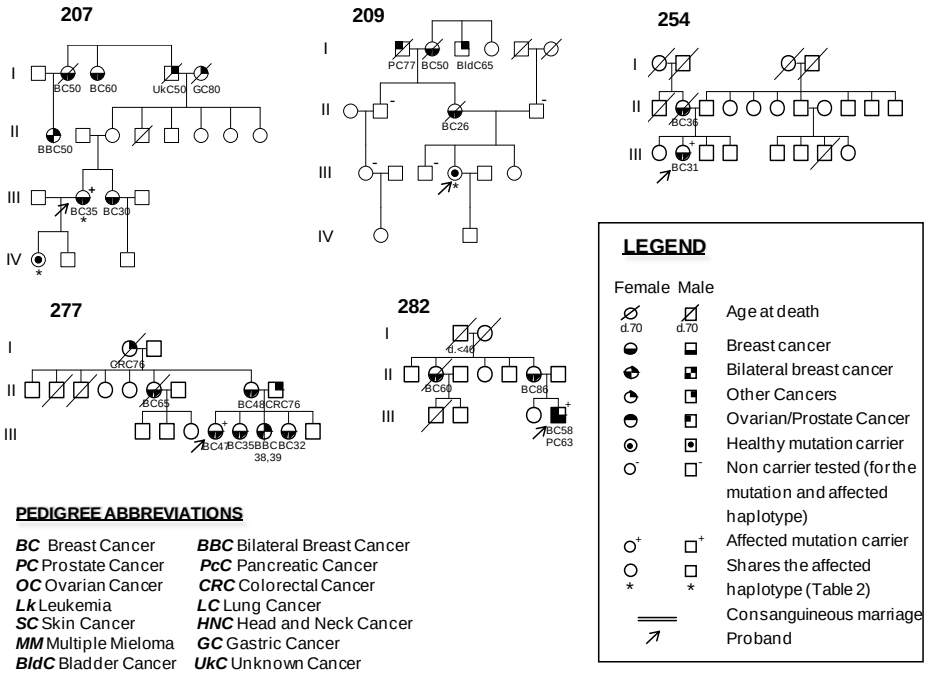


Figure A1. Pedigrees of c.156_157insAlu families. With the exception of families 24, 52, 184, and 282 (with male breast cancer), families 69 and 112 (with breast/ovarian cancer), and families 108 and 129 (with a heterogeneous phenotype) pedigrees revealed a predominance of female breast cancer. Results of c.156_157insAlu screening and segregation of the common haplotype are also shown (families 254, 277, and 282 were not included in haplotype analysis).

Figure A2. Map of Portugal indicating the origin of different families with the c.156_157insAlu rearrangement (●). Fifteen of these families came from the central part of Portugal (districts of Leiria, Santarém, Portalegre, and Évora).

CHAPTER 3.1

FGFR2 is a breast and ovarian cancer site risk modifier in *BRCA1/2* families

Rita D. Brandão, Marinus J. Blok, Andrea Romano, Jeske J.T. van Harssel, Maurice P. Zeegers, Encarna B. Gómez García

Submitted

ABSTRACT

BRCA1/2 mutation carriers show cancer site variability. Since breast and ovaries are the main target organs, we hypothesize that polymorphisms in genes involved in steroid hormone-mediated cell proliferation act as genetic risk modifiers. Therefore, we have studied two functional variants in the progesterone receptor (*PR*) gene: +331G/A and PROGINS and FGFR2, as well as variants described as modifiers of BC risk in the general population: TNRC9, *CASP8* -652 6N ins/del, and *CASP8* D302H.

The study included the probands of *BRCA1/2* mutation positive families and their relatives (both mutation carriers and non-carriers). In total, 548 women were screened (293 carriers, 255 non-carriers) from 124 *BRCA1/2* families (72 *BRCA1*, 52 *BRCA2*). Clinical parameters recorded included: (bilateral) BC, OC and age at cancer diagnosis.

The *FGFR2* variant showed a trend towards increased risk of bilateral BC (OR:1.59, 95%CI:0.926-2.729, p=0.093) which was significant among women above the age of 50. (OR:2.67, 95%CI:1.45-4.92, p=0.002). *FGFR2* was also significantly associated with decreased risk of OC (OR:0.52, 95%CI: 0.28-0.96, p=0.037) among women above 50. The variant alleles of *PR*, *TNRC9* and *CASP8* had no significant association with cancer site.

In our study, a functional polymorphism in the *FGFR2* gene was associated with bilateral BC and OC risks in women above 50 from *BRCA1/2* families. This marker may be relevant in genetic counselling of the women carrying a *BRCA1/2* mutation.

INTRODUCTION

Carriers of a mutation in the *BRCA1* or *BRCA2* genes are at high risk of developing breast (BC) and/or ovarian cancer (OC) and are eligible for risk-reducing interventions (mastectomy and adnexectomy) and/or intensive surveillance programs. Individualized advice about the most suitable intervention among the options available has been hampered by the considerable variability in the BC and OC risks observed amongst these women. It was suggested that the risk variability can be partially explained by the type and location of the *BRCA1* or *BRCA2* mutation. For instance, mutations that lie within the 3' third of the *BRCA1* gene would confer lower OC risk, whereas truncating mutations in a region of *BRCA2*, which has been named Ovarian Cancer Cluster Region (OCCR), correlate with higher risk of OC [1, 2]. However, within and between family cancer history among carriers with the same mutations, as shown that additional factors modify the risk. A Portuguese founder mutation in the *BRCA2* gene was described to give rise to different phenotypes depending on the families [3]. Some families had only female breast cancer cases, other had also male breast cancer, there were families with both breast and ovarian cancer cases and other families had a heterogeneous phenotype. Among Ashkenazi Jewish families, where three founder mutations account for the majority of the mutations identified, the site-specific cancer risk in a carrier can be better predicted by knowing the tumour sites in her relatives than by knowing the mutation [4]. These studies indicate the presence of modifier risk factors that segregate in the families. In addition, there is no explanation for the relatively high frequency of BC in women who tested negative for the familial mutation (i.e. the phenocopies), as compared with the general population [5]. Epidemiological studies have suggested that familial clustering of cancer can be better explained if additional genetic factors exist that may influence the cancer risks of individuals from *BRCA1/2* families [6].

In *BRCA1/2*-female mutation carriers, the main affected organs are the breast and the ovaries, which are hormone-related organs. Polymorphisms in the progesterone receptor (PR) gene have been reported to affect either the expression (+331G/A) or the activity (PROGINS) of the two receptor isoforms: PR-A and PR-B [7, 8]. Although both receptors bind to and mediate progesterone activity, the responses of ligand-activated PR-A and PR-B strongly depend on the cellular context. The G to A change at position +331 results in the introduction of a TATA-box, which exclusively enhances transcription of the PR-B isoform, thereby increasing the ratio between PR-B and PR-A. Furthermore, PROGINS, which consists of a haplotype of three genetic variations in complete linkage disequilibrium: G3432T, C3764T, and an Alu-S insertion in intron G, shows altered stability and transactivation activity in both PR-A and PR-B isoforms compared with the most common allele [8]. Our previous study, consisting of patients with a positive family history for BC and OC (n=220), of which only 23 were *BRCA1/2* mutation carriers, showed a marginal association of the +331A allele with OC (p=0.07) [9]. To confirm our previous results, we extended the study to *BRCA1/2* probands and their relatives (mutation carriers

and non-carriers) and we included other polymorphisms reported to act as modifiers of BC risk in sporadic cases: *FGFR2*, *TNRC9*, and *CASP8* -652 6N ins/del and *CASP8* D302H.

In this study we have used family members as controls. This allows to control, at least for some extent, for genetic background, since family members partly share their genes. Consequently, using family relatives allows to reduce the residual noise variance and enhance the power to detect relevant genetic risk modifiers.

METHODS

Patients

The study included female probands and their female relatives from 124 unrelated families (72 BRCA1 and 52 BRCA2) who had undergone genetic counselling and DNA-testing at our Centre. The methods used for *BRCA1/2* mutation screening were described previously [10]. In total, 548 women, from 124 families, were genotyped (293 carriers, 255 non-carriers). Four relatives per family were genotyped on average (proband excluded). There were 24 probands with no relatives tested, whereas, there were two large families (20 and 22 relatives tested). Clinical records from each woman were examined for cancer history and age at diagnosis of: breast cancer (BC), bilateral BC (BBC), ovarian cancer (OC) or other cancer sites.

Genotyping

The polymorphisms genotyped were: +331G/A (rs10895068), PROGINS (rs1042838), FGFR2 (rs2981582), TNRC9 (rs3803662), CASP8 D302H (rs1045485) and CASP8 -652 6Nins/del (rs3834129). Primers for pyrosequencing analysis were designed using the PSQ Assay Design software version 1.0.6 (Biotage). Primers and amplification conditions are available on request. Samples were analyzed on a PSQ HS 96A system (Biotage) using Pyro Gold SNP reagents kit (Biotage) according to the manufacturer's instructions.

Statistical analysis

The allele and genotype frequencies of each SNP were assessed for BC, BBC and OC phenotypes. Mutation carriers with cancer above the age of 80 were included in the non-affected phenotype (n=1). The statistical analysis was performed with Stata 10. Deviation from Hardy-Weinberg equilibrium for the genotypes of all markers and linkage disequilibrium between marker alleles were tested by X^2 tests within the group of non-carriers without cancer for all SNPs. We calculated odds ratios (OR) and corresponding 95% confidence intervals (95%CI) using logistic regression for allelic and genotypic analysis of the six SNPs. For allelic analyses, robust standard errors were calculated to model familial clustering of alleles within individuals. Analyses were performed among the whole group (n=548), adjusting the analysis for age and BRCA1/2 mutation status, and for

a subgroup that included only the women older than 50 (age at last genetic counseling) (n=196).

RESULTS

Clinical features and genotype frequencies of the population studied

From the 548 women that were genotyped in this study, 293 were *BRCA1/2* mutation carriers, whereas 255 were non-carriers (Table1). Within this population, 27% had BC (mean age: 44, range: 26-76), 6% had bilateral BC (2^{nd} BC, mean age: 47, range: 33-71), 6% had OC (mean age: 53, range: 24-75) and 2% had both BC and OC. The clinical features of the women included in the study according to their mutation status are also presented in Table 1. Among the 116 non-carriers from *BRCA1* families 5% had BC, 3% had BBC and 1% BC and OC, and among the 139 non-carriers from *BRCA2* families 3% had BC and 1% BBC.

The observed allele frequency of the SNPs was the following: *FGFR2*: C- 61,4%, T- 38,6%; PROGINS: G- 88,1% and T- 11,9%; +331G/A: G- 93,6% and A- 6,4%; TNRC9: C- 73.7% and T- 26.3%; CASP8 652 6Nins/del: ins-51% and del- 49% and CASP8 D302H: G- 87.1% and C- 12.9%. All the SNPs were within Hardy-Weinberg Equilibrium.

Genotype-phenotype association

PR polymorphisms

The polymorphisms in the *PR* gene were not associated with BC risk (or bilateral BC), as shown in Table 2. We also did not find evidence of an association between +331G>A or PROGINS and OC risk (Table 3). Given our current sample size, for the SNP +331G>A, the study was only sufficiently powered to detect allelic odds ratios (OR) of about 2.04 and 3.33 or higher for breast and ovarian cancer, respectively, assuming a power of 80%, a type-1 error rate of 5% and a binominal distribution. Under the same assumptions, for the PROGINS polymorphism we would only been able to detect odds ratios higher than 1.65 or 2.37 for breast and ovarian cancer, respectively.

Table 1. Clinical characteristics of the women from the study according to *BRCA1/2* mutation status

	carriers (%)	mean age at diagnosis (range)	non-carriers (%)	mean age at diagnosis (range)
BRCA1				
BC	48 (34%)	41 (26-65)	6 (5%)	55 (41-75)
BBC	16 (11%)	1st: 36 (32-54) 2nd: 42 (33-54)	3 (3%)	1st: 55 (47-61) 2nd: 59 (50-71)
OC	11 (8%)	49 (38-63)	0	
BC + OC	5 (4%)	BC: 48 (29-60) OC: 57 (47-61)	1 (1%)	BC: 60 OC: 75
BBC+OC	2 (1%)	1st BC: 32 (30-34) 2nd BC: 44 (33-54)	0	
Other cancer*	6 (4%)		3 (3%)	
no cancer	54 (38%)		103 (89%)	
TOTAL	142 (100%)		116 (100%)	
BRCA2				
BC	47 (31%)	47 (28-76)	4 (3%)	54 (48-61)
BBC	8 (5%)	1st: 47 (39-53) 2nd: 52 (40-60)	1 (1%)	1st: 43 2nd: 46
OC	10 (7%)	54 (24-71)	0	
BC + OC	1 (1%)	BC: 49 OC: 66	0	
BBC+OC	1 (1%)	1st BC: 55 2nd BC: 55	0	
Other cancer*	2 (1%)		2 (1%)	
no cancer	82 (54%)		132 (95%)	
TOTAL	151 (100%)		139 (100%)	

* Other Cancer sites include: colon, cervix, endometrium, skin and lung.

FGFR2

The minor allele of *FGFR2* rs2981582 (T allele) did not significantly affect the risk of unilateral BC in our study. However, the T allele showed a trend for association with the development of BBC (p=0.093), which became clearly significant (p=0.002) within the subgroup of older women (defined as those who were above the age of 50 when they were counselled) with an OR of 2.67 (95%CI:1.45-4.92), as presented in Table 2. Furthermore, in the group of women counselled above the age of 50, carriers of the T allele of *FGFR2* had significant less risk of OC (OR:0.52, 95%CI:0.28-0.96, p=0.037) as it is shown in Table 3.

Table 2. Genotype-phenotype associations: breast cancer

SNP	Per allele OR		95% CI	p-value
BC				
PR +331G/A (rs10895068)	A allele	1.237	0.633-2.420	0.534
PR PROGINS (rs1042838)	T allele	1.265	0.744-2.151	0.385
FGFR2 (rs2981582)	T allele	1.149	0.854-1.545	0.359
TNRC9 (rs3803662)	T allele	0.852	0.586-1.237	0.399
CASP8 D302H (rs1045485)	C allele	0.944	0.479-1.861	0.869
CASP8 -652 6Nins/del (rs3834129)	Del allele	0.809	0.597-1.096	0.172
Bilateral BC				
PR +331G/A (rs10895068)	A allele	1.158	0.349-3.844	0.811
PR PROGINS (rs1042838)	T allele	0.650	0.290-1.457	0.296
FGFR2 (rs2981582)	T allele	1.590	0.926-2.729	0.093
FGFR2 (rs2981582) *	T allele	2.67	1.450 -4.920	0.002
TNRC9 (rs3803662)	T allele	1.205	0.655-2.215	0.549
CASP8 D302H (rs1045485)	C allele	0.726	0.192-2.748	0.638
CASP8 -652 6Nins/del (rs3834129)	Del allele	0.900	0.535-1.515	0.693

*Analysis performed in a subpopulation group consisting only of women above the age of 50, n=179

TNRC9 and *CASP8* polymorphisms

It can be seen from the data in Tables 2 and 3 that the variant alleles of *TNRC9, CASP8* D302H, and *CASP8* -652 6N ins/del were not significantly associated with (bilateral) breast or ovarian cancer in our population. With the sample size used in this study, we can detect allelic odds ratios of 1.19 (for BC) and 1.34 (for OC) or higher with minor allele frequencies of 50% (e.g. *CASP8* -652 6N ins/del) or lower (e.g. *TNRC9* and *CASP8* D302H) for breast and ovarian cancer, respectively, assuming a power of 80%, a type-1 error rate of 5% and a binominal distribution.

DISCUSSION

This is the first association study performed within *BRCA1/2* families that attempts to clarify the underlying genetic risk factors behind the individual and familial cancer propensity for either breast or ovaries. This research assessed the influence of relevant genetic variants in four genes, *PR, FGFR2, TNRC9,* and *CASP8,* in the phenotypic variability observed in cancer risks among individuals and families with a known deleterious *BRCA1/2* mutation.

Table 3. Genotype-phenotype associations: ovarian cancer.

SNP		Per allele OR	95% CI	p-value
PR +331G/A (rs10895068)	A allele	1.392	0.606-3.196	0.436
PR PROGINS (rs1042838)	T allele	0.664	0.264-1.669	0.384
FGFR2 (rs2981582)	T allele	0.760	0.402-1.436	0.398
FGFR2[a] (rs2981582)	T allele	0.52	0.28-0.96	0.037
TNRC9 (rs3803662)	T allele	.066	0.557-2.040	0.847
CASP8 D302H (rs1045485)	C allele	.400	0.111-1.442	0.162
CASP8 -652 6Nins/del (rs3834129)	Del allele	1.279	0.710-2.303	0.412

[a]Analysis performed in a subpopulation group consisting only of women above the age of 50, n=179

PR polymorphisms

In a previous study, which consisted of 211 patients with a positive family history for breast/ovarian cancer, with only a minority being from *BRCA1/2* mutation families (n=23), we found that the association of the rare allele +331A with OC was of borderline significance [9]. Risch *et al.* [11] have found that the minor allele +331A associates with increased risk of epithelial OC in postmenopausal women, while others have found an association with OC among women under the age of 51 [12] or within the endometrioid subtype of OC [13]. No significant association was found between the +331G/A SNP and the risk of BC in our study, which confirms previous results from other studies [14, 15]. However, the association of the SNP +331G/A with the risk of BC and OC within *BRCA1/2* mutation carriers has not been subject of investigation in those studies. The results of our study do not confirm an association of the polymorphism +331G/A with the risk of OC. This may be due to lack of power, since our study had only enough power to detect allelic odds ratios of 3.33 or higher, which is similar to those observed in some of the previous studies [9, 12]. However, we lacked power to detect ORs in the range of those reported in other studies [11, 13] and in our own study. Therefore, further studies in a larger population of *BRCA1/2* carriers would be relevant to establish the risk associated with this polymorphism.

We also did not find any significant associations between PROGINS and the BC risks, which is in agreement with other association studies on BC [9, 12, 15-18]. PROGINS was also not associated with OC in our study. We had only enough power to detect OR =0.42 or lower. Our result is in agreement with the study of Pierce *et al*. [13], the largest association study performed so far. Conversely, some studies did find an association of PROGINS with risk of sporadic OC, as the study of Leite *et al*. [19] and references therein.

The ORs reported vary between 2.2 and 4.5, so it is possible that our study was underpowered.

The above-mentioned discordant results amongst studies may rely on differences related to the population studied, such as age, histology of the tumour, or history of use of oral contraceptives.

FGFR2

Easton *et al*. have reported increased risk of BC in sporadic cases among the women carrying the *FGFR2* rare allele in a genome-wide association study [20]. This SNP lies in intron 2 of *FGFR2*, close to putative transcription-factor binding sites, the minor allele of this variant increases FGFR2 expression in breast tumours [21]. FGFR2 and PR interact with each other, i.e., upon binding of FGF-2, FGFR2 activates the PR, which in turn stimulates mammary tumour growth [22]. Interestingly, FGFR2 is not only present in breast cells, but also in ovarian tissue [23]. The effect of this SNP in *BRCA1/2* mutation carriers was studied by Antoniou *et al*. [24]. The authors reported that the *FGFR2* rare allele also increased risk of BC within *BRCA1/2* mutation carriers [24] as well as in women with a first-degree relative with BC [20]. The authors reported a significant association between the SNP rs2981582 and BC in *BRCA2* but not in *BRCA1* carriers. This study included 489 Dutch *BRCA1* carriers but no *BRCA2* carriers. However, among these Dutch *BRCA1* carriers no significant association with BC was observed [24]. In another study, an association between the minor allele and increased risk of BC was also found among women with a family history of breast cancer but without *BRCA1/2* mutations. However, as in our study, no significant association was found within women carrying B*RCA1/2* mutations [25].

Regarding *FGFR2* and the risk of BBC in our study, there was a trend towards an association with increased risk of BBC (p=0.093). Since the mean age of developing the 2nd BC in our study was 47 years, we did this analysis in the subgroup of older women (above age 50), which revealed that the trend found in the total population was due to the subgroup of older women (p= 0.002). Easton *et al*. also observed that the minor allele of *FGFR2* increased the risk of BBC in their study on sporadic BC [20]. Unfortunately, neither the study of Antoniou *et al.* nor the study of Latif *et al*. have analyzed separately the BBC cases in *BRCA1/2* mutation carriers [24, 25]. The fact that this SNP was only relevant for the development of the second tumour may reflect that the effect of hormone-related risk factors is age-dependent and most prominent in the peri- and post-menopausal period.

In contrast to its effect on the risk of BC, we found that the rare T allele of *FGFR2* SNP rs2981582 significantly decreased the risk of OC amongst the group of women older than 50. The fact that the association with OC was only evident among older women is also not surprising since the mean age of OC was 53 years. Three other studies, one performed among *BRCA1/2* mutation carriers (52 OC cases) [26], among 1,905 sporadic epithelial OC cases [27], and the other among 2,513 invasive OC cases did not observe an association

between this SNP and the risk of OC. Due to the relevance of our results, we suggest to clarify this in a larger population of *BRCA1/2*-mutation carriers.

As FGFR2 activates the PR [22, 28], it may be that the observed contrasting tissue-specific effects of FGFR2 in BC and OC are mediated through this receptor. It is known that progesterone exposure and subsequent PR activation are protective for endometrial cancer and OC, whereas it is a risk factor for BC [29-31].

TNRC9 and *CASP8* polymorphisms

In addition to *PR* and *PR* activator FGFR2, we have also studied the SNP rs3803662 *TNRC9* and two polymorphisms in the *CASP8* gene *(*-652 6N ins/del *CASP8* D302H). *TNRC9* was initially found to be a risk factor for BC in the same genome-wide association study which identified the *FGFR2* SNP [20], and later it was found to increase the risk of BC within *BRCA1/2* carriers [24]. Polymorphisms in *CASP8* were studied by candidate-gene approaches. *CASP8* -652 6N ins/del and *CASP8* D302H are associated with decreased risk of several tumours in the general population [32-34]. *CASP8* D302H was found to be also associated with decreased risk of BC in familial cases without *BRCA1/2* mutation [25]. *CASP8* D302H, *CASP8* -652 6Ndel allele and *TNRC9* SNP were not associated with BC or OC risk in our study, as opposed to previous reports in sporadic BC cases [20, 32, 33] and in familial cases [17]. Latif *et al.* reported an association of the *TNRC9* variant with a protective effect of OC in *BRCA1/2* mutation positive cases, whereas the protective effect of *CASP8* D302H for OC was only observed among the *BRCA1/2* mutation negative cases [26].

In conclusion, functional genetic variants in the *FGFR2* gene contribute to site-specific cancer risk observed in members of *BRCA1/2* families. *FGFR2* SNP rs2981582 (T allele) was associated both with an increased risk of BBC and with a decreased risk of OC in our families (in women above the age of 50). These contrasting effects may be explained by the fact that FGFR2 activates PR, which has tissue-specific effects. We do not find evidence for an association of PR polymorphisms with disease risk, most probably due to the fact that our sample size is relatively small for the low frequency of the variants in the population and the effect sizes observed (ORs < 1.65). This indicates that the study was underpowered to detect associations for most of the candidate polymorphisms. We suggest to further investigate the OC and BBC risk *for the FGFR2* and *PR* polymorphisms in larger *BRCA1/2* populations, as they may be useful for an individualized assessment of the cancer risks of these patients. In order to improve these association studies, we also suggest that families are used in subsequent analysis to demonstrate linkage of the site-specific cancer with the polymorphisms. This would allow to test the robustness of the cancer-site prediction in the presence of the candidate genetic risk modifier variants.

ACKNOWLEDGEMENTS

We thank the Kerry Stichting for financing the laboratory tests for the study.

REFERENCES

1. Gayther, S.A., et al., *Germline mutations of the BRCA1 gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation.* Nat Genet, 1995. **11**: p. 428-433.
2. Gayther, S.A., et al., *Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene.* Nat Genet, 1997. **15**: p. 103-105.
3. Machado, P.M., et al., *Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes.* J Clin Oncol, 2007. **25**(15): p. 2027-2034.
4. Simchoni, S., et al., *Familial clustering of site-specific cancer risks associated with BRCA1 and BRCA2 mutations in the Ashkenazi Jewish population.* Proc Natl Acad Sci USA, 2006. **103**: p. 3770-3774.
5. Goldgar, D., et al., *BRCA phenocopies or ascertainment bias?* J Med Genet, 2007. **44**: p. e86.
6. Antoniou, A.C. and D.F. Easton, *Models of genetic susceptibility to breast cancer.* Oncogene, 2006. **25**: p. 5898-5905.
7. Vivo, I.D., et al., *A functional polymorphism in the promoter of the progesterone receptor gene associated with endometrial cancer risk.* Proc Natl Acad Sci U S A, 2002. **99**: p. 12263-12268.
8. Romano, A., et al., *The PROGINS polymorphism of the human progesterone receptor diminishes the response to progesterone.* J Mol Endocrinol, 2007. **38**: p. 331-350.
9. Romano, A., et al., *Impact of two functional progesterone receptor polymorphisms (PRP): +331G/A and PROGINS on the cancer risks in familial breast/ovarian cancer.* Open Cancer J, 2007. **1**: p. 1-8.
10. Gómez García, E., et al., *Patients with an unclassified genetic variant in the BRCA 1 or BRCA2 genes show different clinical features from those with a mutation.* J Clin Oncol, 2005. **23**: p. 2185-2190.
11. Risch, H.A., et al., *PGR +331 A/G and increased risk of epithelial ovarian cancer.* Cancer Epidemiol Biomarkers Prev, 2006. **15**: p. 1738-1741.
12. Romano, A., et al., *Two functionally relevant polymorphisms in the human progesterone receptor gene (+331 G/A and progins) and the predisposition for breast and/or ovarian cancer.* Gynecol Oncol, 2006. **101**: p. 287-295.
13. Pearce, C.L., et al., *Progesterone receptor variation and risk of ovarian cancer is limited to the invasive endometrioid subtype: results from the ovarian cancer association consortium pooled analysis.* Br J Cancer, 2008. **98**(2): p. 282-288.
14. Feigelson, H.S., et al., *No Association between the Progesterone Receptor Gene +331G/A Polymorphism and Breast Cancer.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(6): p. 1084-1085.
15. Vivo, I.D., et al., *The progesterone receptor Val660->Leu polymorphism and breast cancer risk.* Breast Cancer Res, 2004. **6**(6): p. R636-R639.
16. Spurdle, A.B., et al., *The progesterone receptor exon 4 val660leu G/T polymorphism and risk of breast cancer in Australian women.* Cancer Epidemiol Biomarkers Prev, 2002. **11**: p. 439-443.
17. Dunning, A.M., et al., *A systematic review of genetic polymorphisms and breast cancer risk.* Cancer Epidemiol Biomarkers Prev, 1999. **8**: p. 843-854.
18. Pearce, C.L., et al., *Clarifying the PROGINS allele association in ovarian and breast cancer risk: a haplotype-based analysis.* J. Natl. Cancer Inst., 2005. **97**(1): p. 51-59.
19. Leite, D.B., et al., *Progesterone receptor (PROGINS) polymorphism and the risk of ovarian cancer.* Steroids, 2008. **73**: p. 676-680.
20. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**: p. 1087-1093.
21. Meyer, K.B., et al., *Allele-Specific Up-Regulation of FGFR2 Increases Susceptibility to Breast Cancer.* PLoS Biology, 2008. **6**(5): p. e108.
22. Giulianelli, S., et al., *Carcinoma-associated fibroblasts activate progesterone receptors and induce hormone independent mammary tumor growth: A role for the FGF-2/FGFR-2 axis.* Int J Cancer, 2008. **123**: p. 2518-2531.
23. Ben-Haroush, A., et al., *Expression of basic fibroblast growth factor and its receptors in human ovarian follicles from adults and fetuses.* Fertility and Sterility, 2005. **84**(Supplement 2): p. 1257-1268.

24. Antoniou, A.C., et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. **82**: p. 937-948.
25. Latif, A., et al., *Breast cancer susceptibility variants alter risks in familial disease.* J Med Genet, 2010. **47**: p. 126-131.
26. Latif, A., et al. (2010) *Breast cancer susceptibility variants alter risk in familial ovarian cancer*. Familial Cancer, DOI: doi:10.1007/s10689-010-9349-2.
27. Gates, M.A., et al., *Breast cancer susceptibility alleles and ovarian cancer risk in 2 study populations.* Int J Cancer, 2009. **124**: p. 729-733.
28. Cerliani, J.P., et al., *Interaction between FGFR-2, STAT5, and progesterone receptors in breast cancer.* Cancer Res, 2011. **71**(10): p. 3720-3731.
29. Ho, S.-M. (2003) *Estrogen, progesterone and epithelial ovarian cancer*. Reprod Biol Endocrinol **1**, 73 DOI: 10.1186/1477-7827-1-73.
30. Charles, N.J., P. Thomas, and C.A. Lange, *Expression of membrane progesterone receptors (mPR/PAQR) in ovarian cancer cells: implications for progesterone-induced signaling events.* Horm Canc, 2010. **1**: p. 167-176.
31. Aupperlee, M., et al., *Progestins and breast cancer.* Breast Dis, 2005-2006. **24**: p. 37-57.
32. Sun, T., et al., *A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers.* Nat Genet, 2007. **39**: p. 605-613.
33. Cox, A., et al., *A common coding variant in CASP8 is associated with breast cancer risk.* Nat Genet, 2007. **39**: p. 352-358.
34. MacPherson, G., et al., *Association of a Common Variant of the CASP8 Gene With Reduced Risk of Breast Cancer.* J. Natl. Cancer Inst., 2004. **96**(24): p. 1866-1869.

CHAPTER 3.2

FGFR2 is protective for ovarian cancer in *BRCA1/2* mutation carriers

Rita D. Brandão, Marinus J. Blok**,** A.C. Antonis, D. Barrowdale, G. Chenevix-Trench, Fergus J. Couch, and Encarna B. Gómez García, on behalf of the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA)

In preparation

ABSTRACT

Variability observed in penetrance, age of onset, and tumour site, both among and within *BRCA1/2* families, suggests that other, low-penetrance, genetic variants modify the cancer risk. The fact that the target organs are sex-hormone regulated organs suggests that genes involved in the steroid hormone pathways are candidate gene modifiers. A common polymorphism of the fibroblast growth factor receptor 2 (*FGFR2*), associated with increased FGFR2 expression, was previously reported to increase breast cancer risk among *BRCA1/2*-mutation carriers. We hypothesized that *FGFR2* would also modify the risk of ovarian cancer based on the knowledge that FGFR2 activates the progesterone receptor (PR) and is also present in ovarian tissue.

The *FGFR2* rs2981582 was genotyped in a population of 11,704 *BRCA1* and 8,208 *BRCA2* carriers, of whom 1,839 and 631 were ovarian cancer cases, using the custom array iCOGs.

Our results indicate that *FGFR2* is protective for ovarian cancer among *BRCA2* carriers (HR = 0.67, p-value = 0.005). For BRCA1 this association was borderline significant (HR = 0.86, p-value = 0.090).

We hypothesize that the opposite, tissue-specific, effects of *FGFR2* in OC vs. BC are mediated by the progesterone receptor (PR). FGFR2, like progesterone, activates PR, and progesterone exposure is protective for OC (and endometrial cancer), but a risk factor for BC. Therefore, *FGFR2* appears to be a cancer site risk modifier that contributes to the clinical variability observed in *BRCA1/2* carriers. This marker may be relevant for individualized assessment of cancer risks and counselling regarding preventive options for *BRCA1/2*-mutation carriers.

INTRODUCTION

Individualized advice about the most suitable intervention among the risk-reducing options available has been hampered by the considerable variability in the BC and OC risks observed amongst these women. Notably, studies within populations with founder mutations, as well as epidemiological studies, suggest the presence of genetic modifiers co-segregating with BRCA mutations in the families that explain cancer-site risk. For example, a Portuguese founder mutation in the *BRCA2* gene was described to give rise to different phenotypes which were family dependent [1]. Some families had only female breast cancer cases, others had also male breast cancer, and there were families with both breast and ovarian cancer cases, while others had a heterogeneous phenotype [1]. Additional studies performed in Ashkenazi Jewish families, where three founder mutations account for the majority of the mutations identified, have shown that the site-specific cancer risk in a carrier could be better predicted by knowing the tumour sites in her relatives than by knowing the mutation [2]. Epidemiological studies suggested that additional genetic factors influence the cancer risks of individuals from *BRCA1/2* families [3].

FGFR2 was found to be amplified and/or over-expressed [4] or reduced in up to 67% of breast tumour samples [5, 6]. Additionally, missense mutations of the *FGFR2* gene were found to be present in several tumours, including those from breast and ovarian tissues [7]. Recently, through genome-wide association studies, *FGFR2* rs2981582 was found to be strongly associated with an increased risk of breast cancer (BC) [8, 9]. Additionally, *FGFR2* rs2981582 allele increased risk of bilateral BC among the general population and BC among women with a first-degree relative with BC [8], as well as among *BRCA2* mutation carriers [10].

Despite the fact that the DNA-repair action of BRCA1 and BRCA2 proteins is ubiquitous across the human tissues, it is noteworthy that their inactivation leads predominantly to cancer of the breast and ovaries. In those tissues, cell homeostasis is mainly controlled by the steroid hormones, oestrogen and progesterone, and their respective receptors. Furthermore, the activity of oestrogen and progesterone receptors is regulated by the BRCA1 protein [11, 12].

Interestingly, FGFR2 is a PR activator, is also expressed in ovarian tissue (GeneAtlas, http://biogps.org), [13, 14], and the risk allele creates a putative ER binding site due to the SNP rs10736303 [8], which is in strong linkage disequilibrium with rs2981582. Additionally, functional analysis in breast tumours revealed that the risk allele is associated with increased expression of *FGFR2* [15].

Due to the role of FGFR2 in BC and its role in hormonal pathways, we have hypothesized that *FGFR2* may also influence the risk of ovarian cancer among *BRCA1/2*-mutation carriers. In our population, consisting of 2,470 ovarian cancer cases among 19,912 female *BRCA1* and *BRCA2*-mutation carriers, we report that the minor allele of *FGFR2* is protective for ovarian cancer among *BRCA2* carriers with per-allele hazardous-ratio (HR) of 0.672 (p-value = 0.005).

METHODS

Subjects

The patients included participated in clinical or research studies at the host institutions under ethically approved protocols and data was analysed anonymously. Subjects were *BRCA1* and *BRCA2* mutation carriers recruited by 42 study centres in 22 countries through the CIMBA initiative. The majority of patients were recruited through cancer genetics clinics offering genetic testing, and enrolled into national or regional studies. Women were included in the analysis if they carried mutations that were pathogenic according to generally recognized criteria [16]. Details of the CIMBA initiative were previously reported [17]. Briefly, most carriers were identified by population-based sampling of cases, others through community recruitment (e.g. in Ashkenazi Jewish populations). Only female *BRCA1/2*-mutation carriers above the age of 18 are eligible to participate in the CIMBA study. Information collected includes: the year of birth, mutation description, age at the last follow-up, age of cancer diagnoses, and age or date at bilateral prophylactic mastectomy and adnexectomy.

Genotyping

Genotyping was performed using the iCOGs, an Illumina iSelect custom array designed by the Collaborative Oncological Gene-environment Study (COGS). As an additional genotyping quality-control check, the deviation from Hardy–Weinberg equilibrium (HWE) was evaluated. A total of 19,912 mutation carriers (11,704 *BRCA1* and 8,208 *BRCA2* carriers) from 42 studies had an observed genotype and were therefore included in this study. Of these patients, 1,839 *BRCA1*- and 631 *BRCA2*-mutation carriers developed OC.

Statistical analysis

The aim of the analysis was to evaluate the association between each genotype and the risk of ovarian cancer. The phenotype of each individual was defined by the cancer site and the age at diagnosis of cancer or the age at the last follow-up. Details of the analysis have been described previously [18]. The effect of the *FGFR2* SNP is given as a per-allele HR (multiplicative model), estimated on the logarithmic scale and the HRs were assumed to be independent of age (i.e. we used a Cox proportional-hazards model).

RESULTS AND DISCUSSION

Here we have assessed the risk of ovarian cancer in 11,704 and 8,208 *BRCA1*- and *BRCA2*-mutation carriers, respectively. Our data provides evidence that the minor T allele of *FGFR2* rs2981582 is significantly protective for ovarian cancer among *BRCA2* carriers (p-value = 0.005) (Table 1). Among *BRCA1* mutation carriers there is a borderline association

Table 1. Per-allele associations with ovarian cancer per mutation status

Population studied	Per-allele HR	p-value
BRCA1	0.86	0.090
BRCA2	0.67	0.005

(p-value = 0.09). Previous studies did not find an association between *FGFR2* T allele and the risk of sporadic ovarian cancer (n=1,095 and 2,513) [19, 20] or among *BRCA1/2* carriers (n=52) [21]. In another study, *FGFR2* was found to be protective for endometrial cancer (n=652) [22]. The differences observed may be related to differences between BRCA1/2 and general populations regarding OC etiology. The only other study performed among BRCA1/2-mutation carriers had a small sample size, and was probably underpowered.

FGFR2 has mainly two isoforms, FGFR2-IIIb and FGFR2-IIIc, which contain mutually exclusive exons 9 and 10, respectively. Differential exclusion of these exons only leads to different FGF ligands specificity. Since the risk allele lies in intron 2 of the gene, it is not expected that it influences the splicing pattern of the gene. But it increases the overall expression of the gene [15], which is expected to occur independently of the tissue-specific isoform produced.

The pathways that involve FGFR2 protein include FGF signalling, AKT, MAPK, NF-kB, PTEN, and stem cell signalling. Therefore, FGFR2 plays a role in important biological processes such as embryonic development, apoptosis, cell differentiation, proliferation and tissue repair, especially of bone and blood vessels. Deregulation of the FGFR2 signalling will affect cell homeostasis and may lead to carcinogenesis. Interestingly, FGFR2 is also involved in hormonal cellular response and we believe that this link provides an explanation for the opposite effects of *FGFR2* in breast and ovarian tissues among *BRCA1/2*-mutation carriers. In human breast cancer cells, FGFR2 activates progesterone receptor (PR) leading to stimulation of cell growth [13]. More specifically, FGFR2 is a co-activator of PR together with STAT5, binding to it in the nucleus, when PR is bound to DNA progesterone responsive elements [14]. Since the risk allele is associated with increased expression of *FGFR2* [15], it is expected that it leads to increased PR activation amplifying its effect on the target cells. In addition, progesterone exposure, and subsequent PR activation, is protective for OC and endometrial cancer, while it is a risk factor for BC [23-26]. Therefore, the link of FGFR2 and PR can explain the risk of BC previously described and the protective effect on OC that we observe. Epidemiological studies have shown that women with early menarche, late menopause, and women exposed to oral contraceptives and progesterone analogs in hormone replacement therapy have increased breast cancer risk, whereas those who removed their ovaries have a reduced breast cancer risk [26]. These observations are consistent with PR activation being a risk

factor for breast cancer. In contrast, ovarian cancer risk is reduced in women who were exposed to oral contraceptives and increased progesterone levels due to multiparity or twin pregnancy and increased risk of ovarian cancer is observed in women with progesterone deficiency and after menopause, which is associated with decreased levels of progesterone [23, 25, 27].

The link between FGFR2 and progesterone pathways might also explain the differences of the risk of ovarian cancer between BRCA1 and BRCA2 carriers. *FGFR2* was also more strongly associated with breast cancer risk among *BRCA2*-mutation carriers than *BRCA1* carriers [10, 28]. And within each group, this association is stronger for oestrogen receptor (ER) and PR-positive breast tumours [29]. In a large group of *BRCA1/2* carriers, it was confirmed that *BRCA1*-related breast tumours are more frequently ER- and PR-negative than those from *BRCA2* [30]. Therefore, it is not surprising that the association with breast cancer is stronger for *BRCA2* carriers. Regarding *BRCA1/2*-related ovarian cancer pathology, to the best of our knowledge, the status of the hormone receptors is unknown. Most ovarian tumours arising in *BRCA1/2* carriers are invasive epithelial cancers of serous histology [30], and in the general population these tumours have significantly lower PR- and ER-positivity compared with normal ovarian tissue [31], but they are not completely hormone receptor negative. Perhaps *BRCA1*-associated ovarian tumours are also more frequently ER- and PR-negative compared with those from *BRCA2*-carriers, as observed in breast tumours. This would explain the stronger association observed for the protective effect of *FGFR2* among *BRCA2*-mutation carriers.

Table 2. Additional OC genetic modifiers and their associations per mutation status

	BRCA1		BRCA2	
Polymorphism	Per-allele HR	p-value	Per-allele HR	p-value
rs3814113 (*BNC2*)	0.75	4.8×10^{-9}	0.78	5.5×10^{-4}
rs67397200 (19p13.1)	1.16	3.8×10^{-4}	1.30	1.8×10^{-3}
rs10088218 (8q24)	0.89	0.029	0.81	0.033
rs2665390 (3q25)	1.25	6.1×10^{-4}	1.48	1.8×10^{-4}
rs717852 (2q31)	1.06	0.16	1.25	6.6×10^{-4}
rs9303542 (17q21)	1.08	0.06	1.16	0.026

There are no reports yet on ovarian cancer risk modifiers among *BRCA1* and *BRCA2* carriers from genome-wide associations studies; however, several studies analyzed the loci identified in sporadic ovarian cancer cases. Results from these studies confirmed that these variants also modify ovarian cancer risk among *BRCA1/2* carriers (Table 2) [32-34]. It is noteworthy that the locus 19p13 had also been previously found to be associated with breast cancer risk among *BRCA1* female carriers.

The results from our study indicate that *FGFR2* contributes to the cancer-site variability observed in members of *BRCA1/2* families. The mechanism of action and involved proteins warrant further studies. Additionally, other genes are likely to contribute to the different phenotypes observed among the *BRCA1/2*-mutation carriers. The cumulative effect of the several risk alleles remains to be studied. Clarifying the genetic profile that contributes to specific cancer-site risks will allow a more accurate estimation of the risk for each individual and, consequently, lead to personalized clinical management.

REFERENCES

1. Machado, P.M., R.D. Brandão, B.M. Cavaco, et al., *Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes.* J Clin Oncol, 2007. **25**(15): p. 2027-2034.
2. Simchoni, S., E. Friedman, B. Kaufman, et al., *Familial clustering of site-specific cancer risks associated with BRCA1 and BRCA2 mutations in the Ashkenazi Jewish population.* Proc Natl Acad Sci USA, 2006. **103**: p. 3770-3774.
3. Antoniou, A.C. and D.F. Easton, *Models of genetic susceptibility to breast cancer.* Oncogene, 2006. **25**: p. 5898-5905.
4. Adnane, J., P. Gaudray, C. Dionne, et al., *BEK and FLG, two receptors to members of the FGF family, are amplified in subsets of human breast cancers.* Oncogene, 1991. **6**: p. 659-63.
5. Luqmani, Y.A., M. Graham, and R.C. Coombes, *Expression of basic fibroblast growth factor, FGFR1 and FGFR2 in normal and malignant human breast, and comparison with other normal tissues.* Br J Cancer, 1992. **66**(2): p. 273–280.
6. Zhu, X., S.L. Asa, and S. Ezzat, *Genetic and epigenetic mechanisms down-regulate FGF receptor 2 to induce melanoma-associated antigen A in breast cancer.* Am J Pathology, 2010. **176**(5): p. 2333-2343.
7. Katoh, M., *Cancer genomics and genetics of FGFR2 (Review).* Int J Oncol, 2008. **33**: p. 233-237.
8. Easton, D.F., K.A. Pooley, A.M. Dunning, et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**: p. 1087-1093.
9. Hunter, D.J., P. Kraft, K.B. Jacobs, et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nat Genet, 2007. **39**(7): p. 870-874.
10. Antoniou, A.C., A.B. Spurdle, O.M. Sinilnikova, et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. **82**: p. 937-948.
11. Zheng, L., L.A. Annab, C.A. Afshari, W.-H. Lee, and T.G. Boyer, *BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor.* Proc Natl Acad Sci U S A, 2001. **98**(17): p. 9587-9592.
12. Calvo, V. and M. Beato, *BRCA1 counteracts progesterone action by ubiquitination leading to progesterone receptor degradation and epigenetic silencing of target promoters.* Cancer Research, 2011. **71**(9): p. 3422-3431.
13. Giulianelli, S., J.P. Cerliani, C.A. Lamb, et al., *Carcinoma-associated fibroblasts activate progesterone receptors and induce hormone independent mammary tumor growth: A role for the FGF-2/FGFR-2 axis.* Int J Cancer, 2008. **123**: p. 2518-2531.
14. Cerliani, J.P., T. Guillardoy, S. Giulianelli, et al., *Interaction between FGFR-2, STAT5, and progesterone receptors in breast cancer.* Cancer Res, 2011. **71**(10): p. 3720-3731.
15. Meyer, K.B., A.-T. Maia, M. O'Reilly, et al., *Allele-specific up-regulation of FGFR2 increases susceptibility to breast cancer.* PLoS Biol, 2008. **6**(5): p. e108.
16. Mitchell, G., A.C. Antoniou, R. Warren, et al., *Mammographic density and breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Cancer Res, 2006. **66**: p. 1866 - 1872.
17. Chenevix-Trench, G., R. Milne, A. Antoniou, et al., *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, 2007. **9**(2): p. 104.
18. Antoniou, A.C., O.M. Sinilnikova, J. Simard, et al., *RAD51 135G-->C Modifies Breast Cancer Risk among BRCA2 Mutation Carriers: Results from a Combined Analysis of 19 Studies.* The American Journal of Human Genetics, 2007. **81**(6): p. 1186-1200.
19. Gates, M.A., S.S. Tworoger, K.L. Terry, et al., *Breast cancer susceptibility alleles and ovarian cancer risk in 2 study populations.* Int J Cancer, 2009. **124**: p. 729-733.
20. Song, H., S.J. Ramus, S.K.g. Kjaer, et al., *Association between invasive ovarian cancer susceptibility and 11 best candidate SNPs from breast cancer genome-wide association study.* Hum Mol Genet, 2009. **18**(12): p. 2297-2304.

21. Latif, A., H. McBurney, S. Roberts, et al., *Breast cancer susceptibility variants alter risk in familial ovarian cancer.* Familial Cancer, 2010, doi: doi:10.1007/s10689-010-9349-2
22. McGrath, M., I.-M. Lee, J. Buring, D.J. Hunter, and I.D. Vivo, *Novel breast cancer risk alleles and endometrial cancer risk.* Int J Cancer, 2008. **123**: p. 2961-2964.
23. Ho, S.-M., *Estrogen, progesterone and epithelial ovarian cancer.* Reprod Biol Endocrinol, 2003. **1**: p. 73, doi: 10.1186/1477-7827-1-73
24. Romano, A., B. Delvoux, D.-C. Fischer, and P. Groothuis, *The PROGINS polymorphism of the human progesterone receptor diminishes the response to progesterone.* J Mol Endocrinol, 2007. **38**: p. 331-350.
25. Charles, N.J., P. Thomas, and C.A. Lange, *Expression of membrane progesterone receptors (mPR/PAQR) in ovarian cancer cells: implications for progesterone-induced signaling events.* Horm Canc, 2010. **1**: p. 167-176.
26. Aupperlee, M., A. Kariagina, J. Osuch, and S. Haslam, *Progestins and breast cancer.* Breast Dis, 2005-2006. **24**: p. 37-57.
27. Collaborative Group on Epidemiological Studies of Ovarian Cancer, V. Beral, R. Doll, et al., *Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23?257 women with ovarian cancer and 87?303 controls.* Lancet, 2008. **371**(9609): p. 303-314.
28. Antoniou, A.C., J. Beesley, L. McGuffog, et al., *Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction.* Cancer Res, 2010. **70**(23): p. 9742-9754.
29. Mulligan, A., F. Couch, D. Barrowdale, et al., *Common breast cancer susceptibility alleles are associated with tumour subtypes in BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2.* Breast Cancer Res, 2011. **13**(6): p. R110.
30. Mavaddat, N., D. Barrowdale, I.L. Andrulis, et al., *Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA).* Cancer Epidemiol Biomarkers Prev, 2012. **21**(1): p. 134-147.
31. Lindgren, P.R., S. Cajander, T. Bäckström, et al., *Estrogen and progesterone receptors in ovarian epithelial tumors.* Mol Cell Endocrinol, 2004. **221**(1â€“2): p. 97-104.
32. Ramus, S.J., C. Kartsonaki, S.A. Gayther, et al., *Genetic Variation at 9p22.2 and Ovarian Cancer Risk for BRCA1 and BRCA2 Mutation Carriers.* J Natl Cancer Inst, 2011. **103**(2): p. 105-116.
33. Couch, F.J., M.M. Gaudet, A.C. Antoniou, et al., *Common Variants at the 19p13.1 and ZNF365 Loci Are Associated with ER Subtypes of Breast Cancer and Ovarian Cancer Risk in BRCA1 and BRCA2 Mutation Carriers.* Cancer Epidemiol Biomarkers Prev, 2012.
34. Ramus, S.J., A.C. Antoniou, K.B. Kuchenbaecker, et al., *Ovarian cancer susceptibility alleles and risk of ovarian cancer in BRCA1 and BRCA2 mutation carriers.* Hum Mutat, 2012: p. n/a-n/a, doi: 10.1002/humu.22025
35. Song, H., S.J. Ramus, J. Tyrer, et al., *A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2.* Nat Genet, 2009. **41**(9): p. 996-1000.
36. Goode, E.L., G. Chenevix-Trench, H. Song, et al., *A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24.* Nat Genet, 2010. **42**(10): p. 874-879.
37. Bolton, K.L., J. Tyrer, H. Song, et al., *Common variants at 19p13 are associated with susceptibility to ovarian cancer.* Nat Genet, 2010. **42**(10): p. 880-884.

CHAPTER 4.1

Characterisation of unclassified variants in the *BRCA1/2* genes with a putative effect on splicing

Rita Dias Brandão, Kees van Roozendaal, Demis Tserpelis, Encarna Gómez García, Marinus J. Blok

Breast Cancer Research and Treatment 2011, 129(3):971-82

ABSTRACT

A subset of the unclassified variants (UVs) identified during genetic screening of *BRCA1/2* genes may affect splicing. We assessed at RNA level the effect of 4 *BRCA1* and 10 *BRCA2* UVs with a putative splice effect, as predicted *in silico*.

The variants selected for this study were beyond the positions -1,-2 or +1, +2 from the exon and were not previously described (n=8) or their effect on splicing was not assessed previously (n=6). Lymphocytes from UV carriers and healthy controls were cultured and treated with puromycin to prevent nonsense-mRNA mediated decay. The relative contribution of each allele to the various transcripts was assessed using combinations of allele-specific and transcript-specific primers.

BRCA2 c.425G>T, c.7976+3_7976+4del, and c.8754+3G>C, give rise to aberrant transcripts *BRCA2*Δ4, *BRCA2*Δ17 and retention of 46nt of intron 21, respectively, and were considered pathogenic. *BRCA1* c.4987-3C>G gives rise to *BRCA1*Δ17 which is likely pathogenic, however, residual expression of the full-length transcript from the variant allele could not be excluded. *BRCA1* c.692C>T, c.693G>A and *BRCA2* c.6935A>T, besides expressing the full-length transcript, increased expression of *BRCA1*Δ11 and *BRCA2*Δ12, respectively. As these are natural occurring isoforms, also observed in controls, the clinical relevance is unclear. The seven remaining UVs did not affect splicing and three intronic variants were therefore classified as neutral.

In conclusion, the RNA analysis results clarified the clinical relevance of six of the fourteen studied UVs and thereby greatly improves the genetic counselling of high risk breast/ovarian cancer patients carrying these classified variants.

INTRODUCTION

Approximately half of the variants reported in BIC database [1] are unclassified variants (UVs) (46% *BRCA1* and 55% *BRCA2*). These consist mostly of rare missense variants, intronic variants and silent variants. A subset of UVs may affect splicing by disturbing the recognition of the donor and acceptor splice sites (DSS and ASS, respectively) or by disrupting intronic and exonic cis-elements necessary for the regulation of splicing, such as exonic splice enhancer motifs (ESEs) [2].

The splice effect of a variant can be predicted *in silico* using several freely available online algorithms or using commercially available software such as Alamut (Interactive Biosoftware), which integrates six online algorithms. For determination of DSS/ASS, Alamut contains the SpliceSiteFinder-like, which is based on the former algorithm of Alex Dong Li's Splice Site Finder [3], MaxEntScan [4], NNSplice [5] and GeneSplicer [6]. The detection of ESEs is performed with ESEFinder and RESCUE-ESE [7, 8]. However, RNA analysis remains necessary to confirm the *in silico* predictions, in particular when it concerns positions beyond the highly conserved splice donor (GT) and acceptor motifs (AG).

In this study, we assessed the effect on mRNA splicing of a panel of 14 UVs, 4 *BRCA1* and 10 *BRCA2* variants, which were predicted *in silico* to affect a splice site or putative ESE motifs. Their effect on splicing was not assessed before, even for those variants previously described (n=6). We describe four variants that give rise to new aberrant transcripts and are therefore considered to be pathogenic (n=3) or likely pathogenic (n=1). Additionally, three variants increase the expression of naturally occurring isoforms without affecting the expression of the full-length transcript and, because their critical expression levels and functions are unknown, their clinical relevance of remains elusive.

METHODS

Patients and splicing prediction software

The selection criteria of high-risk breast and/or ovarian cancer families and the *BRCA1/2* mutation screening methods were previously described [9]. The putative splice variants included in this study were detected during genetic screening of 1800 unrelated probands. No other mutation was identified in the families carrying these variants and their effect on splicing was predicted using Alamut (version 1.5, Interactive Biosoftware). We selected variants that were predicted to affect DSS or ASS, create new splice sites, or could have an effect on ESEs. We excluded mutations at positions -1,-2 or +1, +2 from the exon and variants previously described to have an effect on splicing. The personal cancer history and the cancer history of the family of the individuals studied are shown in Supplementary Table 1. The variants studied and their respective predictions are described in Table 1. The

formula used to calculate the change of the scores of the splice sites was as follows: (WT splice site score in the presence of the variant – WT splice site score)/ maximum of the scale x 100. This allows to compare the changes between different algorithms, as they have different scales. Table 2 summarizes the number of times that the UVs were reported in the online databases BIC [1] and LOVD [10].

Nomenclature

The description of the variants follows the Human Genetic Variation Society (HGVS) approved guidelines, where c.1 is the A of ATG translation initiation codon [11]. The accession numbers used for *BRCA1* and *BRCA2* mRNA and protein were U14680.1/NP_009225.1 and U43746.1/NP_000050.2, respectively.

PBL cultures

White blood cells were isolated from fresh whole blood (collected in EDTA tubes) and used either fresh or frozen in FCS with 10% of DMSO for subsequent culture in a complete medium consisting of: RPMI 1640 supplemented with L-glutamine (Gibco) and 12.5% FCS, 1x L-glutamine, 0.8mM sodiumpyruvate (Gibco), 17mM Hepes buffer (Gibco), $4.2x10^{-2}$ mM 2-mercaptoethanol (Gibco), 42 units/mL penicillin-streptomycin and 0.21 g/mL amphotericin B solution (Sigma). Lymphocyte growth was stimulated with 50 μL/mL PHA (Gibco) and 10 units/mL of IL-2 (Roche). At day 7, 4-6h before harvesting the cells, each culture was split evenly and one part was treated with 200μg/mL of puromycin (Sigma). Puromycin can enrich for transcripts that contain a premature stop codon (PTC), because it inhibits degradation of these transcripts by nonsense-mediated mRNA decay (NMD) [13].

RNA isolation and cDNA synthesis

Total RNA was isolated using TRIzol (Invitrogen) or TRIpure (Roche) reagent and first-strand cDNA was obtained with reverse transcriptase II (Invitrogen), using a combination of random hexamers and oligo-dT primers (Invitrogen) or random hexamers alone, according to the manufacturers' instructions.

(AS)PCR

PCRs were performed using primers that flank the regions of interest to identify insertions or deletions by fragment size analysis on 2% agarose gels stained with ethidium bromide and sequenced in an ABI PRISM 3730 (Applied Biosystems) instrument using Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Allele-specific (AS)PCR was performed to determine the relative contribution of each allele to the synthesis of full-length or alternative transcripts. For ASPCR, the 3' terminus of the primers was designed to correspond to the position of a heterozygous sequence variant, matching to either one of the alleles. To detect only full-length or alternative transcripts, primers were located within unique regions of these transcripts, e.g. within skipped exons or across exon-exon

boundaries. Additional intentional mismatch in the first or second penultimate 3' nucleotide was included to increase the specificity of the allelic analysis [14]. Supplementary Table 2 describes the primers used in this study.

Table 1. Splice site prediction changes obtained with the Alamut software[a].

	Splice site prediction algorithm					
	SpliceSiteFinder-like [0-100]	MaxEntScan ASS : [0-16] DSS : [0-12]	NNSplice [0-1]	GeneSplicer [0-15]	ESEfinder	Rescue-ESE
BRCA1						
c.692C>T	NA	NA	NA	ASS: 7.5 to 7.0 (-3%)	3 ESEs disrupted	1 ESE created
c.693G>A	NA	Cryptic ASS site: 0 to 2.7 (+16.9%)	NA	ASS: 7.5 to 7.0 (-3%)	3 ESEs disrupted	1 ESE created
c.4092C>G	NP; cryptic site: 0 to 72.7 (+72.7%)	NA	DSS: 0.5 to 0 (-50%)	DSS: 0.5 to 0 (-3%)	3 ESEs disrupted	NA
c.4987-3C>G	ASS: 84.5 to 73.7 (-11.8%)	ASS: 6.7 to 0.6 (-40.7%)	ASS: 0.6 to 0 (-60%)	ASS: 1.4 to 0 (-9.3%)	NA	1 ESE disrupted, 1 ESE created
BRCA2						
c.425G>T	DSS: 84.2 to 71.6 (-12.6%)	DSS: 9.1 to 2.6 (-54%)	DSS: 0.9 to 0 (-90%)	DSS: 0.9 to 0 (-6%)	1 ESE disrupted	1 ESE disrupted
c.794-11T>C	ASS: 88.8 to 84.7 (-4.1%)	ASS: 9.6 to 9.3 (-2%)	ASS: 0.9 to 0.8 (-10%)	ASS: 4.9 to 3.2 (-11%)	1 ESE disrupted, 1 ESE created	NA
c.6935A>T	DSS: 73.7 to 0 (-73.7%)	DSS: 4.8 to 3.2 (-13%)	DSS: 0.6 to 0 (-60%)	NP	4 ESEs disrupted, 1ESE created	NA
c.6938-3T>C	ASS: 83.3 to 89.5 (+6.2%)	ASS: 5.5 to 5.2 (-2%)	NP	NP	2 ESEs changed, 2 ESEs created	NA
c.6943A>G	NA	NA	ASS: 0 to 0.4 (+40%)	NP	1 ESE disrupted, other 2 increase scores	NA
c.7976+3_7976+4del	DSS: 100.0 to 0 (-100%)	NP	NP	NP	1 ESE created	NA
c.8350C>T	NA	NA	NA	NP	3 ESEs disrupted	NA
c.8662C>T	NA	NA	NA	NA	2 ESEs disrupted	NA

	SpliceSiteFinder-like [0-100]	MaxEntScan ASS : [0-16] DSS : [0-12]	NNSplice [0-1]	GeneSplicer [0-15]	ESEfinder	Rescue-ESE
c.8754+3G>C	DSS: 87.3 to 81.7 (-5.6%)	DSS: 7.7 to 5.2 (-20.8%)	DSS: 1.0 to 0.6 (-40%); cryptic DSS: 0 to 0.6 (+60%)	DSS: 3.2 to 0 (-21.3%); Cryptic DSS: 6.7 to 6.6. (-0.7%)	3 ESEs created	1 ESE disrupted
c.8953+13A>G	NA; cryptic DSS: 71.1 to 83.3 (+12.2%)	NA; cryptic DSS: 0 to 6.7 (+55.8%)	NA; cryptic DSS: 0 to 0.9 (+90%)	DSS changes from 4.2 to 4.3 (+0.7%)	NA	NA

In square brackets is shown the range of the scores used by each algorithm.
Abbreviations used: DSS- donor splice site; ASS- Acceptor splice site; NA- not affected; NP- not predicted
The percentage of the decrease/increase of the WT splice site was calculated using the following formula:
(WT splice site score in the presence of the variant – WT splice site score) / maximum of the scale x 100
a. The nomenclature used is according to the HGVS approved guidelines [12], where +1 is the nucleotide A of the ATG translation initiation codon.

RESULTS

Variants affecting splicing

***BRCA1* c.692C>T, c.693G>A:** The variants *BRCA1* c.692C>T and c.693G>A lie within exon 11 and were identified in two different patients. These variants localise 22 and 23 nt downstream of the ASS, respectively. In addition to full-length transcript, there are two other known isoforms in this region: *BRCA1*Δ11 and *BRCA1*Δ11q [15], being both in-frame deletions. *BRCA1*Δ11 lacks the complete exon 11 and *BRCA1*Δ11q only has the first 120 bp of exon 11 (GenBank accession number in NM_007298).

Firstly, we amplified cDNA of patients and controls with primers hybridising to exon 10 and exon 12 (Fig. 1A). The full-length transcript was not observed under the conditions used. In both patients and controls, amplification resulted in a band corresponding in size to that expected for the *BRCA1*Δ11q transcript. In addition, both the carriers of c.692C>T and c.693G>A variants showed an additional strong band that was determined, by sequencing, to be the *BRCA1*Δ11 transcript. The controls show a faint band of the same size.

Secondly, we aimed at determining the contribution of normal and variant alleles to the expression of the *BRCA1*Δ11 transcript, using a forward primer across the exons 10/12 boundary in combination with a reverse primer in exon 14 (Fig. 1B). The polymorphism c.4308C>T in exon 13 (allele frequency = 0.31, BIC database), for which both patients were heterozygous, was used to characterise the contribution of each allele to the expression of the transcript. In both patients' samples, sequencing results show that the only nucleotide observed at position c.4308 is a thymine (Fig. 1C). This indicates that the c.692C>T and c.693G>A carriers have monoallelic expression of the *BRCA1*Δ11 transcript. Because in this family, the variant c.693G>A segregates with the polymorphism c.4308T

(results not shown), it is clear that the allele giving rise to the alternative *BRCA1*Δ11 transcript contains the variant c.693A. For the variant c.692C>T, segregation analysis to confirm if the polymorphism c.4308T and c.692T are on the same allele could not be performed.

Thirdly, we determined the contribution of normal and variant alleles to the expression of the full-length transcript using a forward primer hybridising to the 3' of exon 11, in combination with a reverse primer in exon 14. Sequencing results of the full-length transcript showed equimolar biallelic expression of the full-length transcript in the patients (Fig. 1C). Thus, none of the variants affects the expression level of the full-length transcript. As a consequence, the variant c.692C>T gives rise to a full-length transcript with an amino acid change (p.T231M). The variant c.693G>A is a silent variant.

Table 2. RT-PCR results of the variants with putative effect on splicing analysed in this study.

Patient no.	Sequence variant[a]	exon	No. times reported in BIC [1] /LOVD [10]	mRNA effect	Protein effect[a]
BRCA1					
P1	c.692C>T	11	2/0	Partial ex 11 skipping	p.Ala224_Leu1365del + p.T231M
P2	c.693G>A	11	0/0	Partial ex 11 skipping	p.Ala224_Leu1365del + p.=
P3	c.4092C>G	11	0/0	No effect	p.Asn1364Lys
P4	c.4987-3C>G	17	0	Ex 17 skipping	p.Val1665SerdelfsX9
BRCA2					
P5	c.425G>T	4	0/0	Ex 4 skipping	p.Gly106ValdelfsX9
P6	c.794-11T>C	10	3	No effect	p.=
P7	c.6935A>T	12	5/0	Partial ex 12 skipping	p.Gly2281_Asp2312 del + p.Asp2312Val
P8	c.6938-3T>C	13	0	No effect	p.=
P9	c.6943A>G	13	2/0	No effect	p.Ile2315Val
P10	c.7976+3_7976+4del	17	0	Ex 17 skipping	p.Arg2602_Tyr2658del
P11	c.8350C>T	19	5/1	No effect	p.Arg2784Trp
P12	c.8662C>T	21	4/1	No effect	p.Arg2888Cys
P13	c.8754+3G>C	21	0	Retention 46nt of intron 21	p. Tyr2920ArgfsX3
P14	c.8953+13A>G	22	0	No effect	p.=

a. The nomenclature used is according to the HGVS approved guidelines [11], where +1 is the nucleotide A of the ATG translation initiation codon.

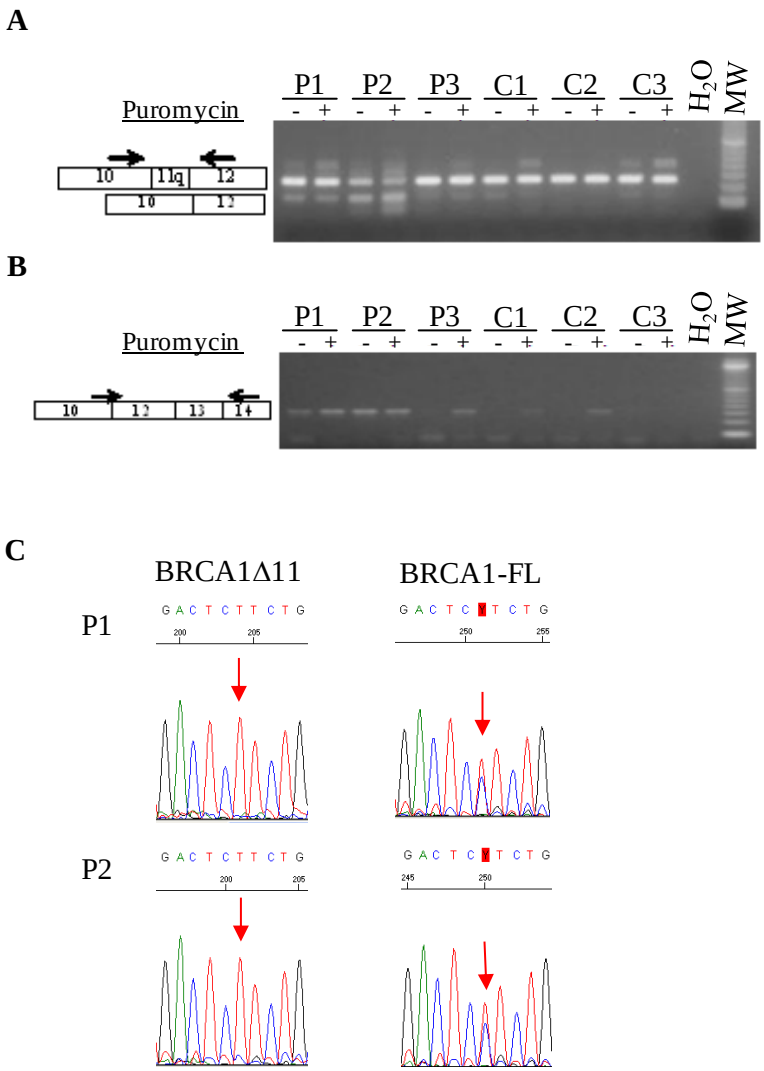


Figure 1. RT-PCR analysis of the variants *BRCA1* c.692C>T (P1), c.693G>A (P2) and c.4092C>G (P3). A) Analysis by electrophoresis of the PCR products obtained using primers flanking exon 11. B) PCR with forward primer specific for the *BRCA1*Δ11 transcript. C) Sequencing results of the contribution of each allele to the expression of each transcript in the puromycin non-treated fractions. On the left side, the results for the *BRCA1*Δ11 and on the right side, the results for the full-length transcript (using a forward primer at the 3' side of exon 11). The arrow indicates the polymorphism c.4308C>T in exon 13. The puromycin-treated samples were sequenced and the results were similar (not shown). C: control sample; H_2O: negative PCR control; MW: molecular weight ladder XIV (Roche). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

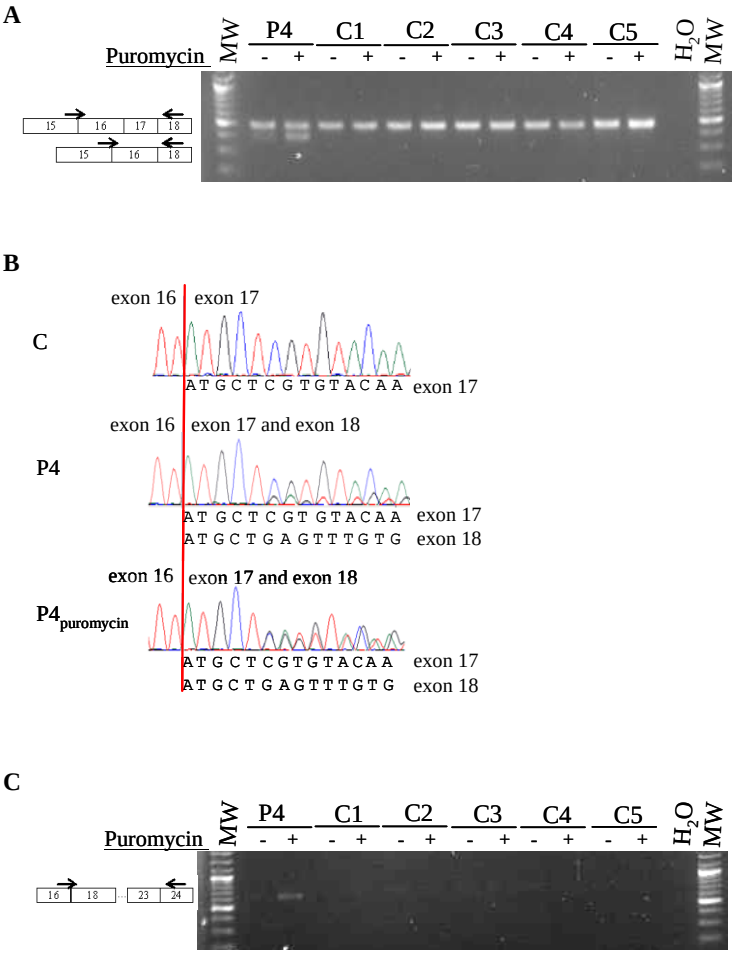


Figure 2. RT-PCR analysis of the variant *BRCA1* c.4987-3C>G (P4). A) Electrophoresis of the PCR products obtained using exonic primers flanking exon 17. B) Sequencing of the PCR products shown in panel A. C) Amplification products obtained with forward primer specific for the *BRCA1*Δ17 transcript and reverse primer in exon 24. C: control sample; H_2O: negative PCR control; MW: molecular weight ladder: O'GeneRuler DNAladder mix (Fermentas). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

***BRCA1* c.4987-3C>G:** We amplified a region flanking exon 17 of *BRCA1* (Fig. 2A) and sequencing results revealed exon 17 skipping in the c.4987-3C>G carrier (Fig. 2B), which was not observed in other controls (Fig. 2A). This event was also confirmed by amplification with a transcript-specific primer that hybridises across exons 16 and 18 (Fig. 2C). As the patient does not carry a heterozygous polymorphism, it was not possible to perform ASPCR to exclude that the allele with the variant still gives rise to full-length

transcript. Similar results were obtained for a daughter of the initially tested index (data not shown), however, the daughter also did not carry a heterozygous polymorphism.

***BRCA2* c.425G>T:** Amplification of the region flanking exon 4 revealed an additional smaller band in the patient's samples (Fig. 3A). Sequencing confirmed that the variant c.425G>T gives rise to exon 4 skipping (Fig. 3B). This frameshift event leads to a PTC (Table 2). The *BRCA2*Δ4 transcript was not present in controls even when a primer specific for this transcript was used for amplification (Fig. 3C). Amplification and sequencing of the full-length transcript, using a primer hybridising to exon 4, confirmed that the variant allele does not give rise to the full-length transcript (Fig. 3D).

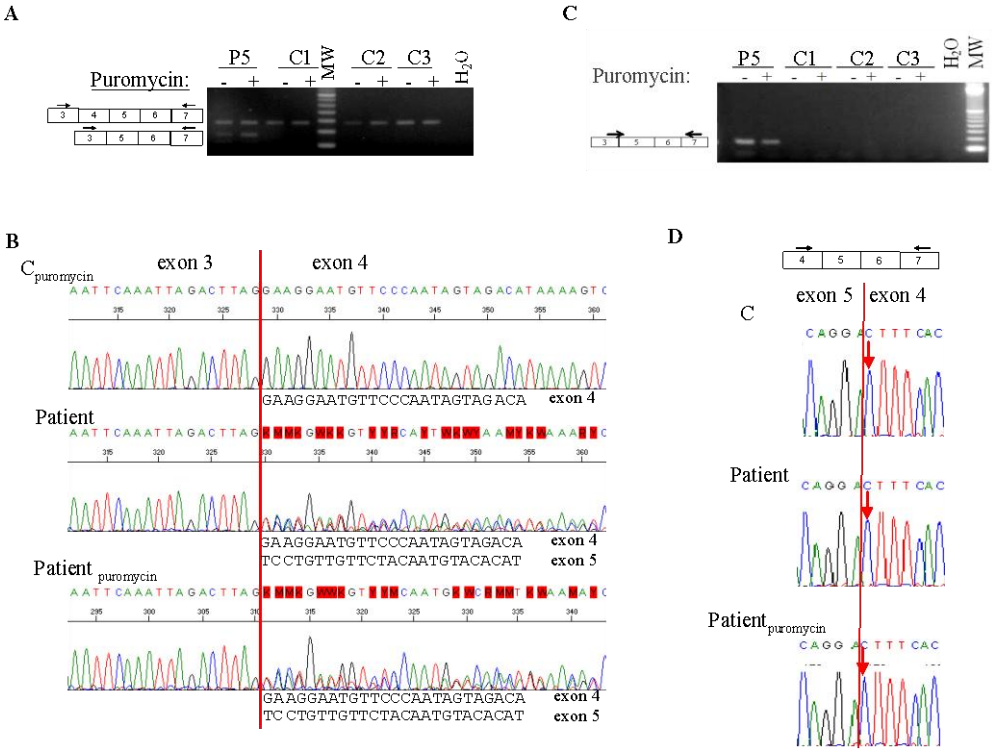


Figure 3. RT-PCR analysis of the variant *BRCA2* c.425G>T (P5). A) Electrophoresis of the products obtained after amplification using exonic primers flanking exon 4. B) Sequencing of some of the PCR products shown in panel A. C) Products of the PCR with primer forward specific for the *BRCA2*Δ4 transcript. D) Reverse sequences of the full-length specific PCR using a forward primer in exon 4. The red arrow indicates the c.425 position. C: control sample; H_2O: negative PCR control; MW: molecular weight ladder XIV (Roche). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

***BRCA2* c.6935A>T:** This variant lies three nucleotides upstream the DSS of intron 12. Amplifying the region flanking exon 12 (Fig. 4A), we observed a known isoform, *BRCA2*Δ12 [16], both in the patients and controls.

To analyse the allelic expression of each of the isoform transcripts, we used either a forward primer specific for the full-length transcript (hybridising to exon 12) or a *BRCA2*Δ12 transcript-specific forward primer (hybridising across the exon 11/13 boundary) in combination with a reverse primer in exon 14 (Fig. 4B). We determined the contribution of each allele to either one of the transcripts using the ratio of the polymorphism c.7242A>G (allele frequency = 0.21, BIC database), for which the patient and a control are heterozygous.

We observed a clear allelic imbalance of *BRCA2*Δ12 expression in the patient. It was not possible to determine in this family which of the alleles of the c.7242A>G polymorphism was co-segregating together with the variant c.6935T and, therefore, it was not possible to determine unambiguously that increased expression of *BRCA2*Δ12 transcript

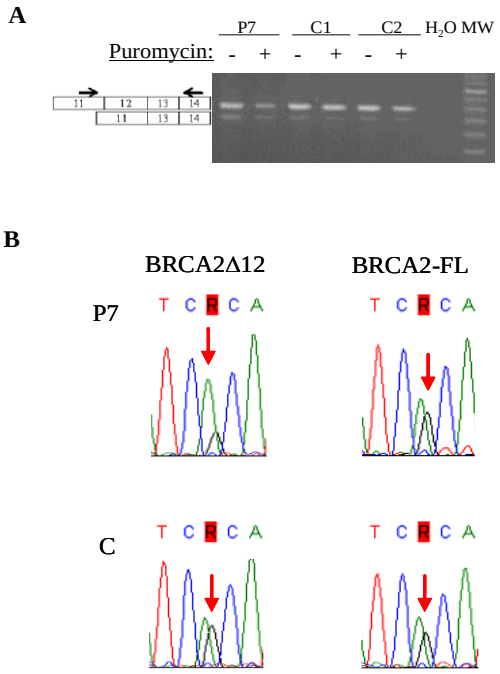


Figure 4. RT-PCR analysis of the variant *BRCA2* c.6935A>T (P7). A) Electrophoresis of amplification products, from cDNA of patient and control samples with primers flanking exon 12. B) Sequencing results of transcript-specific PCRs using either a forward primer hybridising across exon 11 and exon 13 boundary or a forward primer in exon 12 in combination with a reverse primer in exon 14. The red arrow indicates the polymorphism c.7242A>G. The results of the puromycin-treated controls were similar (not shown). C: control sample; H_2O: negative PCR control; MW: molecular weight ladder: O'GeneRuler DNAladder mix (Fermentas). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

comes from the c.6935T allele. However, because it is predicted to have this effect and the increased exon 12 skipping is only observed in the patient sample, we strongly believe that there is a direct causal relationship. Consequently, besides higher expression levels of the in-frame exon 12 deletion transcript, the c.6935T allele expresses full-length transcript at similar expression levels compared with the wild type (WT) allele and control alleles. This results in full-length protein synthesis with the amino acid change p.D2312V (Table 2).

***BRCA2* c.7976+3_7976+4del:** The DSS of intron 17 has a rare consensus splice-motif (AG|GC) and only SpliceSiteFinder-like, takes these non-canonical splice sites into consideration (Table 1). It was predicted that the variant would disrupt this donor site. Amplification of the cDNA using primers flanking exon 17 (Fig. 5A) and sequencing revealed exon 17 in-frame deletion in the patient's sample (Fig. 5B) (Table 2). Subsequently, we performed ASPCR together with transcript-specific primers. We used allele-specific primers for polymorphism c.7242A>G, for which both patient and controls were heterozygous, in combination with primers specific for each transcript: one primer hybridising to exon 17 and another primer hybridising across the exons 16/18 boundary to amplify the full-length transcript and the aberrant transcript, respectively. We observed that only the allele with the variant c.7976+3_7976+4del (present on the allele with c.7242G, as determined by segregation in the family, not shown) gives rise to *BRCA2*Δ17. In contrast, the WT allele of the patient only gives rise to the full-length transcript (Fig. 5C). The *BRCA1*Δ17 transcript was not amplified in the control samples.

***BRCA2* c.8754+3G>C:** Amplification of the region flanking exon 21 (Fig. 6A) and sequencing revealed the use of a downstream cryptic splice-site at c.8754+46, resulting in retention of part of intron 21 (Fig. 6B). Consequently, a PTC is introduced at c.8754+10 (Table 2). Afterwards, we performed ASPCR (using the polymorphism c.7242A>G) in combination with primers specific either for the full-length transcript (across exons 21/22 boundary) or for the aberrant transcript (hybridising in the retained intronic region). In the family of this patient, the c.7242A allele segregates with the variant (data not shown). In Fig. 6C it is observed that the control sample has biallelic expression of the full-length transcript and no expression of the aberrant transcript. The patient sample has monoallelic expression of the full-length transcript and the variant allele (with c.7242A) gives rise only to the transcript with intron retention.

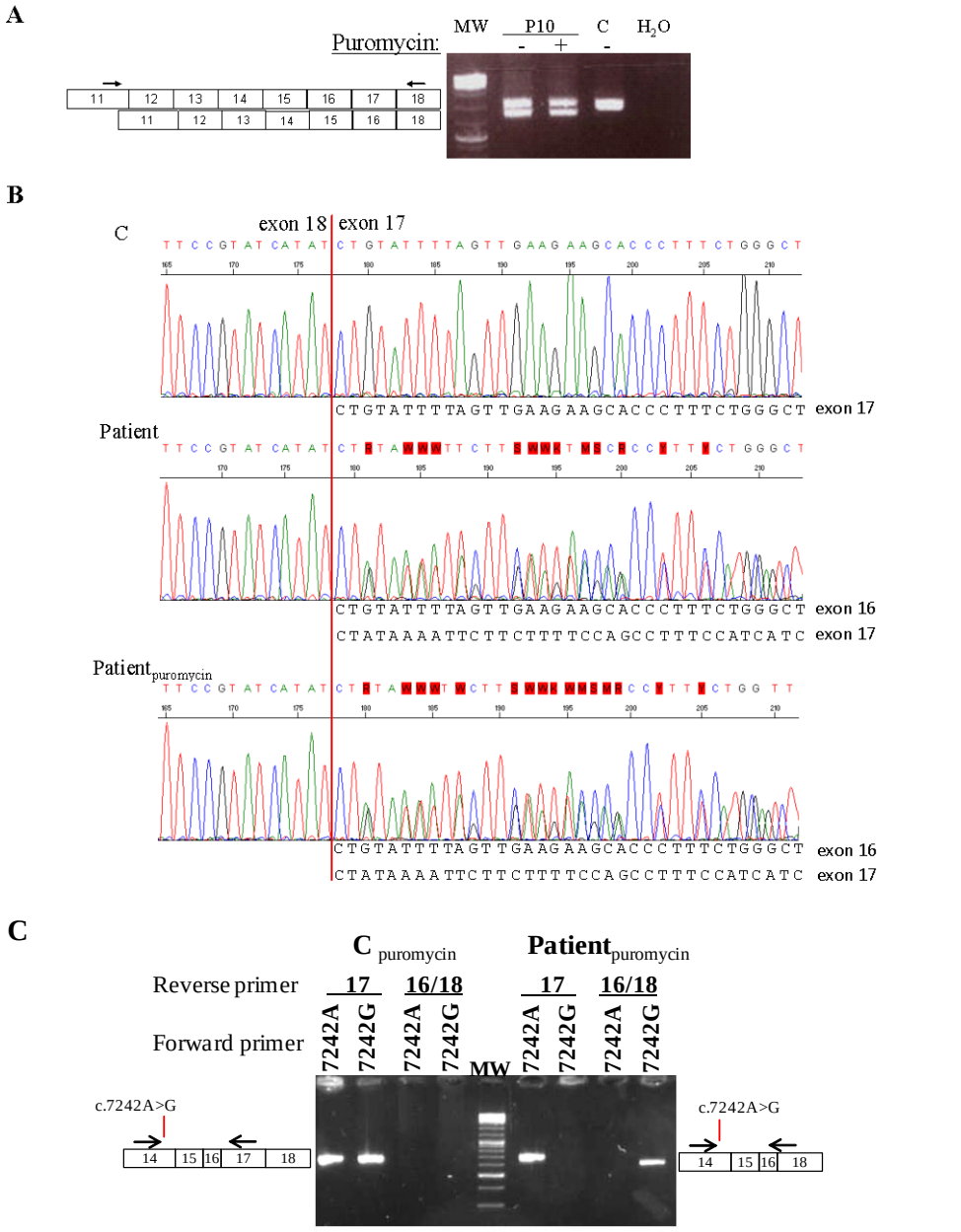


Figure 5. RT-PCR analysis of the variant *BRCA2* c.7976+3_7976+4del (P10). A) Results of the PCR flanking the exon 17. B) Reverse sequence of amplification products using primers flanking exon 17. C) Electrophoresis of the products obtained after amplification with allele-specific primers for 7242A and 7242G in combination with transcript-specific primers. Puromycin non-treated samples gave the same results (not shown). The names of the primers are shown. C: control sample; H_2O: negative PCR control; MW: molecular weight ladder: O'GeneRuler DNAladder mix (Fermentas). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

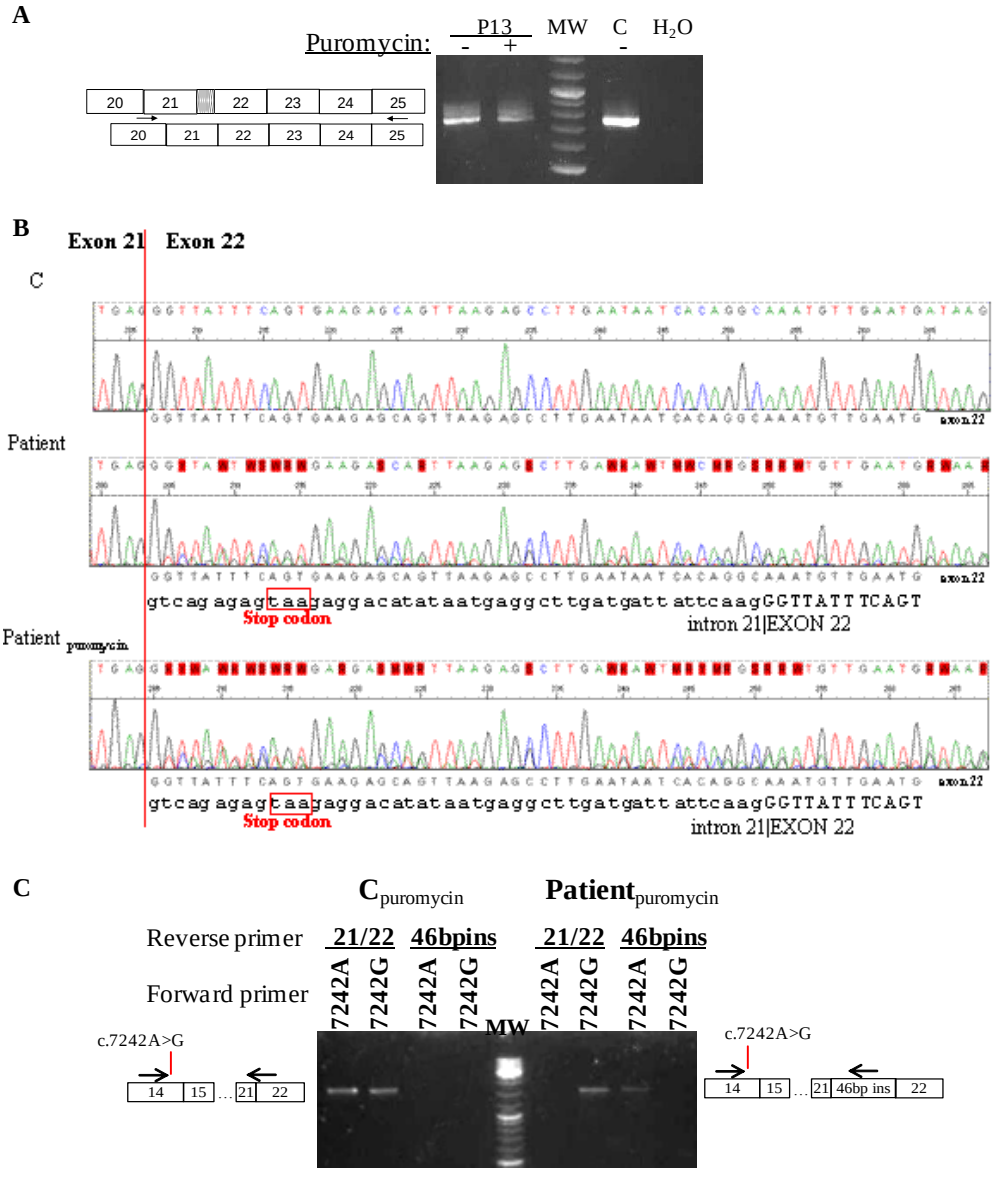


Figure 6. RT-PCR analysis of the variant *BRCA2* c.8754+3G>C (P13). A) Results obtained upon PCR with primers flanking exon 21. B) Sequencing of amplification products using primers flanking exon 21 showing the retention of 46 nt of intron 21. C) Electrophoresis of amplification products, when we used allele-specific primers for 7242A and 7242G in combination with transcript-specific primers. Puromycin non-treated samples gave the same results (not shown). The names of the primers are shown above the picture. C: control sample; H_2O: negative PCR control; MW: molecular weight ladder: O'GeneRuler DNAladder mix (Fermentas). Puromycin non-treated (-) and treated (+) samples are indicated in the pictures. Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

Variants with no effect on splicing:

The variants *BRCA1* c.4092C>G, *BRCA2* c.794-11T>C, c.6938-3T>C, c.6943A>G, c.8350C>T, c.8953+13A>G and c.8662C>T did not show aberrant transcripts that could be detected in our experiments, for both the puromycin treated and non-treated fractions (c.4092C>G shown in Fig. 1, data for other variants not shown). Biallelic expression was checked by amplification of the region of interest including a heterozygous polymorphism. This approach enabled the confirmation of biallelic expression for the majority of the variants, but was not possible for the intronic *BRCA2* c.8953+13A>G variant due to the absence of a heterozygous polymorphism.

DISCUSSION

Pathogenic variants affecting splicing

Variants could be clearly classified as pathogenic if they met two criteria: 1) monoallelic expression of the full-length transcript, which should be from the normal allele; 2) the variant allele gives rise to an aberrant transcript containing a PTC. The variant *BRCA2* c.425G>T, which causes a frameshift exon 4 deletion, met these criteria. So did the variant c.8754+3G>C, for which our results are in agreement to those obtained for other variants affecting the DSS of exon 21 (Table 3) [17-19]. One variant that does not lead to a PTC but, nevertheless, is also considered pathogenic is *BRCA2* c.7976+3_7976+4del. We observed monoallelic expression of in-frame exon 17 deletion, an event previously described [20-22] (Table 3) and classified pathogenic based on functional data. Wu and colleagues have shown that BRCA2 protein lacking exon 17 has impaired function [22]. Namely, they have shown that the localization of the GFP-tagged mutant protein is mainly in the cytoplasm in >90% of the transfected cells and the conclusion of the MMC hypersensitivity, homology-directed repair and induction of centrosome amplification assays was that exon 17 deletion in BRCA2 protein inactivates its function and therefore the authors classified the variant *BRCA2* c.7976G>A (R2659K), which leads to complete exon 17 skipping, as deleterious (table 3).

Additionally, we consider the variant *BRCA1* c.4987-3C>G to be likely pathogenic, as only the carrier of this variant showed *BRCA1*Δ17, which introduces a PTC. However, we could not exclude residual expression of full-length transcript from the variant allele since an allele-specific analysis could not be performed. Additional evidence that supports its pathogenicity comes from the fact that 100% *BRCA1*Δ17 was also previously observed due to the presence of c.4987-5T>A (Table 3), which is located slightly further from the exon than c.4987-3C>G and was considered pathogenic [23].

Table 3. Splice variants with similar effects on splicing as variants reported in this paper and conclusions from literature.

Variants reported	Splicing event description	In-frame/ frameshift	Literature	Reported clinical relevance
BRCA1 c.4987-5T>A	Exon 17 skipping	Frameshift	[23]	Deleterious
BRCA2 c.6853A>G	Increased expression of isoform *BRCA2*Δ12	In-frame	[24]	Likely neutral
BRCA2 c.7976G>A and c.7976G>C	Exon 17 skipping	In-frame	[20-22]	Deleterious
BRCA2 c.8754G>A, c.8754+1G>C, c.8754+1G>A , c.8754+4 A>G, c.8754+5G>A, and c.8754+5G>T	Retention of 46bp of intron 21	Frameshift	[1, 17-19]	Deleterious

Unclear variants affecting splicing

The variants *BRCA1* c.692C>T and c.693G>A lie within a region previously described as "critical region" which spans between codons 200 to 300 [25]. Orban and Olah suggested, based on *in silico* analysis, that there were two ESE motifs in this region, one of these covering positions c.690 to c.695 [26]. Our experimental data now provide evidence for the existence of an ESE motif in this region, since both c.692C>T and c.693G>A affect splicing by giving rise to increased expression of the *BRCA1*Δ11 transcript. This in-frame transcript was previously shown to be present in several human tissues, including T-cells and normal breast tissue [15]. In addition, the variant c.692C>T causes the amino acid change p.T231M in the full-length transcript, whereas c.693G>A is a silent variant. With the current knowledge, there is also not sufficient evidence that the missense change is pathogenic (Supplementary Table 3).

Another variant that remains unclear is the *BRCA2* c.6935A>T, which was reported five times in BIC database as a missense variant (p.D2312V) of unknown clinical relevance (Table 1 and Supplementary Table 3). Easton and colleagues reported for this variant absence of severe cancer history or strong co-segregation of the variant with the disease in families studied [27]. Our patient with the variant c.6935A>T developed cancer above the age of 50, as well as her mother and sister (Supplementary Table 1). However, we were unable to perform co-segregation analysis in the relatives, since DNA was unavailable. In this study, we observed that, besides the missense substitution, this variant gives rise to significantly increased expression of the naturally occurring *BRCA2*Δ12 isoform. Recently, similar effect was detected due to another variant (Table 3) [24]. Based on experimental data, the authors suggested that the *BRCA2* exon 12 is functionally redundant and that missense changes in this exon are therefore likely to be neutral.

Nevertheless, it is possible that BRCA2 exon 12 encoded region plays an important role in other biological processes that were not assessed.

The clinical relevance of the above mentioned variants remains unclear as the critical expression level of isoforms in the cells remains elusive. Furthermore, there is the possibility of a combined effect due to the presence of the full-length transcript with a missense variant, of which the previously reported variant *BRCA2* c.7988A>T [28] is an example.

Variants with no effect on splicing

The variants *BRCA1* c.4092C>G, *BRCA2* c.794-11T>C, c.6938-3T>C, c.6943A>G, c.8350C>T, c.8953+13A>G and c.8662C>T did not have an effect on splicing in our study. We assume that with our experimental set-up, including NMD-inhibition, we were able to detect possible aberrant transcripts, even without a heterozygous polymorphism to confirm biallelic expression, as for *BRCA2* variant c.8953+13A>G.

The intronic variants without a splicing effect are neutral. However, for missense variants, we cannot exclude an effect on protein structure and/or function. For these variants, we analysed the degree of conservation of the residue across species, changes in polarity, GMS (Grantham Matrix Score) [29], effect of the amino acid substitution on the protein and literature reports (Supplementary Table 3). Supplementary Table 1 shows the BRCAPRO scores of the probands used in our study, their personal cancer history and cancer history of their first- and second-degree relatives. We were not able to perform informative co-segregation studies, as relatives were either not available or not willing to participate in the study. These data do not allow to definitely classify these missense substitutions as either neutral or pathogenic. Further functional studies such as centrosome amplification, homologous recombination repair activity or gene expression profiling could be useful to draw conclusions [21, 22, 30, 31].

Transcript enrichment by NMD inhibition

The most significant enrichment because of NMD inhibition was seen for the variant *BRCA1* c.4987-3C>G, leading to exon 17 skipping. Whereas for two other variants also leading to PTCs, no clear (*BRCA2* c.425G>T) or only modest (*BRCA2* c.8754+3G>C) enrichment of aberrant transcripts was observed in puromycin-treated fractions. NMD normally degrades transcripts that contain a PTC located more than 50-55 nucleotides upstream of the last exon-exon junction [32]. Therefore, it was striking that the *BRCA2* c.425G>T variant, which introduces a PTC in codon 115, did not trigger NMD. A possible explanation for this is that the PTC lies close to the translation start codon and might escape NMD initiation because of a closed-loop structure and/or translation is initiated downstream of the PTC, as also observed for c.68_69del (185delAG) and c.71_81del (188del11), which cause a PTC in exon 3 of *BRCA1* [33]. Additionally, NMD does not usually downregulate the expression of aberrant transcripts completely or transcripts escape NMD [34-36]. Nevertheless, NMD-inhibition should be included in the experimental set-

up, especially in those cases for which biallelic mRNA expression cannot be confirmed, in order to avoid false-negative results.

Naturally occurring isoforms

The lack of knowledge regarding the expression level and function of the naturally occurring isoforms was a limitation in this study as we could not predict pathogenicity of three out of seven splice variants. So far, 18 *BRCA1* [37] and references therein and 8 *BRCA2* [16, 38-40] transcript isoforms of unknown function have been identified both in normal and tumour tissues. However, little is known about their critical expression levels, expression variation among controls and their functions. In this study, we have identified three variants (c.692C>T, c.693G>A and c.6935A>T) that increase the expression level of isoforms also present in controls. Although the Sanger sequencing method is semi-quantitative, we analysed the same event with different primer combinations in an allele-specific way. The difference in expression of the isoform transcript was determined relative to the full-length transcript. Based on results pointing in the same direction we concluded that the expression of the isoforms was increased in the presence of the variant. These results indicate the need to comprehend in more detail the complete *BRCA1* and *BRCA2* transcript isoforms repertoire in order to understand and determine the effect of individual sequence variants and its possible consequences for the tumour suppressor function of BRCA1/2. New methodologies, in particular RNA-seq approaches will provide more detailed information about the naturally occurring isoforms and respective expression levels [41].

Correlation between *in silico* prediction and experimental outcome

The variants that were correctly predicted to have an effect on splice sites had three algorithms predicting at least 10% decrease of the WT score or at least two prediction algorithms showing that the WT splice site score decreased at least 20% of the total score of the scale. The exception was the variant c.7976+3_7976+4del, which affected a non-canonical splice site, only detected by one algorithm. Interestingly, exonic variants that affected ESEs were found to be conserved in mammals. However, we also found one conserved exonic variant that did not appear to affect splicing (c.4092C>G).

It was also reported that splice regulatory elements, such as ESEs, might be more relevant in weak splice sites and are more abundant in exons with weak ASSs [42]. In our data, we did not find a correlation between putative ESE variants without an effect on splicing with strong ASSs, or vice-versa. In conclusion, ESE prediction algorithms are sensitive but not specific. Currently, there are no practical guidelines that allow discriminating real relevant changes in ESEs from neutral ones. With the used algorithms almost every substitution could have an effect on splicing since changes in ESE motifs are frequently indicated, as also observed by others [43, 44]. Nevertheless, several variants within ESE sites, affecting splicing, have been described for *BRCA1/2* [21, 45-48].

In conclusion, the clinical relevance for six out of 14 putative splice variants was clearly clarified. This greatly improves the genetic counselling of high-risk breast/ovarian

cancer patients carrying the classified variants. One additional variant is likely pathogenic but 100% exon skipping for the variant allele needs to be confirmed. Three other splice events remain unclassified because they do not affect the expression levels of full-length transcripts, but affect the expression level of transcripts that are also present in control samples, albeit at lower expression levels. To improve the assessment of the clinical relevance of such variants, the *BRCA1* and *BRCA2* normal transcription repertoire needs to be mapped in more detail both qualitatively and quantitatively, e.g. using RNA-seq methodology.

REFERENCES

1. The Breast Information Core Database, http://research.nhgri.nih.gov/bic/.
2. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-298.
3. Shapiro, M.B. and P. Senapathy, *RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implcations in gene expresson.* Nucleic Acids Res, 1987. **15**(17): p. 7155-7174.
4. Yeo, G. and C.B. Burge, *Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals.* J Comput Biol, 2004. **11**(2-3): p. 377-394.
5. Reese, M., F. Eeckman, D. Kulp, and D. Haussler, *Improved splice site detection in Genie.* J Comput Biol, 1997. **4**(3): p. 311-23.
6. Pertea, M., X. Lin, and S.L. Salzberg, *GeneSplicer: a new computational method for splice site prediction.* Nucleic Acids Res, 2001. **29**(5): p. 1185-1190.
7. Cartegni, L., J. Wang, Z. Zhu, M.Q. Zhang, and A.R. Krainer, *ESEfinder: a web resource to identify exonic splicing enhancers.* Nucleic Acids Res, 2003. **31**(13): p. 3568-3571.
8. Fairbrother, W.G., G.W. Yeo, R. Yeh, et al., *RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons.* Nucleic Acids Res, 2004. **32**(Web Server Issue): p. W187–190, doi: 10.1093/nar/gkh393
9. Gómez García, E., T. Ambergen, M.J. Blok, and A. Wijngaard, *Patients with an unclassified genetic variant in the BRCA 1 or BRCA2 genes show different clinical features from those with a mutation.* J Clin Oncol, 2005. **23**: p. 2185-2190.
10. LOVD - Leiden Open Variation Database, http://chromium.liacs.nl/LOVD2/cancer/home.
11. Dunnen, J.T.d. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.
12. Dunnen, J.T.d. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion.* Human Mutation, 2000. **15**(1): p. 7-12.
13. Carter, M.S., J. Doskow, P. Morris, et al., *A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo Is reversed by protein synthesis inhibitors in vitro.* J Biol Chem, 1995. **270**(48): p. 28995-29003.
14. Yu, C.-E., B. Devlin, N. Galloway, E. Loomis, and G.D. Schellenberg, *ADLAPH: A molecular haplotyping method based on allele-discriminating long-range PCR.* Genomics, 2004. **84**(3): p. 600-612.
15. Thakur, S., H.B. Zhang, Y. Peng, et al., *Localization of BRCA1 and a splice variant identifies the nuclear localization signal.* Mol Cell Biol, 1997. **17**(1): p. 444-452.
16. Bieche, I. and R. Lidereau, *Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue.* Cancer Res, 1999. **59**(11): p. 2546-2550.
17. Bonatti, F., C. Pepe, M. Tancredi, et al., *RNA-based analysis of BRCA1 and BRCA2 gene alterations.* Cancer Genet Cytogenet, 2006. **170**(2): p. 93-101.
18. Hansen, T.V.O., M.L. Bisgaard, L. Jønson, et al., *Novel de novo BRCA2 mutation in a patient with a family history of breast cancer*, in *BMC Med Genet*. 2008.
19. Vreeswijk, M., J. Kraan, H.v.d. Klift, et al., *Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs.* Hum Mutat, 2008. **30**(1): p. 107-114.
20. Hofmann, W., D. Horn, C. Huttner, E. Classen, and S. Scherneck, *The BRCA2 variant 8204G>A is a splicing mutation and results in an in frame deletion of the gene.* J Med Genet, 2003. **40**(3): p. e23.
21. Farrugia, D.J., M.K. Agarwal, V.S. Pankratz, et al., *Functional Assays for Classification of BRCA2 Variants of Uncertain Significance.* Cancer Res, 2008. **68**(9): p. 3523-3531.
22. Wu, K., S.R. Hinson, A. Ohashi, et al., *Functional Evaluation and Cancer Risk Assessment of BRCA2 Unclassified Variants.* Cancer Res, 2005. **65**(2): p. 417-426.
23. Bonnet, C., S. Krieger, M. Vezain, et al., *Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene.* J Med Genet, 2008. **45**(7): p. 438-446.

24. Graham, J.D., P.A. Mote, U. Salagame, et al., *DNA replication licensing and progenitor numbers are increased by progesterone in normal human breast.* Endocrinology, 2009. **150**(7): p. 3318-3326.
25. Hurst, L.D. and C. Pál, *Evidence for purifying selection acting on silent sites in BRCA1.* Trends Genet, 2001. **17**(2): p. 62-65.
26. Orban, T.I. and E. Olah, *Purifying selection on silent sites - a constraint from splicing regulation?* Trends Genet, 2001. **17**(5): p. 252-253.
27. Easton, D.F., A.M. Deffenbaugh, D. Pruss, et al., *A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast bancer predisposition genes.* Am J Hum Genet, 2007. **81**(5): p. 873-883.
28. Walker, L.C., P.J. Whiley, F.J. Couch, et al., *Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity.* Hum Mutat, 2010. **31**(6): p. E1484-E1505.
29. Grantham, R., *Amino acid difference dormula to help explain protein evolution.* Science, 1974. **185**(4154): p. 862-864.
30. Lovelock, P., A. Spurdle, M. Mok, et al., *Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants?* Breast Cancer Res, 2007. **9**(6): p. R82.
31. Kote-Jarai, Z., L. Matthews, A. Osorio, et al., *Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage.* Clin Cancer Res, 2006. **12**(13): p. 3896-3901.
32. Brogna, S. and J. Wen, *Nonsense-mediated mRNA decay (NMD) mechanisms.* Nat Struct Mol Biol, 2009. **16**(2): p. 107-113.
33. Nicholson, P., H. Yepiskoposyan, S. Metze, et al., *Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors* Cell Mol Life Sci, 2010. **67**(5): p. 677-700.
34. Perrin-Vidoz, L., O.M. Sinilnikova, D. Stoppa-Lyonnet, G.M. Lenoir, and S. Mazoyer, *The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons.* Hum Mol Genet, 2002. **11**(23): p. 2805-2814.
35. Anczuków, O., M.D. Ware, M. Buisson, et al., *Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins?* Hum Mutat, 2008. **29**(1): p. 65-73.
36. Ware, M.D., D. DeSilva, O.M. Sinilnikova, et al., *Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene?* Oncogene, 2006. **25**(2): p. 323-328.
37. Orban, T.I. and E. Olah, *Emerging roles of BRCA1 alternative splicing.* Mol Pathol, 2003. **56**(4): p. 191-197.
38. Speevak, M.D., S.S. Young, H. Feilotter, and P. Ainsworth, *Alternatively spliced, truncated human BRCA2 isoforms contain a novel coding exon.* Eur J Hum Genet, 2003. **11**(12): p. 951-954.
39. Tesoriero, A.A., E.M. Wong, M.A. Jenkins, et al., *Molecular characterization and cancer risk associated with BRCA1 and BRCA2 splice site variants identified in multiple-case breast cancer families.* Hum Mutat, 2005. **26**(5): p. 495.
40. Diez, O., S. Gutierrez-Enriquez, T. Ramon y Cajal, et al., *Caution should be used when interpreting alterations affecting the exon 3 of the BRCA2 gene in breast/ovarian cancer families.* J Clin Oncol, 2007. **25**(31): p. 5035-5036.
41. Ozsolak, F. and P.M. Milos, *RNA sequencing: advances, challenges and opportunities.* Nat Rev Genet, 2011. **12**(2): p. 87-98.
42. Wu, Y., Y. Zhang, and J. Zhang, *Distribution of exonic splicing enhancer elements in human genes.* Genomics, 2005. **86**(3): p. 329-336.
43. Anczuków, O., M. Buisson, M.-J. Salles, et al., *Unclassified variants identified in BRCA1 exon 11: Consequences on splicing.* Genes Chromosomes Cancer, 2008. **47**(5): p. 418-426.
44. Whiley, P., C. Pettigrew, B. Brewster, et al., *Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts.* BMC Med Genet, 2010. **11**(1): p. 80.

45. Gaildrat, P., S. Krieger, J.-C. Théry, et al., *The BRCA1 c.5434C>G (p.Pro1812Ala) variant induces a deleterious exon 23 skipping by affecting exonic splicing regulatory elements.* J Med Genet, 2010. **47**(6): p. 398-403.
46. Millevoi, S., S. Bernat, D. Telly, et al., *The c.5242C>A BRCA1 missense variant induces exon skipping by increasing splicing repressors binding.* Breast Cancer Res Treat, 2010. **120**(2): p. 391-399.
47. Sanz, D.J., A. Acedo, M. Infante, et al., *A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients.* Clin Cancer Res, 2010. **16**(6): p. 1957-1967.
48. Fackenthal, J.D., L. Cartegni, A.R. Krainer, and O.I. Olopade, *BRCA2 T2722R is a deleterious allele that causes exon skipping.* Am J Hum Genet, 2002. **71**(3): p. 625–631.
49. Parmigiani, G., D.A. Berry, and O. Aguilar, *Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2.* Am J Hum Genet, 1998. **62**: p. 145 - 158.
50. Judkins, T., B.C. Hendrickson, A.M. Deffenbaugh, et al., *Application of embryonic lethal or other obvious phenotypes to characterize the clinical significance of genetic variants found in trans with known deleterious mutations.* Cancer Res, 2005. **65**: p. 10096 - 10113.
51. Easton, D.F., A.M. Deffenbaugh, D. Pruss, et al., *A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes.* Am J Hum Genet, 2007. **81**(5): p. 873 - 883.
52. Farrugia, D.J., M.K. Agarwal, V.S. Pankratz, et al., *Functional assays for classification of BRCA2 variants of uncertain significance.* Cancer Res, 2008. **68**: p. 3523 - 3531.
53. Gomez Garcia, E., J. Oosterwijk, M. Timmermans, et al., *A method to assess the clinical significance of unclassified variants in the BRCA1 and BRCA2 genes based on cancer family history.* Breast Cancer Res, 2009. **11**(1): p. R8.
54. Karchin, R., M. Agarwal, A. Sali, F. Couch, and M.S. Beattie, *Classifying variants of undetermined significance in BRCA2 with protein likelihood ratios.* Cancer Informatics, 2008. **6**: p. 203-216.

SUPPLEMENTARY DATA

Supplementary Table 1. Clinical characteristics and family history of the patients studied

Patient no.	Sequence variant	BRCAPRO scores (%)[a]	Personal clinical characteristics[b]	Family history[b] (1st and 2nd degree)
BRCA1				
P1	c.692C>T	28.2	BBC43,47	na
P2	c.693G>A	60.5	BBC51,54	BC31
P3	c.4092C>G	25.0	asymptomatic	BC35, BC60
P4	c.4987-3C>G	96.9	BBC40,44, OC42	BC69
BRCA2				
P5	c.425G>T	11.9	BC41	BC42
P6	c.794-11T>C	20.8	BC53	BC31; BC92
P7	c.6935A>T	5.2	BC60	BC58; BC58
P8	c.6938-3T>C	37.6	BC41	BBC53,61
P9	c.6943A>G	20.3	BC40	OC35
P10	c.7976+3_7976+4del	37.2	BC35	BC62
P11	c.8350C>T	1.9	BC58	CC65; BC>50; BC>50; BC>50
P12	c.8662C>T	4.9	BC46	BC49
P13	c.8754+3G>C	89.3	BC39	BBC44,52; BC32
P14	c.8953+13A>G	36.5	BBC47,55	BC55

[a] BRCAPRO scores are expressed in percentage (up to 100%) and represent the probability of the proband being a carrier of a deleterious mutation in *BRCA1* or *BRCA2* genes [49].

[b] BC- breast cancer, BBC- bilateral breast cancer, OC- ovarian cancer, CC- colon cancer, with indication of age of onset. na- not available

Supplementary Table 2. Primer Sequences

Gene/ Name	**Sequence** (5' to 3')
BRCA1 primers	
10 Fw	ACAAATCACCCCTCAAGGAACCAG
10/12 Fw	ATTCTGCAAAAAAGGGTGACG
11 Fw	TGATGAAGAAAGAGGAACGGG
12 Rv	AAATGTCACTCTGAGAGGATAGCCC
14 Rv	TTCTGGCTTATAGGGTATTCACTACT
15/16 Fw	GCCAAGGCAAGATCTAGAGG
16/18 Fw	TGACCCCAGAAGAATTTATGCCG
18 Rv	CATTTTCCTCCCGCAATTC
BRCA2 primers	
3 Fw	AACTCCACAAAGGAAACCATC
3/5 Fw	AAATTCAAATTAGACTTAGTCCTGCTG
7 Rv	AGGATCCACCTCAGCTCCTA
8 Fw	AATGAAGAAGCATCTGAAACTGTA
10.01 Rv	TTCCAATGTGGTCTTTGCAG
10.02 Rv	ACGTGGCAAAGAATTCTCTGAAGTAA
11 Fw	ACAGATTCTAAACTGCCAAGTCATG
11 Rv	CAGAATCATTCTGTGAACAGC
11/13 Fw	AGCCCCTTATCTTAGTGGCCA
12 Fw	AAGGCTTCAAAAAGCACTCC
13 Fw	GAAGATTGTTTATGCATCATGTTTCTTTAG
14.01 Fw	ACAACTAAGGAACGTCAAG
14.02 Fw	CACAGAGTTGAACAGTGTGTTAGGA
7242A Fw	TGTTCCACCTTTTAAAACTAAATAA
7242G Fw	TGTTCCACCTTTTAAAACTAAATAG
7242A Rv	CACTGTTCAACTCTGTGAAAACGT
7242G Rv	CACTGTTCAACTCTGTGAAAACGC
14.01 Rv	TCTGCCTGTAGTAATCAAGTGTC
14.02 Rv	GCTTTTGTCTGTTTTCCTCCAA
16/18 Rv	CAATTTCCGTATCATATGTA
17 Rv	CCATAGCTGCCAGTTTCCAT
18 Fw	TGTTTCTGACATAATTTCATTGAGC
18 Rv	GCATACCACCCATCTGTAAGTTC
20 Fw	CGCAATGAAAGAGAGGAAGA
21 Fw	CAAGATGGTGCAGAGCTTTA
46bpIns Rv	TCATCAAGCCTCATTATATGTCC

21/22 Rv	GCTCTTCACTGAAATAACCCACA
22 Fw	TGTCACAACCGTGTGGAAG
24 Rv	TGTCGCTGCTAACTGTATGT
25 Rv	CGTCTGACAAATAGACGAAAGG

The deliberately introduced mismatch nucleotides are underlined. These were used in allele- and transcript-specific primers.

Supplementary Table 3. Prediction of the effect of the missense changes on the BRCA1 and BRCA2 proteins.

Variant	Effect on protein	Polarity change	GMS[a]	Conservation mammals/ other[b]	PolyPhen/ SIFT/ Align GVGD prediction[c]	Combined likelihood ratio (family history, co-segregation, co-occurence, literature)[d]	Classification (literature)
BRCA1							
c.692C>T	p.T231M	Y	81	N/N	benign/tolerated/ neutral	-	Uncertain [50]
c.4092C>G	p. N1364K	N	94	N/N	possibly damaging/ not tolerated/neutral	-	-
BRCA2							
c.6935A>T	p.D2312V	Y	152	Y/N	probably damaging/ tolerated/ neutral	$2.69x10^{-5}$ [51]	Neutral [51]
c.6943A>G	p. I2315V	N	29	N/N	benign/ tolerated/ neutral	-	-
c.8350C>T	p.R2784W	Y	101	Y/Y	probably damaging/ tolerated/ neutral	1.17 [52]	Uncertain [52, 53], deleterious [54]
c.8662C>T	p.R2888C	N	180	N/Y	benign/tolerated/ unclassified	$7.94x10^{-4}$ [51]	Neutral [51, 53, 54]

[a] Grantham matrix score[29]

[b] Alignments were based on following species (named by their common name) and NCBI reference sequences: BRCA1: Human (NP_009225), Chimpanzee (NP_001038958), Gorilla (AAT44835), Orang (AAT44834), Macaque (NP_001108421), Dog (NP_001013434), Rat (NP_036646), Mouse (NP_033894), Cow (NP_848668), Opossum (NP_001029141), Chicken (NP_989500), Frog (AAI70141); BRCA2: Human (NP_000050.2), Chimpanzee (XP_509619), Macaque (XP_001118184), Dog (BAB91245), Cat (NP_001009858), Cow (XP_583622), Rat (AAB71378) Mouse (NP_033895), Opossum (ABP48762), Chicken (NP_989607), Frog (EF508681), See urchin (EF523433). Y=Yes, N=No

[c] Websites for Polyphen: http://genetics.bwh.harvard.edu/pph/; SIFT: http://blocks.fhcrc.org/sift/SIFT.html; Align GVGD: http://agvgd.iarc.fr/agvgd_input.php

[d] Combined likelihood ratio >1000: deleterious variants; combined likelihood ratio <0.01: neutral variants.

CHAPTER 4.2

BRCA1 c.4987-3C>G is a pathogenic mutation

Rita Dias Brandão, Kees van Roozendaal, Demis Tserpelis, Beppy Caanen, Encarna Gómez García, Marinus J. Blok

Breast Cancer Research and Treatment 2012, 131(2):723-5
Letter to the Editor

In our previously published article by Brandão et al. [1] , we reported that the unclassified variant (UV) BRCA1 c.4987-3C>G gives rise to BRCA1D17. However, since we could not exclude residual expression of the full-length transcript from the variant allele, we have classified the variant as likely pathogenic. We now have evidence that the variant is indeed pathogenic.

Allele-specific PCRs are useful to determine the relative contribution of each allele to the synthesis of full-length or alternative transcripts. This can be achieved with, for instance, Sanger sequencing of the RT-PCR products and the use of a heterozygous variant to establish allelic expression ratios [1]. Unfortunately, in this family neither the proband nor the relatives tested for the BRCA1 c.4987-3C>G variant were heterozygous for a polymorphism in the region of the spliced exon. Since the UV is not exonic, it could not be used to determine allelic expression. Consequently, in our previously published article, using RT-PCRs on RNA isolated from primary cultured lymphocytes, we failed to determine whether the UV allele was still giving rise to some full-length BRCA1 transcript. To establish the pathogenicity of this variant, we performed an additional study using an ex vivo assay based on a splicing reporter minigene. We selected the exon-trapping vector pSPL3b [2] (a kind gift from Dr. R Sedlmeier, Ingenium Pharmaceuticals GmbH), previously used for similar studies [3-5]. Initially, BRCA1 exon 17 was amplified including the surrounding intronic regions from the proband's DNA and one additional control, using primers that contained restriction sites XhoI and EcoRV in their 50 ends.

PCR products and vector were digested with the two enzymes and subsequently used in a ligation reaction, after purification with QIAquick PCR Purification Kit (QIAGEN). The structure of the minigenes is shown in Fig. 1a. After transformation of competent E. coli DH5a (Invitrogen), independent colonies were selected with minigenes containing the variant allele from the patient, WT alleles from the patient and from the controls. All constructs were verified by sequence analysis and confirmed the absence of changes in the constructs. These minigenes and the empty vector were transfected into HeLa cells. Transfection was performed in duplicate with FuGENE HD transfection reagent (Promega) according to the manufacturer's protocol. After RNA extraction from the transfected HeLa cells and reverse transcription, the cDNA was amplified using primers in each of the flanking exons of the vector pSPL3b, which we named A and B for clarity (Fig. 1a). The results obtained are shown in Fig. 1b. The WT allele of the patient and of a healthy control show only one band in the agarose gel, with a larger size than that observed for the empty vector. Sequencing confirmed that the RT-PCR products contain exon 17 (Fig. 2). The variant allele of the patient shows a prominent lower band of the same size as the empty vector indicating exon 17 skipping, which was confirmed by sequencing. In addition, a faint unexpected upper band can be observed for the variant allele. Sequencing of the two bands separately, by excision of the bands from the gel, revealed this to be a transcriptcontaining part of intron 17.

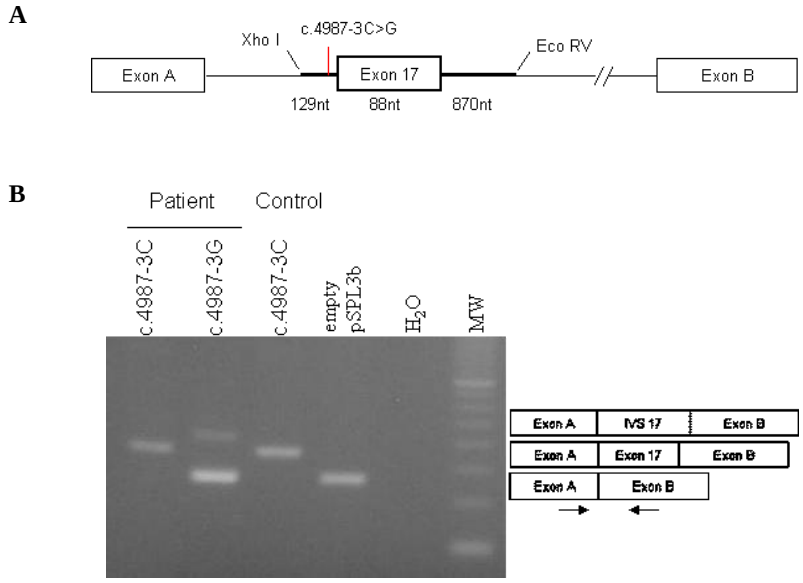


Figure 1. *Ex vivo* assay using pSPL3b vector. a) Structure of the pSPL3b vector containing exon 17 and flanking intronic regions with the mutant or wild-type allele (c.4987-3C>G). The size of the exon, intronic regions and the restriction sites included in the primers are indicated. b) Transcript analysis by electrophoresis of the RT-PCR products obtained after transfection of HeLa cells using primers hybridizing to the exons of the vector. H_2O: negative PCR control; MW: molecular weight ladder XIV (Roche). Boxes next to the PCR bands indicate the exon composition and the position of the primers used is shown with arrows.

In Figure 2, the sequence of this transcript is shown but it contains a background sequence of the transcript lacking exon 17 (the lower band on the gel). This is due to heteroduplex formation between the two fragments, which occurs due to the high similarity between them as both contain exons A and B. The intron 17 retention starts at the beginning of intron 17, position c.5074+1, and the donor splice site is at position c.5074+153. This is possible since the original donor splice site of exon 17 is also predicted to be a strong acceptor splice site (AG|GTATAC, 76% score) by the Splice Site Finder-like algorithm and Human Splice Finder (recently added to the Alamut software), as observed using Alamut (Interactive Biosoftware), which integrates several splice site prediction algorithms. The c.5074+153 is predicted to be a weak donor splice site by two algorithms (5 and 11% by MaxEntScan and GeneSplicer, respectively) but strongly predicted by the Human Splice Finder algorithm (79%). This intronic region is not observed in the WT alleles used in the ex vivo assay, neither in the previous results from the IL2/PHA stimulated lymphocytes. We conclude that the intron retention is an artefact in the in vitro system caused by skipping of exon 17 and activation of cryptic splice sites.

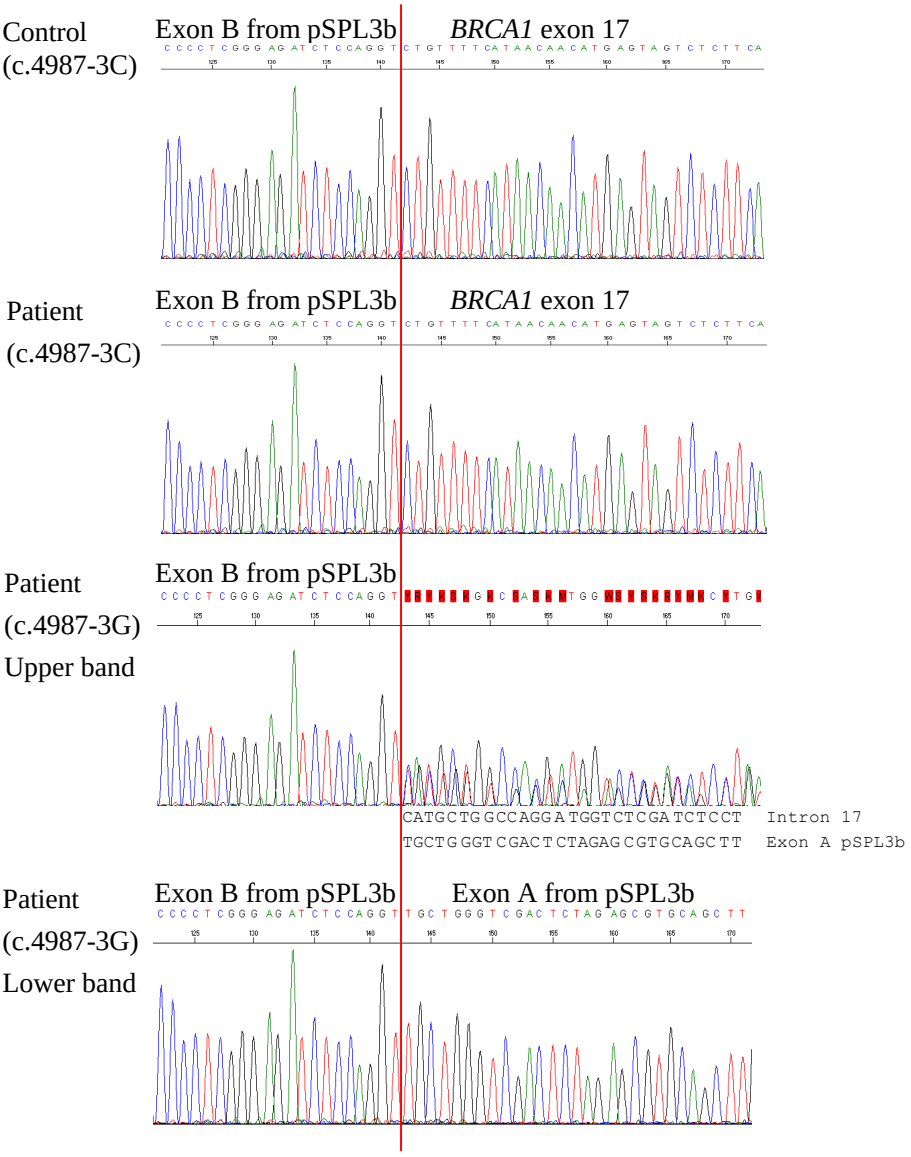


Figure 2. Reverse sequence of amplification products observed in Fig.1B, using primers in the exons A and B of the pSPL3b vector. The two fragments of the sample containing the c.4987-3G variant were excised from the gel and sequenced separately. As expected, both the samples with WT sequence c.4987-3C contain exon 17. Vector containing the c.4987-3G variant gave rise to a transcript where the exons of the vector are adjacent, revealing exon 17 skipping (lower band Fig.1b), and another transcript with inclusion of part of intron 17 (upper band Fig.1b). The latter contained a heteroduplex of the two described transcripts, which formed due to the high similarity between the two fragments.

Summarizing, using an ex vivo assay to complement the previous RT-PCR analysis on RNA from IL2/PHA stimulated lymphocyte cultures [1], we were able to show that the variant allele from the patient results only in exon 17 skipping since a transcript containing exon 17 was not detected. The deletion of BRCA1 exon 17 is a frameshift event that leads to a truncated protein: p.Val1665SerdelfsX9. In combination with the results from the previously reported RT-PCR analysis, we are now confident that the BRCA1 c.4987-3C>G variant is pathogenic and can be genetically counselled as such.

ACKOWLEDGMENTS

We would like to thank Alberto Acedo for the helpful discussion about the *ex vivo* assay.

REFERENCES

1. Brandão, R.D., K. van Roozendaal, D. Tserpelis, E.G. García, and M.J. Blok, *Characterisation of unclassified variants in the BRCA1/2 genes with a putative effect on splicing.* Breast Cancer Res Treat, 2011. **129**(3): p. 971-82.
2. Burn, T.C., T.D. Connors, K.W. Klinger, and G.M. Landes, *Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector.* Gene, 1995. **161**(2): p. 183-187.
3. Schneider, B., A. Koppius, and R. Sedlmeier, *Use of an exon-trapping vector for the evaluation of splice-site mutations.* Mamm Genome, 2007. **18**(9): p. 670-676.
4. Whiley, P., C. Pettigrew, B. Brewster, et al., *Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts.* BMC Med Genet, 2010. **11**(1): p. 80.
5. Hansen, T., A. Steffensen, L. Jønson, et al., *The silent mutation nucleotide 744 G→A, Lys172Lys, in exon 6 of BRCA2 results in exon skipping.* Breast Cancer Res Treat, 2010. **119**(3): p. 547-550.

CHAPTER 5

Detection of exon skipping events in *BRCA1* RNA using MLPA kit P002

Rita D. Brandão, Demis Tserpelis, Encarna Gómez García, Marinus J. Blok

Molecular Biology Reports 2012, 39(7):7429-33

ABSTRACT

A rapid and easy method to screen for aberrant cDNA would be a very useful diagnostic tool in genetics since a fraction of the DNA variants found affect RNA splicing. The currently used RT-PCR methods require new primer combinations to study each variant that might affect splicing. Since MLPA is routinely used to detect large genomic deletions and successfully detected exon skipping events in Duchenne muscular dystrophy in cDNA, we performed a pilot study to evaluate its value for BRCA1 cDNA.

The effect of puromycin, DNase I and two different DNA cleaning protocols were tested in the RNA analysis of lymphocyte cultures. We used two samples from unrelated families with two different *BRCA1* exon deletion events, two healthy unrelated controls and six samples from hereditary breast/ovarian cancer syndrome (HBOC) patients without *BRCA1/2* mutations.

Using RNA treated with DNase I and cleaned in a column system from puromycin-treated fractions, we were able to identify the two *BRCA1* deletions. Additional HBOC patients did not show additional splice events. However, we were not able to get reproducible results.

The cDNA-MLPA technique using kit *BRCA1* P002 is in our hands currently not reliable enough for routine RNA analysis and needs further optimization.

INTRODUCTION

Genetic screening of the *BRCA1 and BRCA2* genes is offered to families with high risk of breast and ovarian cancer. Besides clear pathogenic mutations and polymorphisms, unclassified variants (UVs) of unclear clinical relevance are found. Some of these UVs may result in aberrant splicing, by affecting the donor or acceptor splice sites, or exonic splice site enhancer (ESE) sites [1] as predicted *in silico*. Additionally, deep intronic variants, which are normally ignored, may also affect splicing. One example of a deep intronic pathogenic variant is the variant CDKN2A IVS2-105A>G, which causes retention of intronic sequence [2]. Another example is the mutation c.903+409T>C in the MTRR (methionine synthase reductase) gene, which activates a pseudoexon, causing a frameshift insertion that leads to a premature stop codon [3]. Experimental proof is needed to confirm the predicted changes in RNA splicing. The experiments are usually performed using RT-PCR, for which a set of specific primers targeted to the relevant cDNA region is needed for every new variant [4-7]. It is noteworthy that exon skipping is the most common alternative splice event [8]. After the report of Kesari et al. [9], who were able to detect skipping events on cDNA from the Duchenne muscular dystrophy (DMD) gene using the respective genomic MLPA kit, we sought to evaluate the use of a commercially available *BRCA1* MLPA kit [10] for the detection of exon skipping in cDNA instead of genomic DNA. *BRCA1* MLPA is a multiplex assay based on the hybridization of a large set of primers throughout the entire coding part of the *BRCA1* gene. Therefore the assay should potentially also be able to detect all exon skipping events in cDNA in the presence of a variant affecting splicing, without the need to design a specific RT-PCR assay for each variant. Although these are likely rare events, using a rapid and relatively cheap assay to assess them would be valuable in a diagnostic setting to rule out their presence.

For this pilot study, samples with *BRCA1* exon 13 skipping (c.4242-1643del3835) or exon 22 skipping (c.5333-36del510) [11] were selected. The study also included samples from 2 unrelated healthy controls and 6 samples from patients belonging to high risk families for which no *BRCA1* or *BRCA2* mutation was identified in the standard diagnostic screening. Here we show that the MLPA method was able to detect the skipping events, but it was not reproducible enough for use in clinical testing despite the optimization attempts which are here described.

MATERIALS AND METHODS

Cell culture

White blood cells were isolated and cultured in complete medium consisting of: RPMI 1640 supplemented with L-glutamine (Gibco) and 12.5% FCS with additional supplements and antibiotics. Lymphocyte growth was stimulated with 50 μL/mL PHA (Gibco) and 10 units/mL of IL-2 (Roche). At day 7, 4-6h before harvesting the cells, cultures were treated

with 200 μg/mL of puromycin (Sigma), to enrich for transcripts containing premature stop codons by the inhibition of NMD [12].

RNA isolation, cDNA synthesis and MLPA reaction

Total RNA was isolated using TRIzol (Invitrogen) or TRIpure (Roche) reagent. RNA samples used were either not subjected to DNase I treatment or treated with DNA-free kit (AMBION) or with DNase I followed by purification in the column system RNeasy MinElute Kit (Qiagen). First-strand cDNA was obtained with Reverse Transcriptase M-MUL (Finnzymes) using random hexamers (Invitrogen) following the manufacturers' instructions. The cDNA was amplified with the SALSA MLPA P002 probe mix (MRC-Holland) according to the manufacturer's protocol. Fragment analysis was performed by capillary electrophoresis in an ABI PRISM 3730 automatic sequencer (Applied Biosystems).

Data analysis

The size calling and the peak areas were assessed using the Genemarker software (Softgenetics) and exported to a ".txt" file. The values of the antisense probes were extremely low compared to the sense probes, and they don't have known biological meaning. Therefore, the data was filtered to leave only the data from probes corresponding in sequence to that of sense *BRCA1* mRNA. The normalization of the data was performed using a spreadsheet according to the Manual spreadsheet-based MLPA analysis instructions (available on the MRC-Holland website: www.MLPA.com). The threshold values for deletions and duplications were set to 0.75 – 1.25, respectively, which are also used for DNA analysis [13-16]

RESULTS

With the SALSA MLPA P002 kit, strong signals were obtained for 21 out of 25 probes. These probes contained more than 85% nucleotides hybridizing to the exon sequence in the correct orientation. The signals for the probes with less than 85% matching exonic sequence (exons 1A, 9 and 19) or in antisense (23) were extremely weak and often not even detectable by the software. This also confirms the absence of contaminating genomic DNA in the RNA samples.

Initially, we have compared the results from puromycin-treated and non-treated samples (Figure 1), without DNase I treatment. The results were not optimal, but it was observed that the puromycin-treated samples gave better results than the non-treated. Subsequently, we tested the effect of two different DNase I treatment options: 1) DNase I treatment followed by purification in a column system and 2) DNase I treatment kit that allows to remove the enzyme by precipitation and centrifugation. The results were considerably improved when the RNAs were cleaned in a column system (data not shown),

i.e. variation in the signals among individuals was greatly reduced, at least in two independent experiments.

Six samples from high risk families without a *BRCA1/2* mutation were also analyzed (data not shown) using the puromycin-treated fractions and RNAs treated with DNase I and cleaned in a column system. None of these samples showed an exon skipping event, in the 20 exons tested. However, in an independent third experiment we observed increased interindividual variability in some exon signals. Many exons had normalized values outside the 0.75-1.25 thresholds (Figure 2). This was also observed in healthy control samples. This hampers the evaluation of splicing defects as it suggests duplications or deletions events that would need experimental follow-up or repetitive MLPA analysis to determine reproducibility.

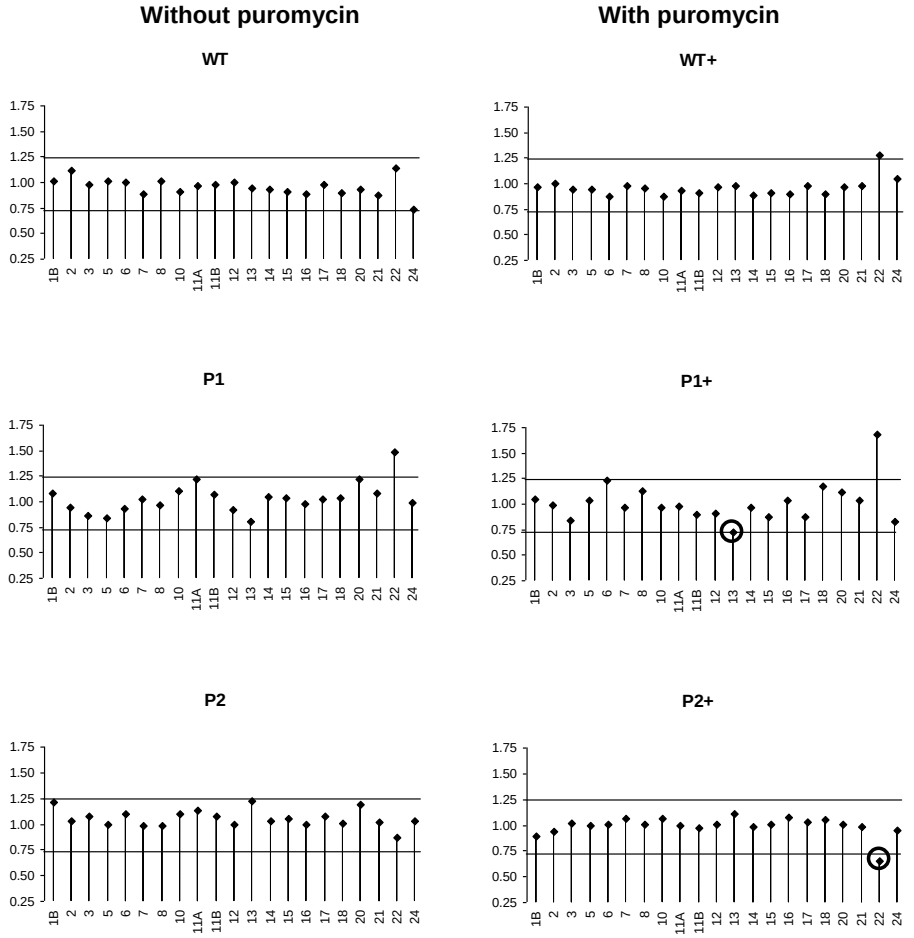


Figure 1. MLPA results obtained using puromycin-treated and non-treated samples as indicated. Healthy controls without BRCA1 mutations are indicated as WT, whereas P1 and P2 are positive controls with exon 13 and exon 22 deletion events, respectively.

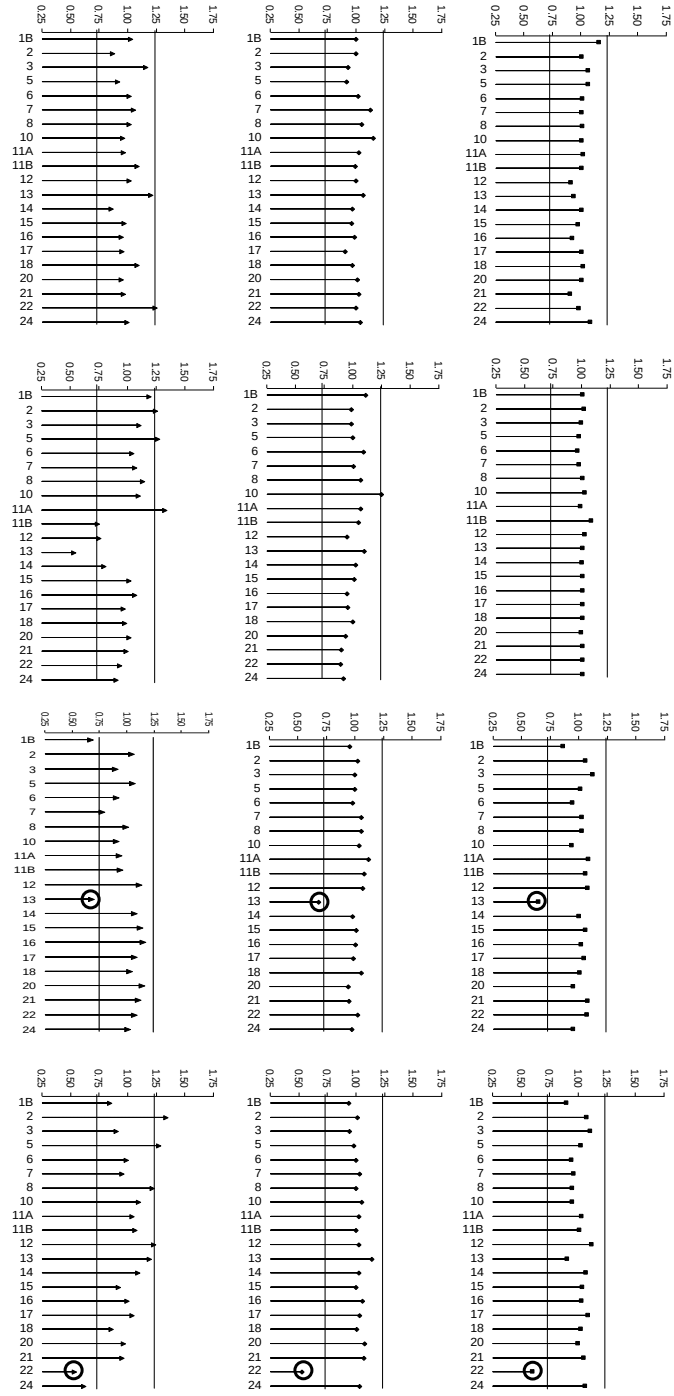


Figure 2. MLPA results obtained in three independent experiments for puromycin-treated fractions. Healthy controls without BRCA1 mutations are indicated as WT; P1 and P2 are positive controls with exon 13 and exon 22 deletion events, respectively.

DISCUSSION

The MLPA method is widely used in diagnostics, mainly to test genomic events such as deletions and duplications. Although there are a few commercial RT-MLPA kits, these are designed to test the expression of genes associated with certain biological processes, MRC-Holland has not developed RT-MLPA kits to test splice events. Besides the use of the MLPA, or other multiplex approaches, to test the effect of genetic variants predicted to affect splicing at the RNA level, it would be useful to test for *BRCA1* and *BRCA2* mutation negative patients with strong breast and/or ovarian cancer history. This group of patients may carry variants outside the screened intronic region flanking the exons which could affect splicing. Since exon skipping is the most common alternative splice event [8], developing a test that allows to screen for exon skipping events would detect the majority of alternative splice events.

One single study has previously shown that MLPA could be used to test exon skipping events in RNA transcripts of the *DMD* gene [9]. Here we report the use of MLPA kit for the analysis of BRCA1 exon skipping events. The most optimal results were obtained from puromycin-treated samples and when RNA was treated with DNase I and subsequently purified in a column system. However, despite efforts to optimize the technique further, we were not able to get reliable, reproducible results for unequivocal interpretation using the kit BRCA1 P002. This variation was also observed in healthy control samples, which showed both deletion and duplication events in one out of three experiments performed.

MLPA test is a flexible multiplex assay which allows for up to a total of 50 probes and in principle, it should be possible to use it for detection of alternative splicing events other than exon skipping. To be able to test also for intron retention or insertion of pseudoexons, probes crossing over exon-exon boundaries should also be included in the assay. Although mRNA-seq technology [17] will also allow to test for aberrant splicing events in patients, MLPA could be a more cost-effective technique. However, it needs to be optimized further for routine use.

REFERENCES

1. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-298.
2. Harland, M., S. Mistry, D.T. Bishop, and J.A. Newton Bishop, *A deep intronic mutation in CDKN2A is associated with disease in a subset of melanoma pedigrees.* Human Molecular Genetics, 2001. **10**(23): p. 2679-2686.
3. Homolova, K., P. Zavadakova, T.K. Doktor, et al., *The deep intronic c.903+469T>C mutation in the MTRR gene creates an SF2/ASF binding exonic splicing enhancer, which leads to pseudoexon activation and causes the cblE type of homocystinuria.* Hum Mutat, 2010. **31**(4): p. 437-444.
4. Bonatti, F., C. Pepe, M. Tancredi, et al., *RNA-based analysis of BRCA1 and BRCA2 gene alterations.* Cancer Genet Cytogenet, 2006. **170**(2): p. 93-101.
5. Bonnet, C., S. Krieger, M. Vezain, et al., *Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene.* J Med Genet, 2008. **45**(7): p. 438-446.
6. Whiley, P.J., L. Guidugli, L.C. Walker, et al., *Splicing and multifactorial analysis of intronic BRCA1 and BRCA2 sequence variants identifies clinically significant splicing aberrations up to 12 nucleotides from the intron/exon boundary.* Human Mutation, 2011. **32**(6): p. 678-687.
7. Brandão, R.D., K. van Roozendaal, D. Tserpelis, E.G. García, and M.J. Blok, *Characterisation of unclassified variants in the BRCA1/2 genes with a putative effect on splicing.* Breast Cancer Res Treat, 2011. **129**(3): p. 971-82.
8. Sammeth, M., S. Foissac, and R. Guigó, *A general definition and nomenclature for alternative splicing events.* PLoS Comput Biol, 2008. **4**(8), doi: 10.1371/journal.pcbi.1000147
9. Kesari, A., L.N. Pirra, L. Bremadesam, et al., *Integrated DNA, cDNA, and protein studies in Becker muscular dystrophy show high exception to the reading frame rule.* Hum Mutat, 2008. **29**(5): p. 728-737.
10. Hogervorst, F.B.L., P.M. Nederlof, J.J.P. Gille, et al., *Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method.* Cancer Res, 2003. **63**(7): p. 1449-1453.
11. Petrij-Bosch, A., T. Peelen, M. van Vliet, et al., *BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients.* Nat Genet, 1997. **17**(4): p. 341-5.
12. Carter, M.S., J. Doskow, P. Morris, et al., *A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro.* J Biol Chem, 1995. **270**(48): p. 28995-29003.
13. Barbaro, M., M. Kotajärvi, P. Harper, and Y. Floderus, *Identification of an AluY-mediated deletion of exon 5 in the CPOX gene by MLPA analysis in patients with hereditary coproporphyria.* Clin Genet, 2011: p. no-no, doi: 10.1111/j.1399-0004.2011.01628.x
14. Harteveld, C.L., A. Voskamp, M. Phylipsen, et al., *Nine unknown rearrangements in 16p13.3 and 11p15.4 causing α and β-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification.* J Med Genet, 2005. **42**(12): p. 922-931.
15. Nyström, A.-M., S. Ekvall, A.-C. Thuresson, et al., *Investigation of gene dosage imbalances in patients with Noonan syndrome using multiplex ligation-dependent probe amplification analysis.* Euro J Med Genet, 2010. **53**(3): p. 117-121.
16. Kellander, M., M. Riley, and C. Liu, *Application Note: GeneMarker® Software for Multiplex Ligation-dependent Probe Amplification (MLPA™)*, SoftGenetics LLC.
17. Levin, J., M. Berger, X. Adiconis, et al., *Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts.* Genome Biol, 2009. **10**(10): p. R115.

CHAPTER 6

Exploring transcriptional changes in IL2/PHA stimulated *BRCA1*$^{+/-}$ lymphocytes after gamma-irradiation to construct a robust genetic classifier

Rita D. Brandão, Lars M.T. Eijssen, Demis Tserpelis, Dennie G.A.J. Hebels, Hubert J.M. Smeets, Encarna B. Gómez García, Marinus J. Blok

Submitted

ABSTRACT

Unclassified variants (UVs) detected during the genetic screening of *BRCA1/2* genes pose a challenge for interpretation and counselling. Agnostic approaches to classify UVs based on gene expression profiles, have not identified reliable genetic classifiers, as these appear to be too specific to the samples used, since they have little overlap and replication is difficult. Additionally, many genes in these classifiers have no functional relationship with the BRCA1/2 proteins.

In this study, the transcriptome of irradiated BRCA1-mutated ($BRCA1^{+/-}$) lymphocytes was explored to identify affected biological processes from which genes were retrieved, aiming at obtaining a more reproducible genetic classifier to distinguish pathogenic sequence variants from neutral ones. White blood cells from *BRCA1*-mutation carriers and controls were stimulated with IL2/PHA. Cells were harvested 2h post-irradiation and RNA was isolated for analysis on Human Gene 1.0st Affymetrix arrays. Differentially expressed genes were identified and used for subsequent pathway and network enrichment analysis.

We observed gene expression changes suggestive of deficient cell cycle arrest, decreased apoptotic activity, decreased immune response, increased chromosomal instability, and decreased mitotic spindle assembly and chromosome separation. A set of genes involved in micronuclei (MN) induction were differentially expressed, consistent with previous studies showing that MN count was a useful test to distinguish $BRCA1^{+/-}$ lymphocytes from controls. An additional group of *BRCA1*-mutation carriers and controls was subjected to the same procedure and analysed separately. Genes differentially expressed in both groups were retrieved to be included in a general applicable genetic classifier. The classifier was evaluated by unsupervised hierarchical cluster analysis including additional data from independently published studies.

The classifier allowed separate clustering of nearly all $BRCA1^{+/-}$ samples and controls from different cell lines and DNA-damage agents. The results in this study illustrate the relevance of exploring the biological processes when aiming for a reproducible genetic classifier.

INTRODUCTION

Germline mutations in the *BRCA1/2* genes lead to breast, ovarian and other cancers [1-4]. Genetic screening is offered to patients from families with high risk of breast and/or ovarian cancer. However, besides clearly pathogenic mutations and polymorphisms (present in at least 1% of the general population and generally considered neutral in terms of disease risk), unclassified variants (UVs) are also being identified. These UVs have unclear clinical significance and therefore hamper genetic counselling of the patient and the relatives at risk. Several functional assays were developed to test BRCA1 protein functions in the presence of mutations [5]. Since BRCA1 has been involved in many different biological processes (DNA damage repair, regulation of gene expression, cell cycle control during the S and G2/M checkpoints, chromatin remodelling, and ubiquitylation) there are many different functional tests. Each of them is time-consuming and requires specific expertise, which hampers use in a routine, clinical setting. In addition, these assays are limited to current knowledge of the BRCA1 functions and usually each of them interrogates only one specific function. As such, agnostic assays to test impaired BRCA1 function in response to gamma irradiation, or other DNA damage-inducing agents, using microarrays have been explored [6-9]. The rationale behind these studies is that the expression of certain genes is affected in the presence of a defective BRCA1 protein as compared to WT protein. These studies aimed at finding a genetic classifier to distinguish pathogenic mutations from neutral variants purely based on the most statistically significant changed genes. These studies, which used either fibroblasts or lymphoblastoid cell lines (LCLs), were rather small for classifier studies, i.e. n= 9 [6], n=10 [7], n=23 [8], n=9 [9], also taking into account the relatively small fold changes observed [10]. Such classifiers are in general too specific to the samples used and are difficult to replicate and, therefore, comparable studies obtain genetic classifiers with different sets of genes with little or no overlap [11, 12].

In this study, we explored the transcriptome of replicating $BRCA1^{+/-}$lymphocytes, in response to irradiation-triggered DNA damage. Our rationale was that irradiation would increase the specificity of the gene expression changes, generating a more accurate and robust classifier. IL2/PHA-stimulated T lymphocytes were used since the collection of lymphocytes is less invasive than skin biopsies for fibroblast culture. This stimulation also avoids EBV-induced immortalisation of B lymphocytes, which is not always successful and was shown to negatively affect the micronuclei induction (MN) test in *BRCA1*-mutated cells [13, 14]. MN test was described to be able to distinguish *BRCA1*-mutated cells from healthy control cells [15] and is affected by EBV-immortalization. Additionally, as other effects of the EBV transformation, e.g. in cell cycle, are not yet fully understood we avoided its use. Our study is the first using stimulated lymphocytes to analyse the transcriptome of *BRCA1*-mutation carriers in response to radiation, in combination with pathway analysis to identify genes from the biological processes involved. Additional genes were identified using a second independent group of samples subjected to the same treatment. Common differentially expressed genes were added to the initial set of genes

identified in pathway analysis. The final set of genes, composed of 160 genes, was used to cluster the samples from our experiments and online available data from previous independently published studies [7-9]. Unsupervised hierarchical clustering performed well across these different data sets, including samples that were treated with mitomycin C, instead of irradiation. This study shows that our approach has more potential to generate a robust classifier that can be used for the classification of UVs, than building classifiers based on the most differentially expressed genes.

MATERIALS AND METHODS

Ethics statement

All human biological material used in this study followed the guidelines of the Medical Ethics Committee of the Maastricht University Medical Centre +. *BRCA1/2*-mutation carriers gave informed consent.

Subjects

Ten *BRCA1*-mutation carriers were randomly selected among women who received genetic counselling and screening at the Maastricht University Medical Centre +. Peripheral whole blood from six anonymous female controls was collected via the local blood bank. The control samples did not carry *BRCA1/2* mutations in the coding and immediate flanking regions. Additional blood samples were collected from 15 *BRCA1*-mutation female carriers and five additional *BRCA1/2*-mutation negative female controls.

Lymphocyte culture

White blood cells were isolated, following erythrocyte lysis, from fresh whole peripheral blood (collected in EDTA tubes) and frozen in liquid nitrogen. Culture conditions were as previously reported [16]. Briefly, lymphocyte growth was stimulated with 50 μL/mL phytohemagglutinin (Gibco) and 10 units/mL of IL-2 (Roche) and cells were kept in RPMI 1640 supplemented with L-glutamine (Gibco) medium and 12.5% FCS, 1x L-glutamine, 0.8mM sodiumpyruvate (Gibco), 17mM Hepes buffer (Gibco), $4.2x10^{-2}$ mM 2-mercaptoethanol (Gibco), 42 units/mL penicillin-streptomycin, 10 units/mL of IL-2 (Roche), and 0.21 g/mL amphotericin B solution (Sigma). After 6 days, cells were irradiated with 10 Gγ and harvested 2h after irradiation.

RNA isolation and microarray expression profiling

Total RNA was isolated using TRIzol (Invitrogen), treated with DNase I using the RNase-Free DNase Set (Qiagen) and subsequently purified using the RNeasy Mini Kit (Qiagen). RNA quantity and purity were determined spectrophotometrically using the Nanodrop ND-1000 (Nanodrop Technologies). RNA integrity was assessed by determining the RNA 28S/18S ratio using the Bioanalyzer 2100 (Agilent Technologies). Two hundred ng of RNA

were reverse transcribed to double-stranded cDNA with the WT Expression kit (Ambion) according to the manufacturer's protocol. The cDNA was subsequently fragmented and labelled using the WT Terminal Labeling Kit (Affymetrix). Hybridisation to Affymetrix Human Gene 1.0 ST arrays and subsequent scanning was performed following the manufacturer's guidelines using the GeneChip scanner 3000 (Affymetrix). Annotation of the probes on the chip was updated by using the freely distributed Ensembl-based chip description file (CDF) of the microarray lab of the University of Michigan (http://brainarray.mbni.med.umich.edu; version 14) [17]. Arrays were scanned for the first group of *BRCA1*-mutation carriers and for the second group of *BRCA1*-mutation carriers in two separate batches, each containing controls. Microarray datasets are publicly available at ArrayExpress database (www.ebi.ac.uk/arrayexpress), accession number E-MTAB-982.

Microarray data analysis

Images of the Human Gene 1.0 ST arrays were quantified with GCOS software (Affymetrix). Thereafter, quality control and RMA normalization of the data was performed using the arrayanalysis.org workflow (www.arrayanalysis.org) for each of the batches separately. All arrays passed quality checks. After this, we have obtained information for 22,732 genes. All statistical analyses presented were performed using the open source program R version 12/13 [18] and publicly available library "limma" [19]. This library was used to fit two statistical models: one to compare the first group of mutation carriers to the controls from their batch, and one to compare the second group to their controls. Average intensities, fold changes and p-values as well as annotations were stored for both comparisons.

Analysis of functional categories

For each of the two comparisons made, genes with a fold-change difference of at least 10% and significantly altered (p-value< 0.05) were classified into categories of biological processes and molecular functions using PathVisio [20] and MetaCore (GeneGo, San Diego, CA, USA). PathVisio analyses for pathway enrichment were performed for all significantly altered genes and for up- and down-regulated genes separately. In PathVisio, the gene database Hs_Derby_20110601.bridge and the pathway collection from WikiPathways [21] were used. PathVisio pathways were ranked by Z-score, which is the standard statistical test under the hypergeometric distribution. Only pathways with a Z-score above 1.96, which corresponds to p-values of 0.05 or lower, were selected as enriched pathways. We removed the pathways for which less than 5 or more than 150 genes were found, as those were considered either too specific or too general for our analyses. MetaCore pathways and networks are ranked by p-value. Only pathways with false discovery rate below 0.2 are shown. Micronuclei induction network was created using with MetaCore, as previously reported [22], and BRCA1 and BRCA1/BARD1 complex were added. Biological processes were evaluated by manually checking the direction of the genes involved and determining the "net effect".

Microarray validation

The transcript levels from a selected set of genes were compared for the above mentioned groups, using Real Time RT-PCR (RT-qPCR) for validation of the microarray data. Total RNA was isolated and purified as described above. First-strand cDNA was obtained from 500ng of total RNA with SuperScript III Reverse Transcriptase kit (Invitrogen) according to the manufacturers' instructions. qPCRs were performed with SensiMix SYBR kit (Quantace), according to the protocol provided, and analysed on the 7900HT system (Applied Biosystems). Ten and 13 genes were selected for analysis of the first and second independent groups, respectively. In addition, *BRCA1*, which was differentially expressed in the first group, was also tested and found to be down-regulated among the second group, despite not reaching significance in the microarrays. Expression of the housekeeping gene *GAPDH* was measured and included during the analysis as reference. The primer sequences are available upon request.

Spearman's rank correlation test was performed to evaluate the results. Our qPCR results (Supplemental Table 6) correlated fairly well with those from the microarrays: R^2=0.791, p-value=0.004 for the first group, and R^2=0.543, p-value=0.045 for the second group. Differences observed are likely to be related with the analysis of different splice isoforms and differences in the dynamics of the microarrays and qPCRs. The arrays used in this study contain probes for each exon of each gene, of which the signals are summarized into an average for the gene, thereby also averaging over all isoforms, whereas the primers designed for qPCR may not target all isoforms or a significant average of these isoforms expression, thereby likely to give other expression levels when compared. This difference is difficult to assess more specifically, since the complete isoform repertoire of most genes is currently unknown.

RESULTS AND DISCUSSION

The first aim was the identification of distinct transcriptional changes in $BRCA1^{+/-}$ lymphocytes compared with $BRCA1^{+/+}$ lymphocytes through pathway/network analysis. To make lymphocytes more dependent on BRCA1 functions, they were stimulated to grow and replicate, while normally these cells are in G0 phase. DNA damage was induced with 10 Gγ of gamma-irradiation and cells were harvested for RNA isolation 2 hours post-irradiation. The irradiation dose was selected based on previous studies [6-8, 23], This resulted in the identification of significant differential expression of 3,280 genes, of which 1,733 were up- and 1,547 were down-regulated in $BRCA1^{+/-}$ lymphocytes compared to controls.

Pathway enrichment analyses

Pathway analysis was performed using different pathway databases since it is known that availability of cell signalling and transcriptional regulatory related content may differ

substantially in different databases [24]. The differentially expressed genes among the first group were found to be over-represented in the pathways/networks listed in Tables 1-3 and Supplemental Tables 1-2.

Enrichment analysis of pathways/networks revealed that the differentially expressed genes affected DNA-damage response, cell cycle, apoptosis, and immune response, as observed in both PathVisio and MetaCore software. Manual assessment of the genes involved, their functions, and the direction of their changes (i.e. up or down) allowed evaluation of the "net effect" on these pathways. We have also investigated other known phenotypic characteristic, i.e. micronuclei induction, known to be associated with *BRCA1*-mutation status, which were not available for assessment using the predefined pathway/network enrichment analysis. A summary of the results is represented in Figure 1.

Table 1. Pathway enrichment results of significantly different expressed genes (p-value <0.05) as determined with PathVisio.

Up- and Down-regulated genes (|FC| ≥ 1.1)					
Pathway	**Positive**	**Measured**	**Total**	**%**	**Z Score**
TGF-beta Receptor Signaling Pathway	42	145	152	28.97%	3.52
IL-2 Signaling Pathway	22	73	76	30.14%	2.74
Cell cycle	25	86	94	29.07%	2.72
TCA Cycle	11	30	45	36.67%	2.68
Keap1-Nrf2	6	13	17	46.15%	2.65
Proteasome Degradation	18	60	66	30.00%	2.45
Cytoplasmic Ribosomal Proteins	22	78	88	28.21%	2.38
miRNAs involved in DDR	11	33	70	33.33%	2.31
IL-3 Signaling Pathway	26	98	102	26.53%	2.24
DNA damage response (only ATM dependent)	23	85	97	27.06%	2.21
G1 to S cell cycle control	19	68	71	27.94%	2.17
DNA damage response	18	65	71	27.69%	2.06
Only up-regulated genes (FC ≥ 1.1)					
Pathway	**positive**	**Measured**	**total**	**%**	**Z Score**
Cytoplasmic Ribosomal Proteins	20	78	88	25.64%	6.08
DNA damage response (only ATM dependent)	15	85	97	17.65%	3.54
Mitochondrial Gene Expression	5	17	23	29.41%	3.4
G Protein Signaling Pathways	14	86	96	16.28%	3.07
TGF-beta Receptor Signaling Pathway	19	145	152	13.10%	2.55
IL-2 Signaling Pathway	11	73	76	15.07%	2.43
Myometrial Relaxation and Contraction Pathways	18	149	161	12.08%	2.11
T Cell Receptor Signaling Pathway	16	131	135	12.21%	2.03
Calcium Regulation in the Cardiac Cell	17	142	153	11.97%	2.01
Only down-regulated genes (FC ≤ -1.1)					

Pathway	positive	Measured	total	%	Z Score
Cell cycle	21	86	94	24.42%	4.33
Proteasome Degradation	14	60	66	23.33%	3.33
DNA Replication	10	41	49	24.39%	2.97
G1 to S cell cycle control	14	68	71	20.59%	2.79
Toll-like receptor signaling pathway	18	97	108	18.56%	2.69
IL-3 Signaling Pathway	18	98	102	18.37%	2.64
miRNAs involved in DDR	8	33	70	24.24%	2.63
DNA damage response	13	65	71	20.00%	2.57
IL-5 Signaling Pathway	13	68	69	19.12%	2.39
One Carbon Metabolism	6	24	39	25.00%	2.36
TCA Cycle	7	30	45	23.33%	2.34
TGF-beta Receptor Signaling Pathway	23	145	152	15.86%	2.22
Senescence and Autophagy	16	94	102	17.02%	2.15
Fluoropyrimidine Activity	7	32	37	21.88%	2.15

Positive indicates the number of genes meeting the criteria; measured indicates the number of genes measured in the pathway; total indicates the number of genes in the respective pathway

Table 2 – Significantly changed genes (p-value <0.05) are overrepresented in the following pathways as determined with MetaCore.

Up- and Down-regulated genes (|FC| ≥ 1.1)			
Maps	**P-value**	**Ratio**	
Development_PIP3 signaling in cardiac myocytes	4.395E-06	21	43
Signal transduction_AKT signaling	1.322E-05	19	39
DNA damage_ATM/ATR regulation of G1/S checkpoint	4.290E-05	16	32
Apoptosis and survival_BAD phosphorylation	6.288E-05	17	36
G-protein signaling_G-Protein alpha-q signaling cascades	1.365E-04	13	25
Apoptosis and survival_HTR1A signaling	1.458E-04	17	38
Immune response_Fc epsilon RI pathway	2.191E-04	19	46
Development_Mu-type opioid receptor signaling	2.279E-04	13	26
Regulation of lipid metabolism_Insulin regulation of glycogen metabolism	3.116E-04	17	40
Immune response_PIP3 signaling in B lymphocytes	3.288E-04	14	30

The top 10 most significant pathways are shown. Supplemental Table 1 contains the complete list.
[a]number of genes affected and number of measured genes in the pathway.

Table 3 – Significantly changed genes (p-value <0.05) are overrepresented in the following networks as determined with MetaCore.

Up- and Down-regulated genes (|FC| ≥ 1.1)			
Networks	**P-value**	**Ratio**[a]	
Cell adhesion_Leucocyte chemotaxis	3.789E-07	61	190
Cell cycle_Mitosis	8.209E-07	57	177
Proteolysis_Ubiquitin-proteasomal proteolysis	2.679E-06	53	166
Cell cycle_Meiosis	9.180E-06	36	102
Immune response_Phagosome in antigen presentation	1.114E-05	65	226
Inflammation_MIF signaling	1.473E-05	39	116
DNA damage_DBS repair	1.505E-05	37	108
Cell cycle_G2-M	4.769E-05	58	204
Cell cycle_S phase	4.910E-05	45	147
Apoptosis_Anti-apoptosis mediated by external signals via NF-kB	1.538E-04	33	102

The top 10 most significant networks are shown. The complete list is in Supplemental Table 2.
[a]number of genes affected and number of measured genes in the network

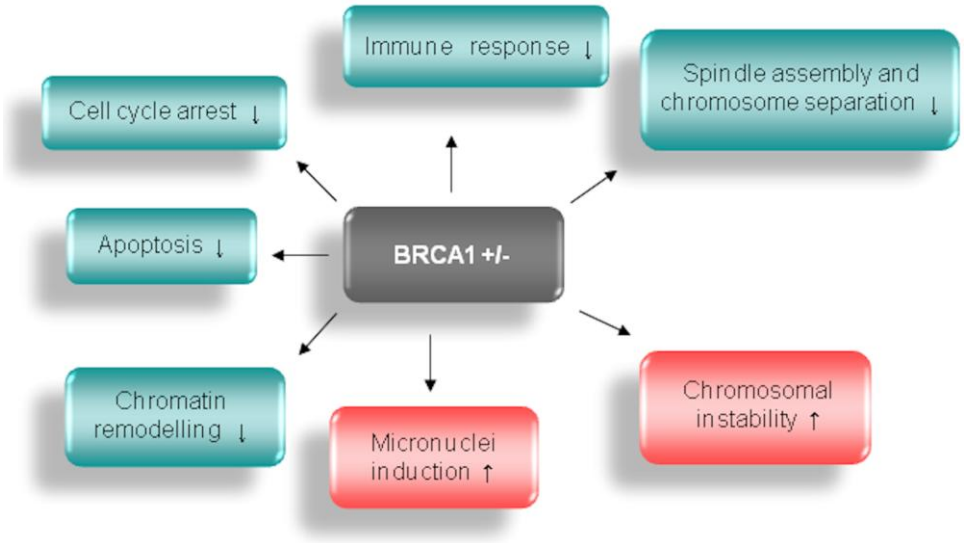


Figure 1. **Pathways predicted to be affected by a *BRCA1* mutation 2h post-irradiation.** Up-oriented arrows indicate processes that are stimulated, whereas down-oriented arrows indicate processes that are inhibited in *BRCA1*-mutated lymphocytes compared with controls.

Cell cycle arrest is impaired in irradiated BRCA1$^{+/-}$ lymphocytes

In response to DNA damage by irradiation, ATM showed to be up-regulated compared to controls. It is known that after γ-irradiation, ATM phosphorylates BRCA1, which in turn will induce expression of members of the KIP family, including p21 (CDKN1A) (Supplemental Figure 1). In our data, many of the ATM downstream proteins, which are normally activated by protein interactions, were found to be down-regulated at the RNA level. Both BRCA1 and p21 (CDKN1A), essential to inhibit cell-cycle progression [25], were down-regulated and MYC, which induces G1 to S phase transition [26], was up-regulated. These events suggest impaired cell cycle arrest. The observed down-regulation of cyclins and CDKs observed is not in agreement with this conclusion, since this can delay cell cycle progression. However, the final outcome is consistent with an impaired cell cycle arrest after irradiation, since chromosomal instability was increased, as discussed below. The observed deficient cell cycle arrest is in agreement with previous studies that reported that BRCA1-mutant HCC1937 breast cancer cell line and fibroblasts from *BRCA1* mutation carriers showed impaired and moderate impaired G1/S cell cycle arrest, respectively, after irradiation [27, 28]. Mouse embryonic fibroblast (MEF) cells from BRCA1$^{\Delta 11/\Delta 11}$ embryos were also shown to have an impaired G2/M cell cycle checkpoint [29].

Decreased apoptotic process

Apoptotic processes related with BAD phosphorylation, HTR1A signalling, and NF-kB were present among the top ten most significantly affected pathways and networks, according to MetaCore (Tables 2 and 3), in addition to other significant apoptosis-related pathways (Supplemental Tables 1 and 2). Supplemental Table 3 lists the differentially expressed genes involved in the apoptotic process. Among the pro-apoptotic genes, 17 genes were down-regulated, whereas 11 were up-regulated. Among the genes that negatively regulate apoptosis, seven were differentially expressed and found to be up-regulated. Deficient cell-cycle checkpoints following irradiation result in genomic instability, which in turn should lead to apoptosis. The IFN-γ mediated apoptosis is an important pathway and is known to depend on functionally intact BRCA1 protein [30]. Although not specifically within the apoptotic processes as present in Supplemental Table 1, IFN-γ signalling was also found to be affected. In total, our results suggest inhibition of the apoptotic process. The observation that thymocytes with BRCA1 haploinsufficiency also showed a decreased apoptotic rate in response to radiation compared with wild-type cells [31], provides support to our observation.

Impaired immune response

Several immune response and inflammation pathways were found to be significantly affected (Tables 1 and 3), including IL2, IL3, IL5 and IL6 signalling, Toll-like receptor signalling and T-cell receptor signalling pathways in PathVisio; Fc epsilon RI (IgE receptor), PIP3 signalling, BCR pathway, CXCR4 and CCR5 signalling in MetaCore.

Among the inflammation and immune response pathways of MetaCore, a total of 117 genes were differentially expressed. Further analysis of this list of genes, using DAVID (Database for Annotation, Visualisation and Integrated Discovery; http://david.abcc.ncifcrf.gov) [32], revealed that they are also involved in ErbB, mTOR, JAK-STAT, MAPK signalling, and apoptotic pathways. Besides inflammation, interleukins are known to be also involved in cell cycle, DNA damage repair and apoptosis in different types of cells [33-37]. For example, NF-κB was recently found to be involved in double-stranded DNA damage repair by homologous recombination [37], besides its known functions in inflammation, cell proliferation and apoptosis. In addition, BRCA1 physically interacts with NF-κB, stimulating the transcription of TNFα and IL1β [38], which are pro-inflammatory cytokines. Although NF-κB expression was not affected in our data, IL1β was strongly down-regulated (FC = -3.2, p-value = 0.01) and TNFα was down-regulated at borderline significance level (FC = -1.3, p-value = 0.05). Other pro-inflammatory cytokines were also down-regulated, i.e. IL8 (FC = -2.82, p-value = 0.008), CCL3 (FC = -1.62, p-value = 0.002), and CCL4 (FC = -1.63, p-value = 0.05), whereas IL2 (FC = 1.14, p-value = 0.05) and IL16 (FC = 1.14, p-value = 0.02) are up-regulated. Anti-inflammatory IL10 is down-regulated (FC = -1.16, p-value = 0.04). Overall, these data suggest impaired immune response in $BRCA1^{+/-}$ lymphocytes compared with controls. Impaired immune response may lead to carcinogenesis, due to deficient removal of abnormal cells [39].

Chromosomal instability is increased

It was previously reported that BRCA1-deficient cells demonstrate increased chromosomal instability compared to controls, including short-term cultures of $BRCA1^{+/-}$ lymphocytes after irradiation [23]. The role of BRCA1 in this process is not limited to DNA-damage repair and cell-cycle control functions. BRCA1 is also responsible for chromosomal stability control through BRCA1/BARD1 E3 ubiquitin-ligase activity, which is required for mitotic spindle-pole assembly and regulation of chromatin dynamics [40, 41]. Moreover, it was reported that BRCA1 associates with the centrosome during mitosis and that its hypophosphorylated form binds to γ-tubulin, which is responsible for microtubule nucleation and mitotic spindle formation [42]. Deregulation of the mitotic spindle assembly by siRNA knock-down of BRCA1/BARD1 resulted in micronuclei induction in HeLa cells [40]. This indicates that the influence of BRCA1 in the micronuclei induction is related to its role in spindle checkpoint. Rothfuss and colleagues showed that induced micronuclei frequency was a useful screening test for carriers of *BRCA1* mutations using peripheral lymphocytes [15]. Later, it was shown that this phenotype was not present in lymphoblastoid cell lines (LCLs), suggesting that the transformation process influences the expression of mutagen sensitivity-related genes [13, 14]. It is also noteworthy that the populations giving rise to LCLs derive from B lymphocytes, whereas PHA stimulates growth of T lymphocytes. A transcriptomics network of micronuclei-related genes recently reported [22], to which we added BRCA1, was used to explore our gene expression data

from irradiated IL2/PHA lymphocytes. Eleven genes out of 27 genes were significantly changed in our dataset (Figure 2), which is significantly more than expected by chance (p-value = 0.01). The differentially expressed genes from this network were BRCA1/BARD1, IL6, DNMT1, BAX, BCL2, CDC20, TP53, CDKN1A (p21), BUB1, and FBXW7 (FBXO30) and were still found to form a network. These genes have relevant roles in DNA damage, cell cycle, apoptosis and spindle assembly checkpoint. Interestingly, in Brca1-deficient MEF cells from mice lower expression levels of Bub1 was also observed [43]. It is also noteworthy that DNMT1 variants with a putative pathogenic effect were found in *BRCA1/2*-negative patients with a family history of breast cancer [44].

Decreased expression of CDKN1A (p21) (FC= -1.2) was observed in $BRCA1^{+/-}$ cells. Its decreased expression is consistent with decreased *BRCA1* expression. These results contradict previous findings showing that up-regulation of CDKN1A significantly correlated with MN frequency in $BRCA1^{+/-}$ LCLs [45]. Increased expression of BAX was

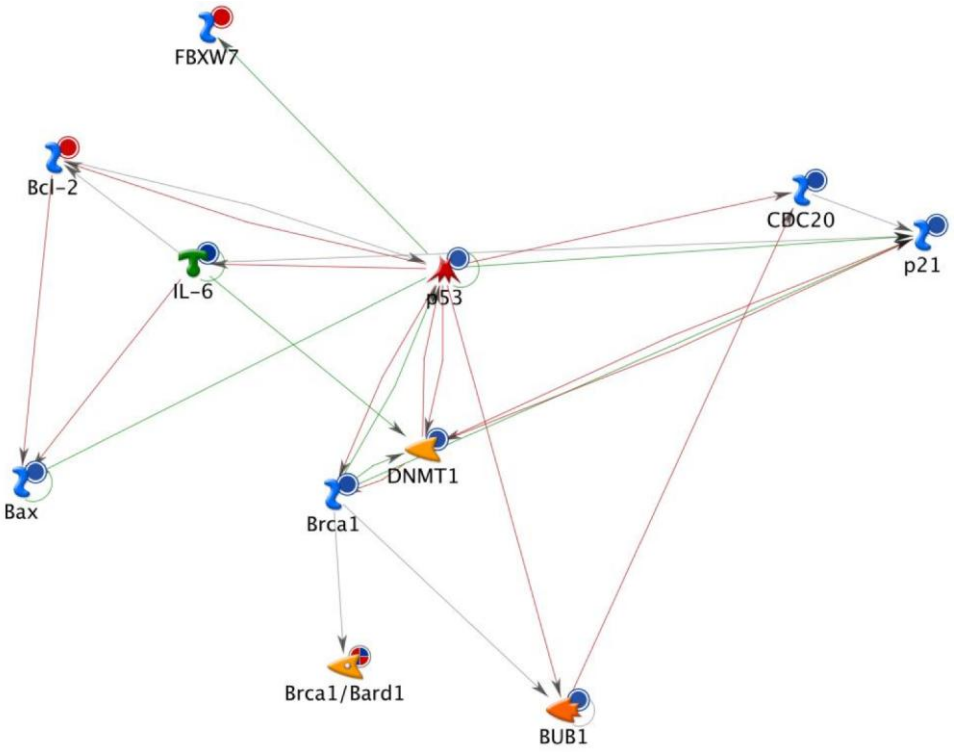


Figure 2. Network of micronucleus (MN) induction. Significantly changed genes from the MN induction network previously reported [22], to which we added BRCA1 and BRCA1/BARD1 complex. Blue and red circles indicate down- and up-regulated genes, respectively. The complete legend is available at http://www.genego.com/pdf/MC_legend.pdf.

also previously showed to be correlated with MN induction frequency in LCLs [45], whereas in our study this tumour suppressor gene was found to be down-regulated. These contrasting results may be due to the fact that, as explained above, LCLs may not be a good cell type to evaluate MN frequency induction in *BRCA1*-mutation carriers.

Chromatin remodelling is down-regulated

During the chromatin remodelling process, BRCA1 interacts with SWI/SNF-related proteins and histone deacetylases [46]. Interestingly, among a group of 26 transcripts involved in this process, 8 were changed, of which 6 were down-regulated. These included SMARCA4, SMARCB1, SMARCAL1, SMARCC2, HDAC8 and HDAC9, whereas the up-regulated genes were SMARCD3 and HDAC5. The number of genes affected is significantly higher than expected by chance (p-value = 0.02) and indicates that BRCA1 not only interacts with these proteins, but likely it also influences their expression. The overall down-regulation of these genes may lead to decreased chromatin remodelling, which affects transcription in general.

Gene-set obtained from affected biological processes

Genes from the relevant biological pathways related to *BRCA1*-mutated cells described above were selected to be part of a genetic classifier. These genes are listed in Table 4.

Independent analysis of a second group of mutation carriers: group 2

To extend our genetic classifier and to validate our findings, microarray analysis of a second independent group of *BRCA1*-mutation carriers was performed. This group included 15 non-related females with three different frequently occurring *BRCA1* mutations from our patient population (Supplemental Table 4). These samples were irradiated and analyzed together with five additional *BRCA1*-mutation negative controls. We further refer to these samples as group 2 and the initial samples described above as group 1.

Analysis of the second group revealed 1,520 significantly changed genes. Of these, 682 and 593 were at least 10% up- and down-regulated, respectively. The Venn diagrams depicted in Figure 3 show the number of genes found in common with group 1. The differentially expressed genes shared between both groups are shown in Supplemental Table 5. Among these genes, several are known to interact with BRCA1 or the BRCA1/BARD1 complex, either directly (PLK1, AURKB, MED21, KDM5B) or indirectly (AHCYL2). Others play a role in the cell cycle (CDC20, CDC20P1, INTS6), TGF-beta signalling (TGFBR1, SMAD2), apoptosis (IL6, C22orf28, BCL2), or DNA damage repair (PARP1, CHRAC1, RFC2, GTF2IRD1). CENPA is also noteworthy as it interacts with the previously mentioned PARP1, AURKB, and CDC20 (DNA damage repair, cell cycle). In addition, GORASP1 and MMP9 are also of potential interest, since they interact with PLK1 and SMAD2 (cell cycle, DNA damage repair, and TGF-beta signalling), respectively. Among the genes from the MN induction network, BCL2 is up-regulated while IL6, CDC20, and BRCA1 are down-regulated, as also observed for group 1. Genes included in this list are

likely relevant in a gene signature to be used as a classifier for *BRCA1*-mutation status, based on the assumption that sets of differentially expressed genes observed in two independent groups of *BRCA1*$^{+/-}$ samples are biologically related to BRCA1 haploinsufficiency.

Table 4 . Genes for the genetic classifier retrieved from the affected biological pathways.

Gene symbol	**Ensembl ID**
BRCA1	ENSG00000012048
CDKN1A	ENSG00000124762
CDKN1B	ENSG00000111276
CCND1	ENSG00000110092
CCND2	ENSG00000118971
CDK4	ENSG00000135446
IL6	ENSG00000136244
DNMT1	ENSG00000130816
BAX	ENSG00000087088
BCL2	ENSG00000171791
CDC20	ENSG00000117399
TP53	ENSG00000141510
BUB1	ENSG00000169679
FBXO30	ENSG00000118496
PLK1	ENSG00000166851
TGFBR1	ENSG00000106799
SMAD2	ENSG00000175387
BAG1	ENSG00000107262
OPA1	ENSG00000198836
XIAP	ENSG00000101966
PIK3R2	ENSG00000105647
MSH6	ENSG00000116062
MYC	ENSG00000136997
IL1R2	ENSG00000115590
TNF	ENSG00000232810

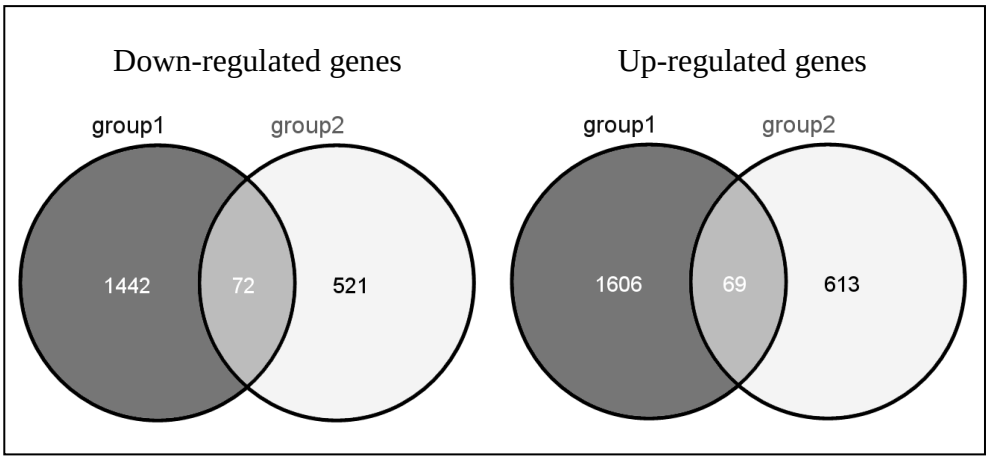


Figure 3. Venn diagrams showing the overlap of the differentially expressed genes between the two groups analysed.

Hierarchical clustering experiments

The shared genes between group1 and 2, listed in Supplemental Table 5 (141 genes), merged with the most relevant differentially expressed genes from the affected pathways discussed above (25 genes) were used in hierarchical cluster experiments. In total, the gene-set is composed of BRCA1-deficiency associated genes. Clustering of the samples was performed using BRB-ArrayTools, which was developed by Dr. Richard Simon and BRB-ArrayTools Development Team (http://linus.nci.nih.gov/BRB-ArrayTools.html). We have also clustered the samples in group 2 using the class prediction gene set from group 1 as identified by using the class prediction function available in the BRB-ArrayTools, and vice-versa (Supplemental Figure 2). Using the BRCA1-deficiency associated genes, the number of samples correctly clustered, the robustness- and discrepancy-ratios improved as compared to results obtained using the class prediction gene-sets. However, even with our gene set, a few samples from mutation carriers clustered together with control samples (Figure 4). Upon investigating this further, we found no clear evidence that these samples were different from the other probands, i.e. regarding mutation localization or phenotype. It is noteworthy that in both groups, we observed 20% false negatives, but no false positives.

In addition to testing our groups of samples, we have retrieved datasets from previous studies and subjected those to hierarchical clustering with our set of genes to verify its robustness in independent data [7-9] (Figure 5). Performance on the study of Kote-Jarai et al. [7] could not be evaluated as most of the genes included in the classifier were not present on the microarray used by the authors. The remaining set (71 genes) was too limited to correctly cluster the samples with and without *BRCA1* mutation (data not shown).

The study of Waddell et al. includes samples with *BRCA1*-truncating mutations, *BRCA1* pathogenic missense mutations, and *BRCA1/2*-negative patients who were screened due to a family history of breast cancer (*BRCA*X). Using 89 genes from our set that were present in this study, the clustering resulted in two major groups, each of them containing predominantly either *BRCA*X samples or *BRCA1*-mutations (Figure 5A). Since, *BRCA*X samples are from patients from high-risk families for whom no *BRCA1* or *BRCA2* mutation was identified, it is still possible that these samples harbour mutations in non-tested regions of *BRCA1* as for example the promoter region, deep intronic, or in miRNAs involved in the regulation of *BRCA1* expression and their binding sites. In this case, part of the few *BRCA*X misclustered samples may in fact be in the correct group. Interestingly, using our approach, samples with missense mutations cluster together with most of the truncating mutations and remained separate from the majority of the *BRCA*X samples, whereas Waddell and colleagues described two distinct classifiers for the two types of mutation.

In the microarray used by Walker et al., 126 genes from our initial gene set could be used for cluster experiments. Using this dataset we were able to cluster separately *BRCA1* mutation carriers (Figure 5B) and healthy controls. This achievement was successful both for IR- and mitomycin C-treated samples.

Results from the unsupervised clustering experiments indicate that the approach used in this study is promising and likely to be more useful to identify a reproducible genetic classifier than using purely agnostic approaches that use the most differentially expressed genes from a dataset

In this study, the transcriptome associated with BRCA1 haploinsufficiency was characterized, by irradiating lymphocytes from *BRCA1*-mutation carriers and controls. Here we confirm for the first time that the transcriptome of normal cells with heterozygous *BRCA1* mutations, when subjected to irradiation, shows alterations in many of the functions/phenotypes which were previously reported to be associated with BRCA1-deficiency. These include deficient cell-cycle arrest, decreased apoptosis, decreased immune response processes, decreased chromatin remodelling, and increased chromosomal instability leading to increased MN induction.

Evaluation of the affected pathways lead to the identification of biologically relevant genes linked to BRCA1 haploinsufficiency (Table 4). These genes were included in a genetic signature that has the potential to distinguish pathogenic *BRCA1* sequence variants from neutral ones. Additional genes for the signature were obtained from the overlap in differentially expressed genes from a second group of *BRCA1*-mutation carriers. In total, our gene-set contains 160 BRCA1-defficiency associated genes. This number may seem high compared to the number of samples used, but we would like to stress that our approach is not based on the most significantly different expressed genes, since this could indeed lead to overfitting on a specific dataset. This is not the case in this study, as we demonstrate its potential to cluster samples from independent datasets correctly. Its performance can most probably still be improved further by analysing additional samples,

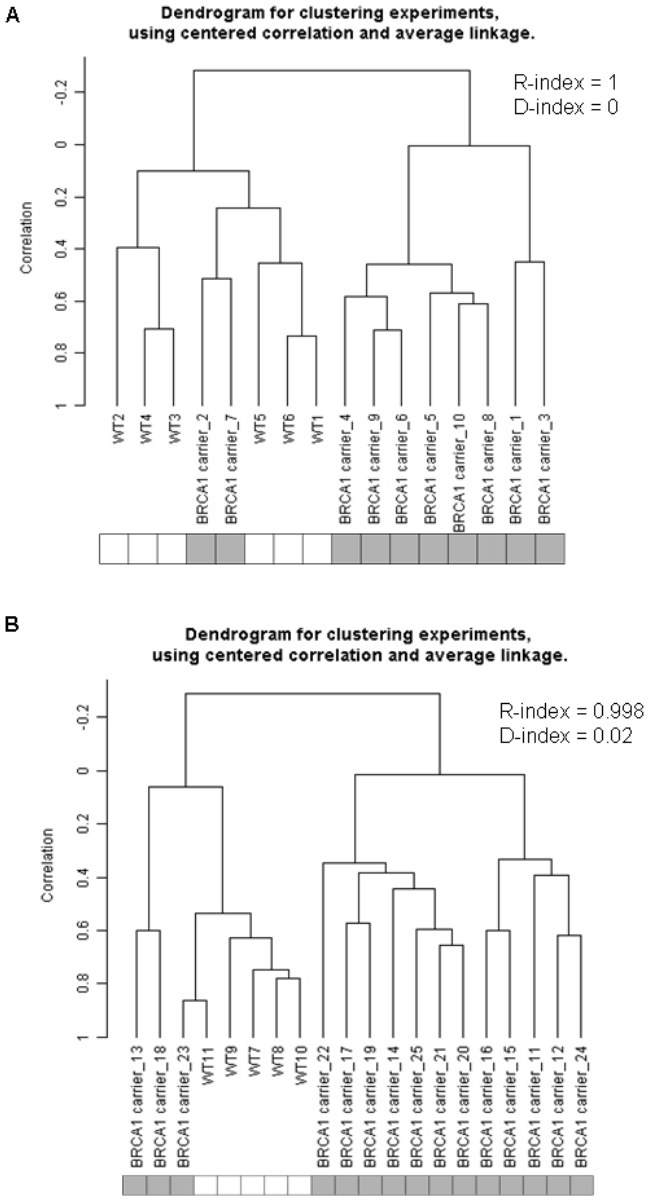


Figure 4. Hierarchical clustering experiments of the two groups of samples analysed in this study. Dendrograms of group 1 (A) and group 2 (B) of *BRCA1*-mutation carriers. Hierarchical cluster analysis of the groups was performed using the genes found to be differentially expressed on both groups (Supplemental Table 5) and the differently expressed genes from the main affected pathways (Table 4). Robustness and discrepancy indexes (R-index and D-index, respectively) are shown. These were calculated based on 100 permutations of the data and considering 2 clusters for group 1 and 4 clusters for group 2. Grey squares represent *BRCA1*-mutation carriers and white squares represent controls.

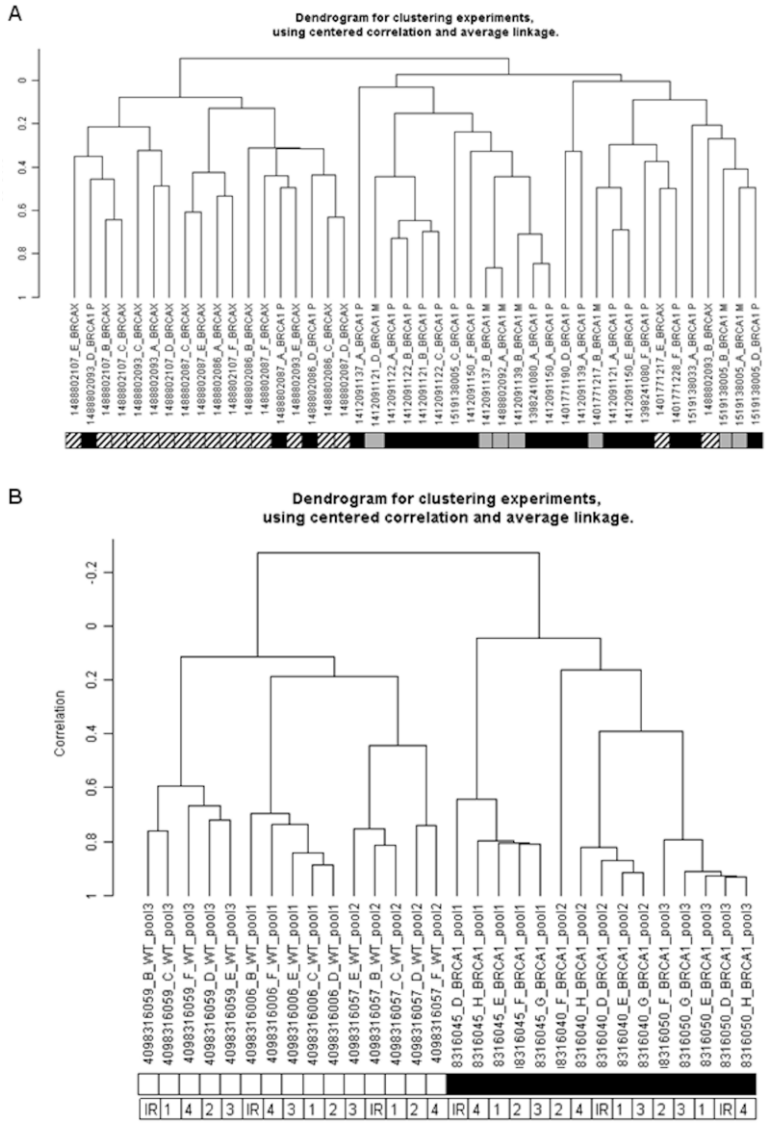


Figure 5. Cluster experiments of samples from independent studies. Dendrograms of the datasets from Waddell et al. (A) and Walker et al. (B). Hierarchical cluster analysis of the groups was performed using the genes found to be differentially expressed on both groups (Supplemental Table 5) and the differently expressed genes from the main affected pathways (Table 4) that were present in these datasets. Robustness and discrepancy indexes (R-index and D-index, respectively) are shown. These were calculated based on 100 permutations of the data and 2-cluster analysis. Black squares represent truncating *BRCA1*-mutation samples, grey squares represent missense pathogenic *BRCA1*-mutation samples, white squares represent healthy controls, striped squares represent *BRCA*X samples (samples from high risk-families without *BRCA1/2* mutations). In panel B, besides the mutation status, The DNA damage source is also indicated. IR stands for ionizing irradiation, 1 to 4 represent the four conditions of mitomycin C (MMC) treatments: 1- 0.4μM t=1h; 2-0.4μM t=2h; 3-1.2μM t=1h; 4- 1.2 μM t=2h.

which would allow to define better the list of differentially expressed genes related to the *BRCA1*$^{+/-}$ status in response to DNA damage. Additionally, it would be of interest to include pathogenic missense mutations in future experiments to ensure that the classifier can indeed be used to test this type of variants, although it already performed well on those from the Waddell *et al*. study.

We foresee that the results from these classification tests based on gene expression are eventually included in a multiparametric approach, which combine other functional tests, biochemical properties and conservation of the amino acids involved with personal and family clinical data, to determine the clinical relevance of UVs [47, 48]. This will contribute to the improvement of the cancer risk assessment for thousands of families carrying these variants.

ACKNOWLEDGEMENTS

We thank Carolien van Oosterhoud, Kim Spee and Roel Brandts for mutation screening of the control samples used in the study, Monique Verluyten for confirming the presence of the mutations in the samples used, and Sabina J. Van Herle for determining the RNA integrity and preparing the samples for the microarray scanning.

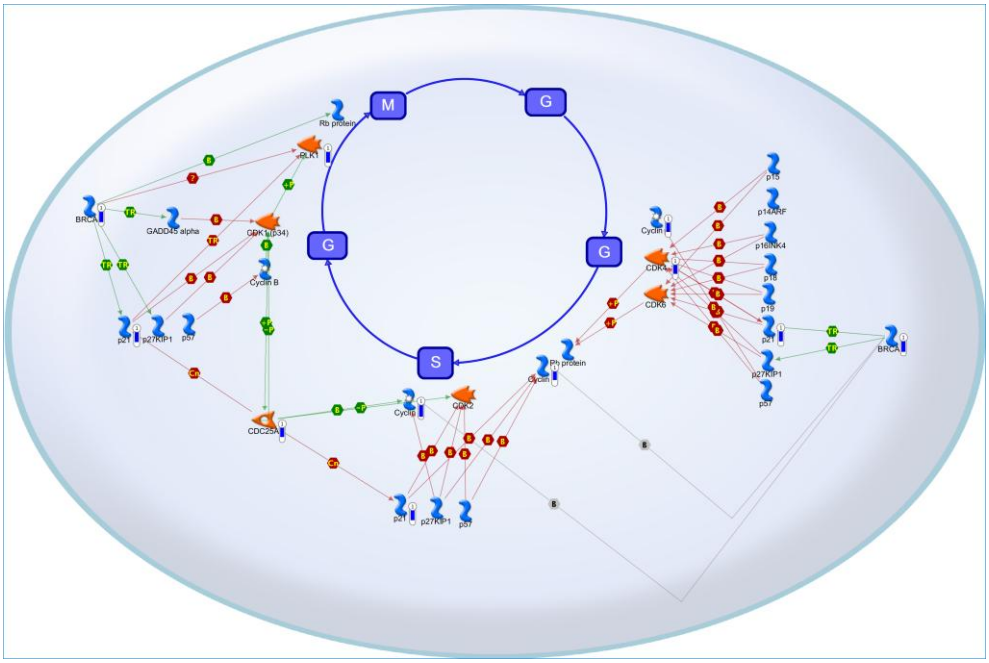
REFERENCES

1. Hall, J.M., M.K. Lee, and J. Morrow, *Linkage analysis of early onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**: p. 1684 - 1689.
2. Miki, Y., J. Swensen, D. Shattuck-Eidens, et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
3. Wooster, R., G. Bignell, J. Lancaster, et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-792.
4. Wooster, R., S. Neuhausen, J. Mangion, et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-2090.
5. Radice, P., S. De Summa, L. Caleca, and S. Tommasi, *Unclassified variants in BRCA genes: guidelines for interpretation.* Ann Oncol, 2011. **22**(suppl 1): p. i18-i23.
6. Kote-Jarai, Z., R.D. Williams, N. Cattini, et al., *Gene Expression Profiling after Radiation-Induced DNA Damage Is Strongly Predictive of BRCA1 Mutation Carrier Status.* Clin Cancer Res, 2004. **10**(3): p. 958-963.
7. Kote-Jarai, Z., L. Matthews, A. Osorio, et al., *Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage.* Clin Cancer Res, 2006. **12**(13): p. 3896-3901.
8. Waddell, N., A. Ten Haaf, A. Marsh, et al., *BRCA1 and BRCA2 missense variants of high and low clinical significance influence lymphoblastoid cell line post-irradiation gene expression.* PLoS Genet, 2008. **4**(5): p. e1000080.
9. Walker, L.C., B.A. Thompson, N. Waddell, et al., *Use of DNA-damaging agents and RNA pooling to assess expression profiles associated with BRCA1 and BRCA2 mutation status in familial breast cancer patients.* PLoS Genet, 2010. **6**(2): p. e1000850.
10. Dobbin, K.K. and R.M. Simon, *Sample size planning for developing classifiers using high-dimensional DNA microarray data.* Biostatistics, 2007. **8**(1): p. 101-117.
11. Dupuy, A. and R.M. Simon, *Critical Review of Published Microarray Studies for Cancer Outcome and Guidelines on Statistical Analysis and Reporting.* Journal of the National Cancer Institute, 2007. **99**(2): p. 147-157.
12. Sotiriou, C. and M.J. Piccart, *Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care?* Nat Rev Cancer, 2007. **7**(7): p. 545-553.
13. Baeyens, A., H. Thierens, K. Vandenbulcke, L. De Ridder, and A. Vral, *The use of EBV-transformed cell lines of breast cancer patients to measure chromosomal radiosensitivity.* Mutagenesis, 2004. **19**(4): p. 285-290.
14. Trenz, K., J. Landgraf, and G.n. Speit, *Mutagen sensitivity of human lymphoblastoid cells with a BRCA1 mutation.* Breast Cancer ResTreat, 2003. **78**(1): p. 69-79.
15. Rothfuβ, A., P. Schutz, S. Bochum, et al., *Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families.* Cancer Res, 2000. **60**(2): p. 390-394.
16. Brandão, R.D., K. van Roozendaal, D. Tserpelis, E.G. García, and M.J. Blok, *Characterisation of unclassified variants in the BRCA1/2 genes with a putative effect on splicing.* Breast Cancer Res Treat, 2011: p. 1-12, doi: 10.1007/s10549-011-1599-7
17. Dai, M., P. Wang, A.D. Boyd, et al., *Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data.* Nucleic Acids Res, 2005. **33**(20): p. e175.
18. Ihaka, R. and R. Gentleman, *R: a language for data analysis and graphics.* Journal of Computational and Graphical Statistics, 1996. **5**(3): p. 299-314.
19. Smyth, G.K., *Limma: linear models for microarray data*, in *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, et al., Editors. 2005, Springer: New York. p. 397-420.
20. van Iersel, M., T. Kelder, A. Pico, et al., *Presenting and exploring biological pathways with PathVisio.* BMC Bioinformatics, 2008. **9**(1): p. 399.

21. Pico, A.R., T. Kelder, M.P. van Iersel, et al., *WikiPathways: Pathway Editing for the People.* PLoS Biol, 2008. **6**(7): p. e184.
22. van Leeuwen, D.M., M. Pedersen, L.E. Knudsen, et al., *Transcriptomic network analysis of micronuclei-related genes: a case study.* Mutagenesis, 2011. **26**(1): p. 27-32.
23. Kote-Jarai, Z., *Increased level of chromosomal damage after irradiation of lymphocytes from BRCA1 mutation carriers.* British Journal of Cancer, 2006. **94**: p. 308-310.
24. Shmelkov, E., Z. Tang, I. Aifantis, and A. Statnikov, *Assessing quality and completeness of human transcriptional regulatory pathways on a genome-wide scale.* Biol Direct, 2011. **6**(1): p. 15.
25. Somasundaram, K., H. Zhang, Y.-X. Zeng, et al., *Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21WAF1/CiPl.* Nature, 1997. **389**(6647): p. 187-190.
26. Obaya, A.J., M.K. Mateyak, and J.M. Sedivy, *Mysterious liaisons: the relationship between c-Myc and the cell cycle.* Oncogene, 1999. **18**(19): p. 2934-2941.
27. Xu, B., S.-t. Kim, and M.B. Kastan, *Involvement of Brca1 in S-Phase and G2-Phase Checkpoints after Ionizing Irradiation.* Mol Cell Biol, 2001. **21**(10): p. 3445-3450.
28. Shorrocks, J., S.E. Tobi, H. Latham, et al., *Primary fibroblasts from BRCA1 heterozygotes display an abnormal G1/S cell cycle checkpoint following UVA irradiation but show normal levels of micronuclei following oxidative stress or mitomycin C treatment.* Int J Radiat Oncol Biol Phys, 2004. **58**(2): p. 470-478.
29. Xu, X., *Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells.* Mol Cell, 1999. **3**: p. 389-395.
30. Andrews, H.N., P.B. Mullan, S. McWilliams, et al., *BRCA1 Regulates the Interferon γ-mediated Apoptotic Response.* J. Biol Chem, 2002. **277**(29): p. 26225-26232.
31. Su, F., L. Smilenov, T. Ludwig, et al., *Hemizygosity for Atm and Brca1 influence the balance between cell transformation and apoptosis.* Radiat Oncol, 2010. **5**(1): p. 15.
32. Dennis, G., B. Sherman, D. Hosack, et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery.* Genome Biology, 2003. **4**(5): p. 3.
33. Lanzavecchia, A., G. Iezzi, and A. Viola, *From TCR engagement to T Cell activation: a kinetic view of T cell behavior.* Cell, 1999. **96**(1): p. 1-4.
34. Khaled, A.R., D.V. Bulavin, C. Kittipatarin, et al., *Cytokine-driven cell cycling is mediated through Cdc25A.* J Cell Biol, 2005. **169**(5): p. 755-763.
35. Moodley, Y.P., N.L.A. Misso, A.K. Scaffidi, et al., *Inverse effects of interleukin-6 on apoptosis of fibroblasts from pulmonary fibrosis and normal lungs.* Am J Respir Cell Mol Biol, 2003. **29**(4): p. 490-498.
36. Boyd, Z.S., A. Kriatchko, J. Yang, et al., *Interleukin-10 receptor signaling through STAT-3 regulates the apoptosis of retinal ganglion cells in response to stress.* Invest Ophthalmol Vis Sci, 2003. **44**(12): p. 5206-5211.
37. Volcic, M., S. Karl, B. Baumann, et al., *NF-kB regulates DNA double-strand break repair in conjunction with BRCA1-CtIP complexes.* Nucleic Acids Res, 2011, doi: 10.1093/nar/gkr687
38. Benezra, M., N. Chevallier, D.J. Morrison, et al., *BRCA1 augments transcription by the NF-κB transcription factor by binding to the Rel domain of the p65/RelA subunit.* J Biol Chem, 2003. **278**(29): p. 26333-26341.
39. Steer, H.J., R.A. Lake, A.K. Nowak, and B.W.S. Robinson, *Harnessing the immune response to treat cancer.* Oncogene, 2010. **29**(48): p. 6301-6313.
40. Joukov, V., A.C. Groen, T. Prokhorova, et al., *The BRCA1/BARD1 heterodimer modulates Ran-dependent mitotic spindle assembly.* Cell, 2006. **127**(3): p. 539-552.
41. Boulton, S.J., *Cellular functions of the BRCA tumour-suppressor proteins.* Biochem Soc Trans, 2006. **34**: p. 633 - 645.
42. Hsu, L.-C. and R.L. White, *BRCA1 is associated with the centrosome during mitosis.* Proc Natl Acad Sci USA, 1998. **95**: p. 12983–12988.
43. Wang, R.-H., H. Yu, and C.-X. Deng, *A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint.* Proc Natl Acad Sci USA, 2004. **101**(49): p. 17108-17113.
44. Caligo, M.A., C. Pepe, V. Pollacchi, et al., *DNMT1 germline allelic variants in early onset familial breast and breast/ovarian cancer patients negative for BRCA1/2 gene mutations.* AACR Meeting Abstracts, 2006. **2006**(1): p. 377.

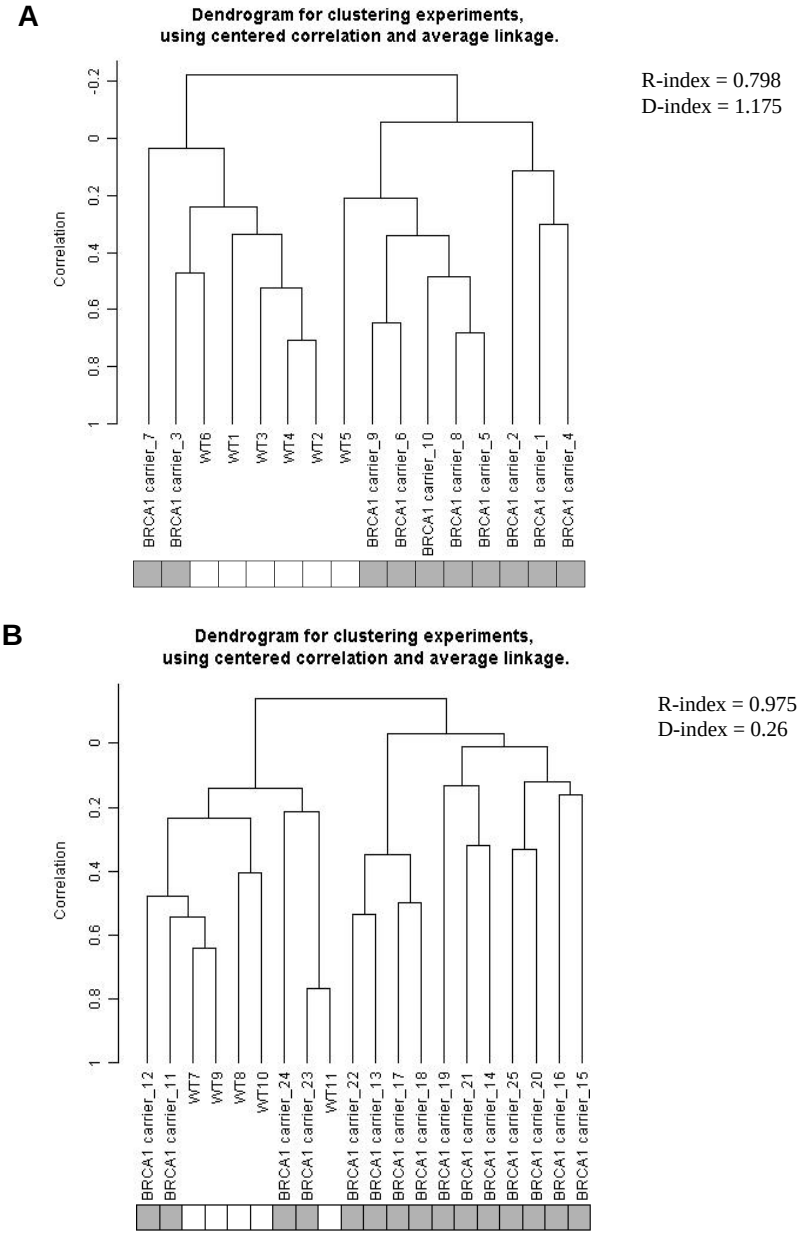
45. Bishay, K., K. Ory, J. Lebeau, et al., *DNA damage-related gene expression as biomarkers to assess cellular response after gamma irradiation of a human lymphoblastoid cell line.* Oncogene, 2000. **19**(7): p. 916-923.
46. Bochar, D.A., L. Wang, H. Beniya, et al., *BRCA1 Is Associated with a Human SWI/SNF-Related Complex: Linking Chromatin Remodeling to Breast Cancer.* Cell, 2000. **102**(2): p. 257-265.
47. Goldgar, D.E., D.F. Easton, A.M. Deffenbaugh, et al., *Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2.* Am J Hum Genet, 2004. **75**(4): p. 535-544.
48. Wu, K., S.R. Hinson, A. Ohashi, et al., *Functional evaluation and cancer risk assessment of BRCA 2 unclassified variants.* Cancer Res, 2005. **65**(2): p. 417-426.

SUPPLEMENTARY DATA



Supplemental Figure 1. Cell cycle regulation - custom pathway.

Cell cycle progress is driven by cyclin-dependent kinases (CDKs) and cyclins. In mammalian cells, Cyclin D, CDK4 and CDK6 are responsible for G1 progression, Cyclin E/CDK2 for the G1/S transition [49], Cyclin A/CDK2 for S phase progression [50], and Cyclin A/CDK1[51] and Cyclin B/CDK1[52] for entry into M phase. The activity of cyclins and CDKs can be inhibited by CKD inhibitors (CDKIs, INK4 and CIP/KIP family members) and reversible phosphorylation. In the presence of DNA damage, the G1/S and G2/M cell cycle checkpoints will delay the cell cycle progression. BRCA1 is phosphorylated by ATM in response to ionizing radiation, and induces G1/S arrest by inducing p21 expression [18] and p27 [53] and interacting with retinoblastoma protein (RB), keeping RB in the hypophosphorylated state [54]. When RB is phosphorylated its action is inhibited and the cell cycle progresses. During the G2/M checkpoint, BRCA1 induces expression of GADD45 [55], in response to ionizing radiation. GADD45 will in turn inhibit the CyclinB-CDC2 complex [56]. BRCA1 also inhibits PLK1 [57], a kinase required for G2 to M transition. Up-ward thermometers have red colour and indicate up-regulated signals and down-ward (blue) ones indicate down-regulated expression levels of the genes compared to controls. The legend is available at http://www.genego.com/pdf/MC_legend.pdf.



Supplemental Figure 2. Hierarchical clustering experiments of the two groups of samples analysed in this study. Dendrogram of group 1 clustered with the 45 genes of class prediction of group 2 (A) and group 2 clustered with the 301 gene composing the class prediction gene set of group1 (B). Class prediction gene list was based on genes p-value threshold of 0.001 and 10% FCs. Robustness and discrepancy indexes (R-index and D-index, respectively) are shown. These were calculated based on 100 permutations of the data and considering 2 clusters for group 1 and 4 clusters for group 2. Grey squares represent BRCA1-mutation carriers and white squares represent controls.

Supplemental Table 1. Pathways from MetaCore with FDR<0.2

Map name	P-value	Ratio	
Development_PIP3 signaling in cardiac myocytes	4.40E-06	21	43
Signal transduction_AKT signaling	1.32E-05	19	39
DNA damage_ATM/ATR regulation of G1/S checkpoint	4.29E-05	16	32
Apoptosis and survival_BAD phosphorylation	6.29E-05	17	36
G-protein signaling_G-Protein alpha-q signaling cascades	1.37E-04	13	25
Apoptosis and survival_HTR1A signaling	1.46E-04	17	38
Immune response_Fc epsilon RI pathway	2.19E-04	19	46
Development_Mu-type opioid receptor signaling	2.28E-04	13	26
Regulation of lipid metabolism_Insulin regulation of glycogen metabolism	3.12E-04	17	40
Immune response_PIP3 signaling in B lymphocytes	3.29E-04	14	30
Role of alpha-6/beta-4 integrins in carcinoma progression	3.68E-04	16	37
Development_VEGF signaling via VEGFR2 - generic cascades	4.56E-04	24	67
Signal transduction_IP3 signaling	5.30E-04	16	38
Immune response_BCR pathway	6.20E-04	17	42
Immune response_CXCR4 signaling via second messenger	6.41E-04	12	25
Cell cycle_Transition and termination of DNA replication	9.95E-04	12	26
Transcription_CREB pathway	1.04E-03	16	40
Immune response_CCR5 signaling in macrophages and T lymphocytes	1.16E-03	17	44
Regulation of lipid metabolism_Insulin signaling:generic cascades	1.42E-03	16	41
Regulation of degradation of deltaF508 CFTR in CF	1.50E-03	12	27
Immune response_Inhibitory action of Lipoxins on pro-inflammatory TNF-alpha signaling	1.53E-03	14	34
Apoptosis and survival_Anti-apoptotic action of Gastrin	1.53E-03	14	34
Development_IGF-1 receptor signaling	1.55E-03	17	45
Cell cycle_Regulation of G1/S transition (part 1)	1.74E-03	15	38
Immune response_IL-15 signaling	1.79E-03	20	57
Development_Regulation of telomere length and cellular immortalization	1.84E-03	13	31
Immune response_MIF - the neuroendocrine-macrophage connector	1.84E-03	13	31
Regulation of degradation of wt-CFTR	2.18E-03	9	18
Immune response_Histamine signaling in dendritic cells	2.36E-03	15	39
Development_Angiotensin activation of Akt	2.60E-03	11	25
Development_A3 receptor signaling	2.90E-03	14	36
Proteolysis_Putative ubiquitin pathway	3.03E-03	10	22
Cell cycle_Role of SCF complex in cell cycle regulation	3.13E-03	12	29
HIV-1 signaling via CCR5 in macrophages and T lymphocytes	3.13E-03	12	29
Immune response_TREM1 signaling pathway	3.46E-03	17	48
NGF activation of NF-kB	3.77E-03	11	26
Development_S1P3 receptor signaling pathway	3.77E-03	11	26
Immune response_ICOS pathway in T-helper cell	3.90E-03	14	37
Translation _Regulation of EIF4F activity	4.42E-03	17	49
Development_Glucocorticoid receptor signaling	4.50E-03	10	23
Chemotaxis_Lipoxin inhibitory action on fMLP-induced neutrophil chemotaxis	4.83E-03	13	34
Signal transduction_cAMP signaling	4.83E-03	13	34

G-protein signaling_Proinsulin C-peptide signaling	5.16E-03	14	38
Cell cycle_Chromosome condensation in prometaphase	5.31E-03	9	20
Apoptosis and survival_Role of CDK5 in neuronal death and survival	5.97E-03	12	31
Translation_Opioid receptors in regulation of translation	6.15E-03	8	17
Translation_IL-2 regulation of translation	6.15E-03	8	17
Transcription_Receptor-mediated HIF regulation	6.43E-03	13	35
Proteolysis_Role of Parkin in the Ubiquitin-Proteasomal Pathway	6.48E-03	10	24
Cell cycle_Nucleocytoplasmic transport of CDK/Cyclins	6.90E-03	7	14
Transcription_Role of heterochromatin protein 1 (HP1) family in transcriptional silencing	7.81E-03	9	21
Development_Activation of astroglial cells proliferation by ACM3	7.81E-03	9	21
Cell cycle_Cell cycle (generic schema)	7.81E-03	9	21
Cytoskeleton remodeling_Thyroliberin in cytoskeleton remodeling	7.81E-03	9	21
Immune response_Fc gamma R-mediated phagocytosis in macrophages	7.99E-03	12	32
Reproduction_Progesterone-mediated oocyte maturation	7.99E-03	12	32
Cell cycle_ESR1 regulation of G1/S transition	7.99E-03	12	32
G-protein signaling_G-Protein beta/gamma signaling cascades	9.06E-03	10	25
Transcription_Transcription regulation of aminoacid metabolism	9.06E-03	10	25
Development_FGF2-dependent induction of EMT	9.32E-03	8	18
Cell cycle_Role of Nek in cell cycle regulation	9.95E-03	11	29
Apoptosis and survival_Beta-2 adrenergic receptor anti-apoptotic action	1.09E-02	7	15
Development_Thyroliberin signaling	1.10E-02	15	45
Development_Activation of ERK by Kappa-type opioid receptor	1.11E-02	9	22
Oxidative stress_Role of ASK1 under oxidative stress	1.11E-02	9	22
DNA damage_ATM / ATR regulation of G2 / M checkpoint	1.24E-02	10	26
Cell adhesion_Chemokines and adhesion	1.35E-02	26	93
Apoptosis and survival_NGF signaling pathway	1.36E-02	8	19
Transport_Aldosterone-mediated regulation of ENaC sodium transport	1.36E-02	8	19
Development_EGFR signaling via PIP3	1.36E-02	8	19
Immune response_PGE2 signaling in immune response	1.36E-02	12	34
Development_VEGF signaling and activation	1.36E-02	12	34
Immune response_IL-13 signaling via PI3K-ERK	1.36E-02	12	34
Immune response_NFAT in immune response	1.38E-02	14	42
Translation_Insulin regulation of translation	1.38E-02	13	38
Immune response_IL-5 signalling	1.38E-02	13	38
Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	1.46E-02	29	107
Development_GH-RH signaling	1.54E-02	9	23
Immune response_IL-4 - antiapoptotic action	1.65E-02	10	27
Cytoskeleton remodeling_FAK signaling	1.68E-02	15	47
Development_GM-CSF signaling	1.68E-02	15	47
Immune response_CD28 signaling	1.71E-02	14	43
Cell cycle_Start of DNA replication in early S phase	1.71E-02	11	31
Development_Ligand-independent activation of ESR1 and ESR2	1.74E-02	13	39
Development_PEDF signaling	1.74E-02	13	39
Development_Endothelin-1/EDNRA transactivation of EGFR	1.74E-02	12	35

Neurophysiological process_NMDA-dependent postsynaptic long-term potentiation in CA1 hippocampal neurons	1.88E-02	17	56
Mechanisms of CFTR activation by S-nitrosoglutathione (normal and CF)	1.91E-02	8	20
Immune response_Lipoxins and Resolvin E1 inhibitory action on neutrophil functions	1.91E-02	8	20
TCA	1.91E-02	8	20
Cytoskeleton remodeling_Role of Activin A in cytoskeleton remodeling	1.91E-02	8	20
Development_Role of HDAC and calcium/calmodulin-dependent kinase (CaMK) in control of skeletal myogenesis	2.04E-02	15	48
Cytoskeleton remodeling_Fibronectin-binding integrins in cell motility	2.16E-02	10	28
Transport_Alpha-2 adrenergic receptor regulation of ion channels	2.16E-02	10	28
Cell cycle_Spindle assembly and chromosome separation	2.19E-02	11	32
Development_EGFR signaling pathway	2.24E-02	17	57
Development_WNT signaling pathway. Part 2	2.35E-02	16	53
Immune response _IFN gamma signaling pathway	2.56E-02	14	45
Cytoskeleton remodeling_ACM3 and ACM4 in keratinocyte migration	2.60E-02	8	21
G-protein signaling_S1P2 receptor signaling	2.60E-02	8	21
Immune response_MIF-mediated glucocorticoid regulation	2.60E-02	8	21
Inhibitory action of Lipoxins and Resolvin E1 on neutrophil functions	2.60E-02	8	21
Immune response_IL-4 signaling pathway	2.65E-02	13	41
Development_Flt3 signaling	2.65E-02	13	41
Development_A2A receptor signaling	2.72E-02	12	37
Transport_ACM3 in salivary glands	2.73E-02	9	25
Development_SSTR2 in regulation of cell proliferation	2.73E-02	9	25
G-protein signaling_G-Protein alpha-12 signaling pathway	2.76E-02	11	33
Transcription_Ligand-dependent activation of the ESR1/SP pathway	2.77E-02	10	29
Immune response_Murine NKG2D signaling	2.77E-02	10	29
Development_TGF-beta receptor signaling	2.94E-02	15	50
Signal transduction_PKA signaling	3.08E-02	14	46
Regulation of lipid metabolism_Insulin regulation of fatty acid methabolism	3.08E-02	14	46
Immune response_Function of MEF2 in T lymphocytes	3.21E-02	13	42
Signal transduction_JNK pathway	3.21E-02	13	42
Signal transduction_PTEN pathway	3.21E-02	13	42
Neurophysiological process_Corticoliberin signaling via CRHR1	3.33E-02	12	38
Transcription_PPAR Pathway	3.33E-02	12	38
Development_A2B receptor: action via G-protein alpha s	3.33E-02	12	38
Muscle contraction_Oxytocin signaling in uterus and mammary gland	3.43E-02	11	34
Neurophysiological process_ACM regulation of nerve impulse	3.43E-02	11	34
G-protein signaling_Regulation of p38 and JNK signaling mediated by G-proteins	3.43E-02	11	34
Apoptosis and survival_nAChR in apoptosis inhibition and cell cycle progression	3.45E-02	8	22
Cell cycle_Role of 14-3-3 proteins in cell cycle regulation	3.45E-02	8	22
Immune response_IL-23 signaling pathway	3.45E-02	8	22
DNA damage_Role of Brca1 and Brca2 in DNA repair	3.50E-02	10	30
Development_EGFR signaling via small GTPases	3.50E-02	10	30
Phospholipid metabolism p.2	3.53E-02	5	11

Development_Role of IL-8 in angiogenesis	3.67E-02	14	47
Immune response_CD16 signaling in NK cells	3.86E-02	16	56
Development_Growth hormone signaling via PI3K/AKT and MAPK cascades	4.04E-02	12	39

Supplemental Table 2. Networks from MetaCore with FDR<0.2

Network name	**P-value**	**Ratio**	
Cell adhesion_Leucocyte chemotaxis	3.79E-07	61	190
Cell cycle_Mitosis	8.21E-07	57	177
Proteolysis_Ubiquitin-proteasomal proteolysis	2.68E-06	53	166
Cell cycle_Meiosis	9.18E-06	36	102
Immune response_Phagosome in antigen presentation	1.11E-05	65	226
Inflammation_MIF signaling	1.47E-05	39	116
DNA damage_DBS repair	1.51E-05	37	108
Cell cycle_G2-M	4.77E-05	58	204
Cell cycle_S phase	4.91E-05	45	147
Apoptosis_Anti-apoptosis mediated by external signals via NF-kB	1.54E-04	33	102
Transcription_Nuclear receptors transcriptional regulation	1.60E-04	52	185
Transcription_Chromatin modification	2.19E-04	38	125
Cell cycle_G1-S Interleukin regulation	2.19E-04	38	125
Translation_Regulation of initiation	3.32E-04	37	123
Cell cycle_G1-S	3.42E-04	46	163
Cell adhesion_Platelet aggregation	4.23E-04	41	142
Reproduction_Progesterone signaling	4.35E-04	52	192
Signal transduction_ESR1-nuclear pathway	5.03E-04	55	207
Immune response_BCR pathway	5.58E-04	37	126
Immune response_Phagocytosis	7.37E-04	52	196
Development_Hemopoiesis, Erythropoietin pathway	8.32E-04	36	124
Cell cycle_G1-S Growth factor regulation	8.72E-04	50	188
Development_Regulation of telomere length	9.26E-04	18	49
Immune response_TCR signaling	1.01E-03	45	166
Inflammation_TREM1 signaling	1.15E-03	36	126
Inflammation_IFN-gamma signaling	1.48E-03	30	101
DNA damage_Checkpoint	1.68E-03	35	124
Reproduction_Feeding and Neurohormone signaling	1.94E-03	52	204
Reproduction_Male sex differentiation	2.46E-03	58	235
Proliferation_Lymphocyte proliferation	2.59E-03	50	197
Apoptosis_Anti-Apoptosis mediated by external signals via MAPK and JAK/STAT	3.52E-03	41	157
Cell adhesion_Integrin priming	3.89E-03	28	98
Cell cycle_Core	5.55E-03	31	114
Cell adhesion_Integrin-mediated cell-matrix adhesion	6.15E-03	51	210
Protein folding_Folding in normal condition	8.31E-03	31	117
Cytoskeleton_Spindle microtubules	8.64E-03	29	108
Transport_Iron transport	9.12E-03	23	81

Immune response_IL-5 signalling	9.19E-03	13	38
Reproduction_FSH-beta signaling pathway	1.04E-02	38	152
Apoptosis_Apoptotic mitochondria	1.06E-02	21	73
Signal transduction_Insulin signaling	1.23E-02	28	106
Muscle contraction_Relaxin signaling	1.23E-02	19	65
Cell adhesion_Cell junctions	1.26E-02	37	149
Inflammation_IL-6 signaling	1.33E-02	30	116
Signal Transduction_Cholecystokinin signaling	1.39E-02	25	93
Neurophysiological process_Long-term potentiation	1.69E-02	17	58
DNA damage_BER-NER repair	1.89E-02	26	100
Transcription_Transcription by RNA polymerase II	1.93E-02	38	158
Development_Regulation of angiogenesis	1.99E-02	48	208
Chemotaxis	2.22E-02	33	135
Apoptosis_Apoptotic nucleus	2.32E-02	37	155
DNA damage_MMR repair	2.55E-02	16	56
Signal transduction_Androgen receptor signaling cross-talk	3.20E-02	17	62
Protein folding_ER and cytoplasm	3.22E-02	13	44
Proteolysis_Proteolysis in cell cycle and apoptosis	3.23E-02	30	124
Translation_Elongation-Termination	3.33E-02	35	149
Translation_Translation initiation	3.34E-02	38	164
Inflammation_IgE signaling	3.39E-02	26	105
Inflammation_IL-4 signaling	3.45E-02	27	110
Neurophysiological process_Corticoliberin signaling	3.46E-02	12	40
Inflammation_IL-2 signaling	3.65E-02	24	96
Apoptosis_Anti-Apoptosis mediated by external signals via PI3K/AKT	3.73E-02	48	216
Signal transduction_ERBB-family signaling	3.88E-02	18	68
Neurophysiological process_Circadian rhythm	4.15E-02	14	50
Autophagy_Autophagy	4.15E-02	14	50
Neurophysiological process_Melatonin signaling	4.39E-02	8	24
Apoptosis_Death Domain receptors & caspases in apoptosis	4.73E-02	29	123
Apoptosis_Anti-Apoptosis mediated by external signals by Estrogen	5.03E-02	18	70
Cytoskeleton_Macropinocytosis and its regulation	5.03E-02	18	70
Transcription_mRNA processing	5.11E-02	36	159
Proliferation_Positive regulation cell proliferation	5.37E-02	47	216
Cytoskeleton_Cytoplasmic microtubules	5.68E-02	27	115
Neurophysiological process_GABAergic neurotransmission	6.02E-02	21	86
Signal transduction_Nitric oxide signaling	6.26E-02	17	67
Signal transduction_NOTCH signaling	6.38E-02	49	229
Protein folding_Protein folding nucleus	6.69E-02	15	58
Signal transduction_Androgen receptor nuclear signaling	7.40E-02	28	123
Cytoskeleton_Regulation of cytoskeleton rearrangement	8.30E-02	39	181
Inflammation_Amphoterin signaling	8.84E-02	26	115
Proliferation_Negative regulation of cell proliferation	9.00E-02	38	177
Inflammation_IL-10 anti-inflammatory response	9.89E-02	19	81
Reproduction_Spermatogenesis, motility and copulation	1.04E-01	45	216

Supplemental Table 3. Apoptosis related genes

	Fold change	P-value
Pro-apoptotic genes		
Via death domain receptors		
DAXX	-1.17	0.005
DEDD2	-1.19	0.011
By DNA damage		
AIFM1	-1.13	0.012
BRCA1	-1.30	0.001
MSH6	1.13	0.018
TP53	-1.15	0.025
By intracellular signals		
CDKN1A	-1.22	0.012
CUL1	-1.15	0.002
CUL5	1.13	0.020
HIPK2	-1.15	0.047
MYC	1.33	$2x10^{-4}$
Other genes		
AKT1	-1.13	0.012
BAX	-1.18	0.006
BCL2L11	1.34	0.012
CASP4	1.15	0.006
CD70	-1.26	0.049
CD27	1.30	0.002
DEDD	-1.14	0.002
IKBKG	-1.10	0.014
NLRP3	1.11	0.004
NUPR1	-1.13	0.035
PIK3R2	-1.24	$2x10^{-4}$
PLAGL2	-1.15	0.011
PMAIP1	1.18	0.009
PPP2R1A	-1.08	0.003
STK17B	1.13	0.004
TNFAIP8	1.17	0.032
ZNF443	1.19	0.040
Anti-apoptotic genes		
BAG1	1.14	0.029
BCL2	1.12	0.023
BNIP3L	1.25	$8x10^{-5}$

MCL1	1.10	0.016
OPA1	1.16	0.019
XIAP	1.13	0.028

Supplemental Table 4. BRCA1 mutations present in the population used in our study

Sample #	BRCA1 mutation
mutation carriers - first group	
BRCA1 carrier_1	c.3549_3550delinsT
BRCA1 carrier_2	c.212+1G>A
BRCA1 carrier_3	c.66dup
BRCA1 carrier_4	c.1504_1508del
BRCA1 carrier_5	c.5277+1G>A
BRCA1 carrier_6	EX1a_7del
BRCA1 carrier_7	c.2269del
BRCA1 carrier_8	c.2197_2201del
BRCA1 carrier_9	c.1115G>A
BRCA1 carrier_10	c.3695del
mutation carriers – second group	
BRCA1 carrier_11	c.2197_2201del
BRCA1 carrier_12	c.2197_2201del
BRCA1 carrier_13	c.2197_2201del
BRCA1 carrier_14	c.2197_2201del
BRCA1 carrier_15	c.2197_2201del
BRCA1 carrier_16	c.2722G>T
BRCA1 carrier_17	c.2722G>T
BRCA1 carrier_18	c.2722G>T
BRCA1 carrier_19	c.2722G>T
BRCA1 carrier_20	c.2722G>T
BRCA1 carrier_21	c.5277+1G>A
BRCA1 carrier_22	c.5277+1G>A
BRCA1 carrier_23	c.5277+1G>A
BRCA1 carrier_24	c.5277+1G>A
BRCA1 carrier_25	c.5277+1G>A

Supplemental Table 5. List of common deregulated genes between group 1 and group 2 of *BRCA1*-mutation carriers.

Down	Up
AC093838.4	AC016683.6
AHCYL2[a]	AC093693.1
AKR1A1	AKAP5
AURKB[a]	ALG6
BPGM	AOAH
C17orf87	AS3MT
C22orf28[d]	ATHL1
CCL3[f]	BCL2[d]
CCR6[f]	BTBD3
CD74[f]	C11orf46
CDC20[b]	C15orf29
CDC42BPB	C5orf42
CENPA	CD55
CHAMP1	CDC37L1
CHI3L1	COMMD6
CHRAC1[e]	DCP2
CLEC6A	DCUN1D4
CNDP2	ESF1
CYBASC3	ETNK1
DENND5B	FAM208B
DUS2L	FGFBP2
DUS3L	G2E3
EHD4	GALNT11
FLOT2	GNAQ
G6PC3	HAVCR1
GLA	HCFC2
GORASP1	INTS6[b]
GTF2IRD1[e]	J01415.16
IDO1	JPX
IL6[d,f]	KCTD9
KIF4A	KDM5B[a]
LIMK1	KIAA2026
LIX1L	KLF12
LONP1	KLRC2
MAP4K2	KLRK1
MMP9	LYST

MYO1C	MAN1A2
N/A (ENSG00000149397)	MAP3K2
NAPSB	MED21[a]
NMRAL1	MGA
NRSN2	MIB1
NUBP1	MIR142
PARP1[e]	MLLT10
PCK2	N/A (ENSG00000249546)
PEX26	PHF14
PFKM	PMCH
PLEK	PMFBP1
PLK1[a]	PRMT10
PRDX1	PURB
PSMB6	RBM41
REEP4	RP11-303G3.6
RFC2[e]	RP11-466F5.9
RFFL	RP4-706A16.3
RNF26	RPS6KA3
RNFT2	SMAD2[c]
RWDD2B	TAOK1
SCAMP3	TAS2R13
SCAMP4	TAS2R14
SECTM1	TGFBR1[c]
SHKBP1	TIMD4
SLCO5A1	TIMM8B
SNRPA	TRIM23
STAP2	U6 (ENSG00000202029)
TMED8	U6 (ENSG00000252444)
TMEM173	URGCP
TMEM176A	XXbac-BPG55C20.1
TNFRSF17	ZNF177
TOMM34	ZNF638
UEVLD	ZNF770
USP6NL	
ZEB2	
ZNF385A	

[a]proteins that interact with BRCA1 or BRCA1/BARD1; proteins involved in [b]cell cycle, [c]TGF-β signalling, [d]apoptosis, [e]DNA damage repair, and [f]inflammation/immune response.

Supplemental Table 6 – Validation of the microarrays by quantitative RT-PCR

gene ID	**Fold Change**	
	microarrays	qPCR results
First group		
MYC	1.3275	1.973
BRCA1	-1.3050	-1.350
IL6	-1.6851	-2.455
CDC20	-1.3405	1.016
IFIT2	1.5338	2.651
SMAD2	1.1795	2.202
CDKN1A	-1.2233	-1.055
TGFBR1	1.2209	2.201
BRCA2	1.1816	1.564
TP53	-1.1468	-1.266
MMP9	-2.1883	-1.104
Second group		
TLR5	-1.6127	-4.0000
IFIT2	-1.6589	-3.0620
FHIT	-1.5435	-2.1080
CCR6	-1.4073	-2.3990
MYC	-1.2373	-2.0280
CDC25C	-1.2547	-1.6600
PTK2	-1.5170	-2.1960
PTGS2	-2.9963	-2.1820
LYZ	-4.5622	-5.7000
IL6	-3.1142	1.0740
TGFBR1	1.1870	-1.0630
MMP9	-3.0087	-2.8750
SMAD2	1.1135	-1.1520
BRCA1	-1.1557	-2.0500

CHAPTER 7

General discussion

Hereditary cases account for approximately 5-15% of all breast cancer (BC) [1-6] and 10-15% of all ovarian cancer (OC) cases [1, 3, 7, 8]. Two major susceptibility genes for breast cancer, *BRCA1* [9] and *BRCA2* [10] were discovered in 1994 and 1995, respectively. These genes account for the largest part of the hereditary breast and ovarian cancers. No other single gene is likely to be found responsible for this syndrome [11, 12]. Inherited highly penetrant pathogenic mutations in these genes lead to high lifetime risks of breast cancer and/or ovarian cancer [13-17]. Therefore, carriers of pathogenic *BRCA1/2*-mutations are eligible for risk-reducing interventions (mastectomy and adnexectomy) and/or intensive surveillance programs. Individualized advice about the most suitable options has been hampered by the considerable variability in the BC and OC risks observed amongst *BRCA1/2* carriers and by the uncertain clinical relevance of variants of undetermined significance (VUS) found during genetic screening. The studies in this thesis aim at improving the risk-assessment for individuals of HBOC (hereditary breast and/or ovarian cancer) families. The identification of genetic factors involved and understanding how they act together in modulating disease risk is essential to reach optimal and personalized risk-assessment. Progress in this field would improve accuracy of advice involving surveillance and risk-reducing strategies and ultimately patient care. To achieve this goal we have identified and analyzed variants in the *BRCA1/2* genes as well as genetic risk modifiers to determine their contribution to breast/ovarian cancer susceptibility and we have developed and optimized new tools for the characterization of genetic variants in *BRCA1/2* with respect to their pathogenicity. In particular, the studies in this thesis have led to:

- The identification and characterisation of *BRCA1/2* mutations in Portuguese families, including the description of a Portuguese *BRCA2* founder mutation, and the observed associated site-specific cancer risks (**chapter 2**);

- Identification and confirmation of *FGFR2* SNP rs2981582 as a protective factor for ovarian cancer among *BRCA1/2*-female mutation carriers (**chapters 3.1 and 3.2**), although it increases the risk of breast cancer;

- Characterisation of VUS with a putative effect on mRNA-splicing using AS-PCRs and an *ex vivo* minigene assay to assess the contribution of each allele to each transcript detected (**chapters 4.1 and 4.2**). Since these techniques require the design of new primer sets for the analysis of each variant and are time consuming, the potential use of MLPA to assess *BRCA1* exon skipping events at RNA level was evaluated (**chapter 5**), but it was not reproducible enough;

- Set up of a robust genetic signature able to distinguish between irradiated cells from BRCA1-mutation carriers and controls (**chapter 6**), which might prove useful to determine the pathogenicity of missense VUS.

In this General Discussion, the implications of the key findings for more accurate cancer risk assessment in *BRCA1/2* families are being discussed and the rationale for the culture system chosen is provided. Considerations for future studies on the identification of

genetic-risk modifiers, RNA splicing, classification of VUS, and missing heritability are being discussed more extensively.

Portuguese founder mutation and mutation age estimation

The Portuguese founder mutation, *BRCA2* c.156_157insAlu, identified in 8% of the families analysed, was initially thought to originate from the central region of Portugal, but later found to be also prevalent in the northern region [18]. The fact that we observed the Alu insertion in *BRCA2* exon 3 in more than one individual raised the possibility of a founder effect, since it seemed unlikely that this rearrangement could have arisen independently. A founder haplotype was identified and the mutation was estimated to have occurred centuries ago. Peixoto et al. have recently concluded that the mutation must have occurred even earlier than first estimated: around 558 ± 215 years ago [19]. They have assessed the presence of the mutation in European, North American, Brazilian and Hindu populations. It is remarkable that such an old mutation seems to be only present in patients of Portuguese origin [19], despite the fact that Portuguese sailors, traders, and emigrants have settled abroad since the 15th century. However, populations from Portuguese-speaking African countries – Angola, Mozambique, Guinea Bissau, Cape Verde, and São Tomé and Príncipe – were not assessed.

Founder mutations have been previously identified in other populations as summarized in Table 1. The identification of founder mutations accelerates the identification of causative mutations in patients from specific populations, by screening of the founder mutations first. For example, screening of the Portuguese founder mutation in an additional group of 157 Portuguese probands allowed the identification of an additional 14 apparently unrelated families. Three more were found during the revision of the article, before screening of the complete coding sequences of *BRCA1/2*. It is noteworthy that the *BRCA2* c.156_157insAlu is the only Portuguese founder mutation identified in high risk breast and/or ovarian cancer families.

Table 1. *BRCA1/2* founder mutations, besides the Portuguese *BRCA2* c.156_157insAlu

Population	Mutation (old nomenclature)	Mutation new nomenclature
Askenazi Jews	BRCA1 185delAG	BRCA1 c.66_67del
	BRCA1 5832insC	BRCA1 c.5713_5714insC
	BRCA2 6174delT	BRCA2 c.5946del
Icelanders	BRCA2 995delG	BRCA2 c.767del
Norwegians	BRCA1 816delGT	BRCA1 c.697_698del
	BRCA11135insA	BRCA1 c.1016_1017insA
	BRCA1 1675delA	BRCA1 c.1556del
	BRCA1 3347delAG	BRCA1 c.3228_3229del
Finns	BRCA1 IVS11+3A>G	BRCA1 c.4096+3A>G
	BRCA2 8555T>G	BRCA2 c.8327T>G
	BRCA2 IVS23+1G>A	BRCA2 c.9117+1G>A
Swedes	BRCA1 3171ins5	BRCA1 c.3052_3053ins5

French	BRCA1 3600del11	BRCA1 c.3481_3491del
Dutch	BRCA1 2804delAA	BRCA1 c.2685_2686del
	BRCA1 IVS12-1643del3835	BRCA1 c.4186-1643_4357+2020del3835
	BRCA2 5579insA	BRCA2 c.5351_5352dupA
	BRCA2 6503delTT	BRCA2 c.6275_6276del
Spanish (Galicia) [20]	BRCA1 330A>G	BRCA1 c.211A>G
Italians (Calabria)	BRCA1 5083del19	BRCA1 c.4964_4982del
Italians (Sardinia)	BRCA2 8765delAG	BRCA2 c.8537_8538del
French-Canadians (Quebec)	BRCA1 4446C>T	BRCA1 c.4327C>T
	BRCA2 3398del5	BRCA2 c.3170_3174del
	BRCA2 8765delAG	BRCA2 c.8537_8538del
Hispanics (South California)	BRCA1 2552delC	BRCA1 c.2433del
	BRCA1 2983C>A	BRCA1 c.2864C>A
Hispanics (Colombia)	BRCA1 3450delCAAG	BRCA1 c.3331_3335del
	BRCA1 5242C>A	BRCA1 c.5123C>A
	BRCA2 3034delACAA	BRCA2 c.2806_2809del
Afro-Americans	BRCA1 943ins10	BRCA1 c.824_825ins10
	BRCA11832del5	BRCA1 c.1713_1717del
	BRCA1 5296del4	BRCA1 c.5177_5180del
	BRCA2 IVS13+1G>A	BRCA2 c.4357+1G>A
South-Africans	BRCA1 2760G>T	BRCA1 c.2641G>T
Iraqi/Iranian Jews	BRCA 1 3053T>G	BRCA1 c.2934T>G
Chinese	BRCA1 1081delG	BRCA1 c.962del
Japanese	BRCA1 307T>A	BRCA1 c.188T>A
	BRCA1 2919C>T	BRCA1 c.2800C>T
	BRCA2 5802delAATT	BRCA2 c.5574_5577del
Malaysians	BRCA1 2846insA	BRCA1 c.2727_2728insA
Filipinos	BRCA1 5454delC	BRCA1 c.5335del
	BRCA2 4265delCT	BRCA2 c.4037_4038del
	BRCA2 4859delA	BRCA2 c.4631del
Pakistanis	BRCA1 4627C>A	BRCA1 c.4508C>A
	BRCA1 5622C>T	BRCA1 c.5503C>T

Adapted from [21]

Genetic modifiers of cancer-site risk and the Consortium of Investigators of Modifiers of BRCA1/2

Current risk-reducing strategies include prophylactic bilateral salpingo-oophorectomy (RRSO) and bilateral mastectomy (RRM), which are irreversible and mutilating procedures which affect future maternity, sexual relations and body image. Additionally, women that undergo RRSO have an increased risk for other diseases [22-24]. Ideally, it should be

possible to provide more personalized clinical advice to these women, which would result in an optimal, individualized risk reducing strategy. Several studies have aimed at finding genetic factors, such as SNPs, in candidate genes that affect the personal risk, sometimes with contradicting results, as indicated by the results of the association studies between polymorphisms in the progesterone receptor (*PR*) gene and the risk of BC and OC, summarized in Table 2. Follow-up studies were based on larger sample sizes to improve the statistical power of the analysis and on agnostic approaches, i.e. genome-wide association studies (GWAS), based on the common SNP-common disease concept. The great expectations that were raised by the possibilities of GWAS are now fading, as it becomes clear that the variants identified by these studies have only small effect sizes and underestimate or neglect the role of rare variants [48, 49].

Table 2. List of publications of association studies between PROGINS and +331G/A in the *PR* gene and breast cancer (BC) and ovarian cancer (OC), presented in chronological order, 1995-2012.

Article	SNP studied	Results[a]	Comments
McKenna et al, 1995[25]	PROGINS	↑ OC	
Garret et al, 1995[26]	PROGINS	≈ BC.	
Manolitsas et al, 1997[27]	PROGINS	≈ BC ; ≈ OC	
Lancaster et al, 1998[28]	PROGINS	≈ OC	
Dunning et al, 1999[29]	PROGINS	↓ BC	meta-analysis study with data from 3 studies. Borderline statistical significance.
Wang-Gohrke et al, 2000[30]	PROGINS	↓ BC	within women under age of 51
Tong et al, 2001[31]	PROGINS	≈ OC	
Runnebaum et al, 2001[32]	PROGINS	↑ OC	among *BRCA* carriers who were never exposed to oral contraceptives
Spurdle et al, 2001[33]	PROGINS	≈ OC	
Spurdle et al, 2002[34]	PROGINS	≈ BC	
De Vivo et al, 2003[35]	+331G>A	↑ BC	among women with BMI ≥25 Kg/m^2
Lancaster et al, 2003[36]	PROGINS	≈ OC	
Feigelson et al, 2004[37]	+331G>A	≈ BC	
de Vivo et al, 2004[38]	PROGINS	≈ BC	
Agoulnik et al., 2004[39]	PROGINS	↑ OC	
Pearce et al., 2005[40]	PROGINS	↑ OC ; ↓ BC	
Romano et al., 2006[41]	PROGINS and +331G>A	≈ BC; ↑ OC	association with OC was observed among women under the age of 51
Risch et al., 2006[42]	+331G/A	↑ epithelial OC	among postmenopausal women
Romano et al., 2007[43]	PROGINS and +331G>A	+331A : ↑ OC ; ↑ BBC PROGINS: ≈ BC; ≈ OC	+331A was associated with BBC among *BRCA* carriers.
Johnatty et al., 2008[44]	PROGINS and +331G>A	PROGINS: ↑ BC +331G/A: ≈ BC	

Leite et al., 2008[45]	PROGINS	↑OC	PROGINS was not in HWE in postmenopausal women with OC
Pearce et al., 2008[46]	PROGINS and +331G>A	≈ OC ; ↑ endometrioid OC	
Kotsopoulos et al., 2009 [47]	+331G/A	↑ BC	

a. Symbols:↓- decreased risk; ↑- increased risk; ≈ - no association with the disease; HWE: Hardy-Weinberg equilibrium; BMI: body mass index.

The Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) is an international group aiming at the identification of genetic modifiers of cancer risk in *BRCA1/2* mutation carriers [50]. Until now this consortium has collected genetic material from over 22,000 mutation carriers, of which approximately half developed breast cancer and 10% ovarian cancer. This is the result of a collaborative effort from several genetic centres from USA, Canada, Australia and several European countries, including the Maastricht University Medical Centre, among other Dutch centres, and the Portuguese Oncology Institute-Porto Breast Cancer Study. Initially, the consortium focused on the analysis of a few candidate gene modifiers (http://www.srl.cam.ac.uk/consortia/cimba/ pubs/pubs.html). More recently, as part of the Collaborative Oncological Gene-environment Study (COGS), which includes the CIMBA consortium, an Illumina iSelect custom array was designed covering thousands of candidate SNPs (named iCOGs), selected by the collaborating groups. *BRCA1/2* carriers and controls included in the CIMBA study were genotyped with the iCOGs chip. *FGFR2* rs2981582 and *PR* +331G>A were included in this array. Our data from this study and from the study among the families counselled in our centre provides evidence that the minor allele of *FGFR2* has a limited protective effect on ovarian cancer among *BRCA1* and *BRCA2* carriers. Others had previously shown that it is associated with increased risk of BC [51]. Thecontrasting tissue-specific effects may be due to the role of FGFR2 in PR activation, which is protective for ovarian cancer and a risk factor for BC [52-55]. Only few studies have assessed the risk of BC and OC among *BRCA1/2* carriers so far for only a limited number of other candidate alleles [43, 56, 57]. The effects of BC-modifier genes on ovarian cancer and OC-modifiers on BC should be assessed for a comprehensive estimation of the cancer-site risk modifiers among *BRCA1/2* carriers. Of note, BRIP1 variants were also associated with opposite effects on breast and ovarian cancer among mutation carriers [57], but the underlying mechanism remains to be elucidated.

The large collaborative studies will contribute to the understanding of the underlying mechanisms of carcinogenesis among *BRCA1/2* women and inter-individual variability. As a result the risk assessment of these women and individualized counselling regarding their risk-reducing options will improve. Nonetheless, the number of variants in, for example, the iCOGs chip is limited. A more comprehensive investigation of all genetic risk modifiers should become feasible by next-generation sequencing (NGS) technology. With the advances made in the NGS, it can be expected that it will soon be possible to

sequence the complete human genome of thousands of samples at a low price and high-throughput. Earlier this year, Life Technologies has announced the new Ion Torrent sequencer that is able to sequence the whole human genome for $1,000 in less than a day. The price is expected to decrease and whole-genome sequencing will become the standard for identifying both common and rare genetic risk-factors. Whole-genome sequencing has advantages over whole-exome, since it allows sequencing all introns and gene-regulatory regions. However, the interpretation of the NGS data will represent a major challenge as thousands of variants with a possible biological function will be identified. For genetic-risk modifiers, current approaches aim at identifying pathways involved, based on the clustering of the biologically relevant variants identified. Even if each variant alone has a small effect on the protein, the combination of several mildly affected proteins within the same pathway/network may explain the pathogenic or protective effect observed and the different phenotypes. Pathway analysis approach was used in GWAS data contributing to the understanding of the etiology of diseases such as hypertension [58] and late onset Alzheimer disease, allowing identification of several known and new biological pathways (abstract in Annual Meeting of the American Society of Human Genetics 2011). For example, dopamine signalling pathway was known to be involved in hypertension, but variants in genes from this pathway had not previously been identified [58]. O'Roak and colleagues successfully integrated pathway/network analysis to filter sporadic autism whole exome sequencing results [59]. Integration of gene expression and proteomic data should further assist to successfully analyse NGS data as they would also lead to identification and confirmation of the pathways and networks identified. A comparable approach was used to identify genes and dysregulated pathways in glioblastoma multiforme [60]. Exome sequencing of candidate genes may be the preferred approach at this moment for reasons of cost and speed.

Several genetic modifiers of breast and ovarian cancer risk among *BRCA1/2* carriers have been traced down to pathways, e.g., hormone-related growth, inflammation, and DNA-damage repair. It is likely that the alleles previously identified, together with other alleles from these pathways, lead to cell homeostasis disruption. Another approach to identify new candidate gene modifiers could be through the use of functional networks. Lee and colleagues have shown the use of such an approach for the identification of genetic modifiers of several genes in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* phenotypes, such as aging, size, fat content, or radiation sensitivity [61]. A network centered around BRCA1, BRCA2, ATM and CHEK2 with 188 genes has been established for the identification of new breast cancer susceptibility genes using –omics approaches [62], resulting in the identification of *HRRM*. The genes in this network are also good candidate-gene modifiers among *BRCA1/2* mutation carriers. If a candidate-gene approach becomes useful for screening genes from functional networks, greater statistical power would be achieved due to the smaller number of genes tested, but obviously this would depend on the underlying genetic heterogeneity and genetic attributable risk.

While the search for gene modifiers continues, the cumulative effect of all the genetic modifiers for individual cancer-site risk remains poorly studied and will be the challenge for the whole-genome sequencing approaches, as it is essential for the understanding of an individual's risk. Some studies reported that the risks associated with some of these SNPs are multiplicative without evidence for an interaction [51, 63, 64]. Additionally, association studies should be complemented with linkage studies using families with several affected members, as they help validating the role of the genetic modifiers in the individual phenotypic expression [65, 66]. In addition to genetic factors, there are also environmental factors and hormonal factors, i.e., age of menarche and number of full-term pregnancies, etc, that also modify the risk of breast and ovarian cancer. Ideally, the current reductionistic approach should be changed into a holistic systems biology based approach and the net effect of carrying a certain combination of genetic modifier variants, their environmental interactions, and hormonal factors should be combined and integrated in a predictive model such as BOADICEA to improve individualized risk estimation [67].

Variants of undetermined significance and splicing

VUS account for a large part of the variants identified in the *BRCA1/2* genes. In the Breast Cancer Information Core Database, almost 2,000 VUS are listed (data from January 2012) [68]. Several studies in this thesis focused on the study of VUS, mainly on those with a putative effect on mRNA splicing. We describe the use of allele-specific PCRs (AS-PCRs), besides normal biallelic amplification of cDNA, to assess the effect of a certain variant on mRNA splicing more accurately. This allowed to determine the residual transcription of full-length transcript from the mutant allele, which is important to determine the clinical relevance of the variants studied. Variants giving rise to incomplete skipping are not necessarily pathogenic, as they still produce some normal protein, which might be sufficient, but it is not known how much normal protein is needed. By using AS-PCRs, we were able to characterize the allelic contributions of 5 out of 6 splice variants in a semi-quantitative way [69]. Alternatively, exon-trapping vectors may be used, as described in chapter 4.2 [70], although the results of such experiments must be interpreted with caution and preferably used in combination with *in vivo* studies, since only a part of the coding region is expressed and the expression is in tumour cells. These *ex vivo* minigene assays are useful when AS-PCRs on patient material are not possible [71-73]. The methods used require new primer sets for the analysis of each variant. In order to assess the effect on splicing of several variants in one single fast experiment we sought to test the commercially available MLPA genomic kit in cDNA samples. Although promising, reproducibility was not as expected and therefore, for use on cDNA samples, the method should be made more robust before it can be applied in a diagnostic setting.

The evaluation of the effect of a set of VUS on mRNA splicing, led to the supplementary identification of naturally occurring isoforms in controls (some of them were not reported previously), as well as VUS that give rise to increased expression levels

of these isoforms. Identifying isoforms of the *BRCA1/2* genes was not surprising per se, since most multi-exon genes have alternatively spliced isoforms [74]. Furthermore, it has been reported that 5% of all exons are differentially spliced among individuals [75]. The generation of diverse mRNA repertoires through alternative splicing is a very powerful mechanism which contributes to expand the proteomic diversity by generating multiple products from one single gene [76]. A fraction of alternative transcripts is also a consequence of noise in splicing, not giving rise to proteins due to mRNA surveillance mechanisms or nonfolding-protein degradation [77]. Isoforms can be tissue specific, e.g. those from the *fibroblast growth factor receptor 2* gene [78], and others are constitutively expressed in most tissue types, as for example those from the NOS gene [79]. As for *BRCA1/2* genes it is currently unknown whether most of these isoforms give rise to stable protein and if the produced protein isoforms are functionally active and/or if they have similar or different functions than the full-length protein. Table 3 summarizes some of the known isoforms, and tissues where they were identified. Variants analyzed in many different tissues, such as BRCA1Δ11 and BRCA1Δ11q, illustrate the different expression

Table 3. Examples of known *BRCA1/2* isoforms

Splice variant	Keeps ORF[a]?	Tissue of expression[b]
BRCA1		
full-length [9, 83]	Yes	breast, ovary, testis, thymus, lymphocytes
BRCA1Δ9,10 [84, 85]	Yes	Normal human mammary epithelial cells, PBMC
BRCA1Δ9,10,11 [9, 86]	Yes	Lymphoblasts
BRCA1Δ9,10,11q [84, 86]	Yes	Normal human mammary epithelial cells, lymphoblasts
BRCA1Δ11 [80, 87]	Yes	peripheral blood lymphocytes, brain, colon, ovary, lung, heart, T-lymphocytes, thymus, testes, and thyroid
BRCA1Δ11q (deletion of the 3' of exon 11) [80, 84]	Yes	Normal human mammary epithelial cells, brain, colon, lung, T-lymphocytes, thymus, testes, small intestine, pancreas, liver, breast
BRCA2		
full-length [88]	Yes	Breast, thymus, lung, ovary, testis, spleen. Lower expression in brain, pancreas, prostate, leukocytes, kidney
BRCA2Δ3 [89]	Yes	Lymphocytes
BRCA2Δ6q,7 unpublished	Yes	Lymphocytes
BRCA2Δ12 [90]	Yes	peripheral blood leukocytes, kidney, smooth muscle, stomach, colon, skin, liver, bone marrow, ovary, placenta, and prostate
BRCA2Δ17,18 [91]	No	LCLs
BRCA2Δ18 [91]	No	LCLs
BRCA2Δinsi20 [92]	No	PBLs/ LCLs and primary lymphocyte culture

a. ORF: Open Reading Frame

b. Note that most isoforms were not tested in all the different tissues

For a review on BRCA1 splice isoforms see reference [93]

patterns observed among tissues [80]. For most isoforms, their expression in different tissues remains to be investigated. Furthermore, the critical expression levels, whether they lead to stable protein, and the functions of these proteins also require further studies. Using next generation sequencing, it is now possible to analyze transcription in-depth at an unprecedented genome-wide scale, using RNA-Seq procedures. By focusing on a smaller subset of genes and their transcripts, the different and even rare mRNA isoforms can be detected and mapped to the genome [81, 82]. This is a fast method to test the effect of variants on splicing, although it is much more expensive than MLPA. Additionally, knowing the normal transcription repertoire both quantitatively and qualitatively in lymphocytes, breast, and ovarian/fallopian tube cells will help to improve interpretation of the clinical relevance of those variants which increase the expression of normal isoforms, without changing expression level of the full-length transcript.

Variants of undetermined significance and genetic classifiers

Most missense changes have unclear clinical relevance. These changes may affect the protein folding and/or function. Therefore, several functional assays were developed to test the impact of VUS on the protein function, but most of them are too specific to certain protein regions and are labour intensive, which makes them less suitable for routine applications [94, 95] An indirect method to study the impact of VUS is by gene expression analysis using microarrays. Most studies have focused on agnostic approaches and used the most statistically significant genes among the differentially expressed to find a genetic classifier [96-99]. The result is that the gene signatures reported in these studies are too specific to the population used [100, 101]. This is mostly a problem in studies with a small number of samples. We used a different strategy to identify relevant genes for a genetic signature (chapter 6) by exploring the $BRCA1^{+/-}$ transcriptome and the affected pathways. Biological pathways and networks were previously integrated in genetic classifiers to, for example, successfully classify [102] or predict prognosis [103-106] of breast tumours. The authors of these studies argued that these biological-based classifiers were more accurate and reproducible. A recent study has claimed that if proper correction for gene set size is employed, then the performance of classical genetic classifiers, based on the most statistically differentially expressed genes, is similar to that of biological-based methods [107]. It is noteworthy that previously reported biological-based classifiers used averaged mRNA expression values [103, 104, 106], the difference of mRNA expression of a centered gene and its interactors [105], or the pattern of pathway activity [102], whereas our approach integrated the pathway and network analysis without loosing expression data of each selected gene. We initially identified relevant deregulated biological processes, which in fact had been previously reported to be affected in the presence of *BRCA1*-mutations [108-114]. Deregulated genes from these pathways were used to generate separate clusters for *BRCA1* mutations and wt *BRCA1* from independent studies, which had used different tissue sources and different DNA-damage agents. Therefore, our gene set showed to be more robust than those obtained through traditional classifier prediction approaches. Yet,

evaluation in a larger validation study is required. A similar approach should be equally useful for the study of *BRCA2* VUS.

The recently established Evidence-based Network for the Investigation of Germline Mutant Alleles (ENIGMA) aims at the classification of VUS in the *BRCA1/2* genes [115]. The Maastricht University Medical Centre is also part of this international consortium. ENIGMA is divided in six main groups focusing on different topics: analysis, clinical, database, pathology, functional, and splicing. The clinical group focuses on the translation of the unclassified-variant information into the clinical practice, e.g. manner of reporting VUS information in counselling, and in collecting family-cancer history and co-segregation data. Table 4 describes the five-class system that was recently proposed for classification of variants analysed by multifactorial method and corresponding testing recommendations [116]. The database workgroup focuses on developing and maintaining the VUS databases for the ENIGMA project. The pathology workgroup aims at the identification of tumour markers that may be used in the multifactorial likelihood model. The functional workgroup aims to use and further development of functional assays that may help to classify *BRCA1/2* VUS. The splicing workgroup focuses on VUS with a putative effect on splicing and aims at comparing different protocols, such as NMD inhibition and cDNA synthesis, and interpretation/classification of splice events. The analysis workgroup aims at classification of variants through statistical analysis, i.e. by multifactorial likelihood models that integrate the data from the different approaches and is currently collecting data. This method was previously used to determine pathogenicity of several variants, integrating tumour-pathology data [117, 118] as well as functional assay results [119, 120]. Further incorporation of the results from gene-expression and splicing studies in the multifactorial methods will be a useful systems biology-based approach to refine the models.

Identification of genes and biological processes affected in *BRCA1/2*-deficient cells may contribute to the identification of new candidate risk-modifier genes and processes, although not directly of the causative variants. These may be identified by NGS of these genes in a population of *BRCA1/2* carriers with different cancer phenotypes.

Table 4. Five-class system and recommendations proposed

Class	Definition	Probability of being pathogenic	Clinical Testing	Surveillance recommendations
5	Definitely pathogenic	>0.99	Test at-risk relatives for the variant	Full high-risk surveillance
4	Likely pathogenic	0.95–0.99		
3	Uncertain	0.05–0.949	Do not use as predictive testing in at-risk relatives	Counsel based on family history and other risk factors
2	Likely not pathogenic	0.001–0.049		Counsel as if no mutation detected
1	Not pathogenic	<0.001		

Adapted from [116]

Alternatively, as discussed above, the gene lists can be integrated in the filtering steps of whole-genome or whole-exome sequence data. To characterize the carcinogenic effect of the causal variants and their interactions in more depth, functional tests are required. Preferably, normal breast and ovarian tissue should be used, since these are the main affected organs. RNA interference experiments, to knockdown *BRCA1/2* and putative gene modifiers, followed by gene-expression studies will contribute to the understanding of the underlying molecular mechanisms of these genetic modifiers. The creation of a biobank to store frozen tissue sections from prophylactic preventive surgeries, would be of great importance for these future studies.

IL2/PHA-stimulated lymphocyte cultures

We have chosen to use primary lymphocyte cultures for splicing assays and irradiation experiments instead of the more widely used lymphoblastoid cell lines (LCLs) in these studies. Stimulation of primary lymphocyte cultures with IL2/PHA has already been used for a very long time in cytogenetic tests to get proliferating cells and metaphases that can be analysed by chromosome banding [121]. When freshly isolated lymphocytes, which are resting cells in the G0 phase, are stimulated with IL2/PHA, lymphocytes will enter in the S phase and proliferate [122]. Proliferating cells are known to accumulate BRCA1/2 proteins, which require active synthesis of mRNAs encoding these proteins. BRCA2 accumulates mostly during G1/S phase [123] and BRCA1 during S and G2 phases [124]. Moreover, Liu and colleagues have recently shown that EBV-transformed cell lines might introduce "illegitimate splicing" compared to fresh samples [125], similar to that found in RNA isolated from "aged" blood samples [92, 126]. It was also shown that EBV-induced immortalisation of B lymphocytes negatively affects the micronuclei induction (MN) test in *BRCA1*-mutated cells [13, 14], and therefore may not be a good cell type to study the impact of *BRCA1* variants in irradiated cells. Finally, the establishment of EBV-transformed cultures is labour intensive and not always successful. The disadvantage of primary lymphocyte cultures is that they are not immortalized like LCLs. This limits the amount of material available, and therefore the number of experiments that can be performed. Within the ENIGMA consortium, a collaborative study to compare EBV-cell lines and IL2/PHA-lymphocyte cultures regarding the different splice isoforms is currently in progress.

Missing heritability

In the majority of HBOC families, no *BRCA1/2* mutation is identified in the genetic screening. Individuals from these families may be tested for mutations in other genes or genetic regions, besides the *BRCA1/2*-coding regions, which are not yet routinely tested. Mutations in other genes such as *BARD1, RAD50, RAD51C, RAD51D, ATM* or *PALB2*, as well as in the 3' UTR of *BRCA1* and primary precursors of microRNAs miR-30c-1 and miR-17, although rare, have been described as the cause of some HBOC families [127-134]. Nowadays, genetic screening of all these genes, including intronic and promoter

regions, can be performed in parallel using next-generation sequencing (NGS), which increases considerably the speed at which mutations in these genes are being identified [135-137]. For the regions of the genes to be tested by NGS, different enrichment strategies exist, which can be commercially available, like hybridization capture methods [136], or based on in-house developed PCRs [137].

Using high-throughput whole-genome/exome sequencing of clinically well documented high-risk individuals and their relatives, it will be possible to identify the causative variants in novel breast/ovarian cancer-related genes. Due to the high number of variants identified using this technology, variant prioritisation is an important initial step. Snape and colleagues have demonstrated that in case-control studies, an average of 10 validated, rare (not present in: controls, individuals with other conditions, in the dbSNP, and/or 1000 Genomes database), protein-truncating variants, possibly disease-causing, were found in individuals with familiar breast cancer [138]. This study shows the potential of the NGS technology, the need for a good experimental design and, most importantly, the need for functional analysis of sequence variants. For autosomal-dominant diseases filtering of variants should focus on the selection of non-synonymous variants, splice acceptor and donor site mutations, and short, frame-shift coding insertions or deletions, which are likely to have a functional impact. They should not be present in controls, and must show segregation with the disease or mapping to candidate regions [139, 140]. The presence of the same mutation in more unrelated affected patients also supports a pathogenic role. In a recent study, whole-exome sequencing was performed together with the above mentioned filtering methods and a complementary filtering option for variants located in genes from DNA repair pathways [141]. This allowed identification of *XRCC2* as a new breast cancer susceptibility gene in HBOC families. Another option would have been to filter for the presence of the gene variants in functional networks. Finally, as an alternative strategy, whole human genome *in vivo* RNA interference was used successfully to identify novel tumour suppressor genes associated with breast cancer susceptibility: e.g.*MNT* and *LIFR* [142]. Mutations in the genes identified, may also account partly for the missing heritability in HBOC families.

Although the advantages of high-throughput genetic screening technology are obvious, the use of this technology will undoubtedly lead to the identification of even more variants with unclear clinical significance. The methods developed in this thesis are more generally applicable to characterize the pathogenicity of VUS from other (onco)genes as well. Efforts to assess the clinical relevance of these variants are preferably done in an international and multidisciplinary setting. The ENIGMA consortium is already considering to extend their studies to VUS from other oncogenes found in high-risk breast/ovarian cancer families. This will improve the knowledge about the gene variants and their causal relation with the disease and, hence, improve the risk assessment of hereditary breast and/or ovarian cancer patients.

Conclusions

This thesis compiles several studies that describe successful approaches for a more accurate cancer risk assessment of individuals from breast and/or ovarian cancer families. Identification of the causative mutation in the *BRCA1/2* genes is the first step (chapter 2), but due to the heterogeneous expression, accurate risk estimation and management is still not possible, requiring a more personalized and comprehensive analysis of additional risk factors involved. This was demonstrated for the genetic modifier *FGFR2* SNP rs2981582 (chapter 3), which was protective for ovarian cancer, while having opposite effects on breast cancer. Many other factors exist and it is important to understand how these factors act together before a screening test is introduced in clinical practice. Risk factors modify the effect of mutations in the *BRCA1/2* genes, but defining the impact of these *BRCA1/2* variants is already a challenge in itself. Many VUS exist, some of which might interfere with the correct mRNA-splicing in a pathogenic way (chapter 4). *In silico* prediction tools are not sufficient and functional studies are required. An assay combining allelic-specific and transcript-specific PCR has been developed to improve the study of VUS and their impact on mRNA-splicing. The results obtained can also be used to improve the interpretation of the results from the *in silico* prediction tools and the certainty thresholds. However, a much larger number of variants has to be analysed in order to confirm the suggested thresholds. MLPA, as a replacement of the more laborious RT-PCR based assay (chapter 5), turned out to be insufficiently reproducible and other methods, equally cheap and fast must be developed. Most VUS in the BRCA1/2 genes do not affect splicing, requiring a more general approach to clarify a possible pathogenic role of missense VUS. We explored the possibility of a robust classifier based on gene expression changes, analyzing the transcriptome of irradiated cells from *BRCA1*-mutation carriers and controls. Such a classifier would be extremely useful to determine the pathogenicity of missense VUS. We showed that a biologically relevant gene signature was more robust than the generally used class prediction approaches. Hierarchical clustering results obtained with our gene signature in datasets from independent studies, which used different cell types, included variants pathogenic missense variants, and DNA-damage treatments, demonstrated its robustness. Although our results are highly promising, the classifier needs to be further refined and validated, before it can be used as a diagnostic tool.

The studies included in this thesis demonstrate considerable progress in accurate cancer risk assessment for carriers of *BRCA1/2* mutations in the last couple of years, but they also reveal the immense complexity of this challenge. Detailed knowledge on the effect of the *BRCA1/2* mutations needs to be complemented with information on all other risk modifiers, interpreted within the individual context. Technical developments will allow the characterization of genetic variants and gene expression levels at an unprecedented scale, providing us with the required complete information for every person. The challenge ahead of us will be to refine existing tools and develop new ones to transform these huge amounts of data into diagnostic and prognostic relevant information.

REFERENCES

1. Lynch, H.T., et al., *Genetics, biomarkers, and control of breast cancer: A review.* Cancer Genet Cytogenet, 1984. **13**(1): p. 43-92.
2. Newman, B., et al., *Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families.* Proc Natl Acad Sci U S A, 1988. **85**(9): p. 3044-8.
3. Ford, D. and D.F. Easton, *The genetics of breast and ovarian cancer.* Br J Cancer, 1995. **72**(4): p. 805-812.
4. Newman, B., et al., *Frequency of breast cancer attributable to BRCA1 in a population-based series of American women.* JAMA, 1998. **279**: p. 915-921.
5. Peto, J., et al., *Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer.* J Natl Cancer Inst, 1999. **91**(11): p. 943-9.
6. Anglian Breast Cancer Study Group, *Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases.* . Br J Cancer, 2000. **83**(10): p. 1301-8.
7. Risch, H.A., et al., *Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer.* Am J Hum Genet, 2001. **68**: p. 700 - 710.
8. Pal, T., et al., *BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases.* Cancer, 2005. **104**(12): p. 2807-2816.
9. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
10. Wooster, R., et al., *Identification of the breast cancer susceptibility geneBRCA2.* Nature, 1995. **378**: p. 789-792.
11. Antoniou, A.C., et al., *Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study.* Genet Epidemiol, 2001. **21**(1): p. 1-18.
12. Smith, P., et al., *A genome wide linkage search for breast cancer susceptibility genes.* Genes Chromosomes Cancer, 2006. **45**(7): p. 646-55.
13. Easton, D.F., D. Ford, and D.T. Bishop, *Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium.* Am J Hum Genet, 1995. **56**(1): p. 265-71.
14. Ford, D., et al., *Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium.* Am J Hum Genet, 1998. **62**(3): p. 676-89.
15. Thompson, D. and D. Easton, *Variation in cancer risks, by mutation position, in BRCA2 mutation carriers.* Am J Hum Genet, 2001. **68**(2): p. 410-9.
16. King, M.C., J.H. Marks, and J.B. Mandell, *Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.* Science, 2003. **302**(5645): p. 643-6.
17. Antoniou, A., et al., *Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies.* Am J Hum Genet, 2003. **72**(5): p. 1117-30.
18. Peixoto, A., et al., *The c.156_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal.* Breast Cancer ResTreat, 2009. **114**(1): p. 31-38.
19. Peixoto, A., et al., *International distribution and age estimation of the Portuguese BRCA2 c.156_157insAlu founder mutation.* Breast Cancer ResTreat, 2011. **127**(3): p. 671-679.
20. Vega, A., et al., *The R71G BRCA1 is a founder Spanish mutation and leads to aberrant splicing of the transcript.* Hum Mutat, 2001. **17**(6): p. 520-521.
21. Ferla, R., et al., *Founder mutations in BRCA1 and BRCA2 genes.* Ann Oncol, 2007. **18**(suppl 6): p. vi93-vi98.
22. Robson, M., et al., *Quality of life in women at risk for ovarian cancer who have undergone risk-reducing oophorectomy.* Gynecol Oncol, 2003. **89**(2): p. 281-287.
23. Rocca, W.A., et al., *Long-term risk of depressive and anxiety symptoms after early bilateral oophorectomy.* Menopause, 2008. **15**(6): p. 1050-1059.
24. Rivera, C.M., et al., *Increased cardiovascular mortality after early bilateral oophorectomy.* Menopause, 2009. **16**(1): p. 15-23.
25. McKenna, N.J., et al., *A germline TaqI restriction fragment length polymorphism in the progesterone receptor gene in ovarian carcinoma.* Br J Cancer, 1995. **71**: p. 451-455.

26. Garrett, E., et al., *Mendelian inheritance of a TaqI restriction fragment length polymorphism due to an insertion in the human progesterone receptor gene and its allelic imbalance in breast cancer.* Cancer Res Ther Control, 1995. **4**: p. 217-22.
27. Manolitsas, T.P., et al., *No association of a 306-bp insertion polymorphism in the progesterone receptor gene with ovarian and breast cancer.* Br J Cancer, 1997. **75**: p. 1398-1399.
28. Lancaster, J.M., et al., *Progesterone receptor gene polymorphism and risk for breast and ovarian cancer.* Br J Cancer, 1998. **78**: p. 277.
29. Dunning, A.M., et al., *A systematic review of genetic polymorphisms and breast cancer risk.* Cancer Epidemiol Biomarkers Prev, 1999. **8**: p. 843-854.
30. Wang-Gohrke, S., et al., *Progesterone receptor gene polymorphism is associated with decreased risk for breast cancer by age 50.* Cancer Res, 2000. **60**(9): p. 2348-2350.
31. Tong, D., et al., *Analysis of the human progesterone receptor gene polymorphism progins in Austrian ovarian carcinoma patients.* Int J Cancer, 2001. **95**: p. 394-397.
32. Runnebaum, I.B., et al., *Progesterone receptor variant increases ovarian cancer risk in BRCA1 and BRCA2 mutation carriers who were never exposed to oral contraceptives.* Pharmacogenetics, 2001. **11**(7): p. 635-638.
33. Spurdle, A.B., et al., *No significant association between progesterone receptor exon 4 Val660Leu G/T polymorphism and risk of ovarian cancer.* Carcinogenesis, 2001. **22**(5): p. 717-721.
34. Spurdle, A.B., et al., *The progesterone receptor exon 4 val660leu G/T polymorphism and risk of breast cancer in Australian women.* Cancer Epidemiol Biomarkers Prev, 2002. **11**: p. 439-443.
35. Vivo, I.D., et al., *A functional polymorphism in the progesterone receptor gene is associated with an increase in breast cancer risk.* Cancer Res, 2003. **63**(17): p. 5236-5238.
36. Lancaster, J.M., et al., *No relationship between ovarian cancer risk and progesterone receptor gene polymorphism in a population-based, case-control study in North Carolina.* Cancer Epidemiol Biomarkers Prev, 2003. **12**: p. 226-227.
37. Feigelson, H.S., et al., *No Association between the Progesterone Receptor Gene +331G/A Polymorphism and Breast Cancer.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(6): p. 1084-1085.
38. Vivo, I.D., et al., *The progesterone receptor Val660->Leu polymorphism and breast cancer risk.* Breast Cancer Res, 2004. **6**(6): p. R636-R639.
39. Agoulnik, I.U., et al., *A germline variation in the progesterone receptor gene increases transcriptional activity and may modify ovarian cancer risk.* J Clin Endocrinol Metab, 2004. **89**(12): p. 6340-6347.
40. Pearce, C.L., et al., *Clarifying the PROGINS Allele Association in Ovarian and Breast Cancer Risk: A Haplotype-Based Analysis.* J. Natl. Cancer Inst., 2005. **97**(1): p. 51-59.
41. Romano, A., et al., *Two functionally relevant polymorphisms in the human progesterone receptor gene (+331 G/A and progins) and the predisposition for breast and/or ovarian cancer.* Gynecol Oncol, 2006. **101**: p. 287-295.
42. Risch, H.A., et al., *PGR +331 A/G and increased risk of epithelial ovarian cancer.* Cancer Epidemiol Biomarkers Prev, 2006. **15**: p. 1738-1741.
43. Romano, A., et al., *Impact of two functional progesterone receptor polymorphisms (PRP): +331G/A and PROGINS on the cancer risks in familial breast/ovarian cancer.* Open Cancer J, 2007. **1**: p. 1-8.
44. Johnatty, S., et al., *Progesterone receptor polymorphisms and risk of breast cancer: results from two Australian breast cancer studies.* Breast Cancer Res Treat, 2008. **109**: p. 91-99.
45. Leite, D.B., et al., *Progesterone receptor (PROGINS) polymorphism and the risk of ovarian cancer.* Steroids, 2008. **73**: p. 676-680.
46. Pearce, C.L., et al., *Progesterone receptor variation and risk of ovarian cancer is limited to the invasive endometrioid subtype: results from the ovarian cancer association consortium pooled analysis.* Br J Cancer, 2008. **98**(2): p. 282-288.
47. Kotsopoulos, J., et al., *+331G/A variant in the progesterone receptor gene, postmenopausal hormone use and risk of breast cancer.* Int J Cancer, 2009. **125**(7): p. 1685-1691.
48. Iles, M.M., *What can genome-wide association studies tell us about the genetics of common disease?* PLoS Genet, 2008. **4**(2): p. e33.

49. Hindorff, L.A., et al., *Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.* Proc Natl Acad Sci USA, 2009. **106**(23): p. 9362-9367.
50. Chenevix-Trench, G., et al., *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, 2007. **9**(2): p. 104.
51. Antoniou, A.C., et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. **82**: p. 937-948.
52. Ho, S.-M. (2003) *Estrogen, progesterone and epithelial ovarian cancer*. Reprod Biol Endocrinol **1**, 73 DOI: 10.1186/1477-7827-1-73.
53. Romano, A., et al., *The PROGINS polymorphism of the human progesterone receptor diminishes the response to progesterone.* J Mol Endocrinol, 2007. **38**: p. 331-350.
54. Charles, N.J., P. Thomas, and C.A. Lange, *Expression of membrane progesterone receptors (mPR/PAQR) in ovarian cancer cells: implications for progesterone-induced signaling events.* Horm Canc, 2010. **1**: p. 167-176.
55. Aupperlee, M., et al., *Progestins and breast cancer.* Breast Dis, 2005-2006. **24**: p. 37-57.
56. Rebbeck, T.R., et al., *Modification of BRCA1-associated breast and ovarian cancer risk by BRCA1-interacting genes.* Cancer Res, 2011. **71**(17): p. 5792-5805.
57. Engel, C., et al., *Association of the Variants CASP8 D302H and CASP10 V410I with Breast and Ovarian Cancer Risk in BRCA1 and BRCA2 Mutation Carriers.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(11): p. 2859-2868.
58. Torkamani, A., E.J. Topol, and N.J. Schork, *Pathway analysis of seven common diseases assessed by genome-wide association.* Genomics, 2008. **92**(5): p. 265-272.
59. O/'Roak, B.J., et al., *Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations.* Nature, 2012. **advance online publication**.
60. Kim, Y.-A., S. Wuchty, and T.M. Przytycka, *Identifying causal genes and dysregulated pathways in complex diseases.* PLoS Comput Biol, 2011. **7**(3): p. e1001095.
61. Lee, I., et al., *Predicting genetic modifier loci using functional gene networks.* Genome Res, 2010. **20**(8): p. 1143-1153.
62. Pujana, M.A., et al., *Network modeling links breast cancer susceptibility and centrosome dysfunction.* Nat Genet, 2007. **39**(11): p. 1338-1349.
63. Antoniou, A.C., et al., *Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers.* HumMol Genet, 2009. **18**(22): p. 4442-4456.
64. Antoniou, A., et al. (2012) *Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers*. Breast Cancer Res **14**, R33 DOI: 10.1186/bcr3121.
65. Bourgain, C., et al., *Are genome-wide association studies all that we need to dissect the genetic component of complex human diseases?* Eur J Hum Genet, 2006. **15**(3): p. 260-263.
66. Clerget-Darpoux, F. and R.C. Elston, *Are linkage analysis and the collection of family data dead? Prospects for family studies in the age of genome-wide association.* Hum Hered, 2007. **64**(2): p. 91-96.
67. Antoniou, A.C., et al., *The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions.* Brit J Cancer, 2008. **98**(8): p. 1457-1466.
68. *The Breast Information Core Database*.
69. Brandão, R.D., et al., *Characterisation of unclassified variants in the BRCA1/2 genes with a putative effect on splicing.* Breast Cancer Res Treat, 2011. **129**(3): p. 971-982.
70. Brandão, R.D., et al., *BRCA1 c.4987-3C>G is a pathogenic mutation.* Breast Cancer Res Treat, 2012. **131**(2): p. 723-725.
71. Schneider, B., A. Koppius, and R. Sedlmeier, *Use of an exon-trapping vector for the evaluation of splice-site mutations.* Mamm Genome, 2007. **18**(9): p. 670-676.
72. Bonnet, C., et al., *Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene.* J Med Genet, 2008. **45**(7): p. 438-446.

73. Sanz, D.J., et al., *A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients.* Clin Cancer Res, 2010. **16**(6): p. 1957-1967.
74. Modrek, B., et al., *Genome-wide detection of alternative splicing in expressed sequences of human genes.* Nucl Acids Res, 2001. **29**(13): p. 2850-2859.
75. Kwan, T., et al., *Heritability of alternative splicing in the human genome.* Genome Res, 2007. **17**(8): p. 1210-1218.
76. Nilsen, T.W. and B.R. Graveley, *Expansion of the eukaryotic proteome by alternative splicing.* Nature, 2010. **463**(7280): p. 457-463.
77. Melamud, E. and J. Moult, *Stochastic noise in splicing machinery.* Nucl Acids Res, 2009. **37**(14): p. 4873-4886.
78. Katoh, M., *Cancer genomics and genetics of FGFR2 (Review).* Int J Oncol, 2008. **33**: p. 233-237.
79. Förstermann, U., J.-p. Boissel, and H. Kleinert, *Expressional control of the `constitutive' isoforms of nitric oxide synthase (NOS I and NOS III).* FASEB J, 1998. **12**(10): p. 773-790.
80. Thakur, S., et al., *Localization of BRCA1 and a splice variant identifies the nuclear localization signal.* Mol Cell Biol, 1997. **17**(1): p. 444-452.
81. Levin, J., et al., *Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts.* Genome Biol, 2009. **10**(10): p. R115.
82. Mercer, T.R., et al., *Targeted RNA sequencing reveals the deep complexity of the human transcriptome.* Nat Biotech, 2012. **30**(1): p. 99-104.
83. Xu, C.-F., et al., *Distinct transcription start sites generate two forms of BRCA1 mRNA.* Hum Mol Genet, 1995. **4**(12): p. 2259-2264.
84. Lu, M., et al., *Characterization of functional messenger RNA splice variants of BRCA1 expressed in nonmalignant and tumor-derived breast cells.* Cancer res, 1996. **56**(20): p. 4578-81.
85. Balz, V., et al., *Analysis of BRCA1, TP53, and TSG101 germline mutations in German breast and/or ovarian cancer families.* Cancer Genet Cytogenet, 2002. **138**(2): p. 120-127.
86. Xu, C.F., et al., *Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families.* Genes Chromosomes Cancer, 1997. **18**(2): p. 102-110.
87. Keaton, J.C., et al., *A biochemical analysis demonstrates that the BRCA1 intronic variant IVS10-2A?C is a mutation.* J Hum Genet, 2003. **48**(8): p. 399-403.
88. Tavtigian, S.V., et al., *The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds.* Nat Genet, 1996. **12**(3): p. 333-337.
89. Santarosa, M., A. Viel, and M. Boiocchi, *Splice variant lacking the transactivation domain of the BRCA2 gene and mutations in the splice acceptor site of intron 2.* Genes Chromosomes Cancer, 1999. **26**(4): p. 381-382.
90. Bieche, I. and R. Lidereau, *Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue.* Cancer Res, 1999. **59**(11): p. 2546-2550.
91. Walker, L.C., et al., *Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity.* Hum Mutat, 2010. **31**(6): p. E1484-E1505.
92. Speevak, M.D., et al., *Alternatively spliced, truncated human BRCA2 isoforms contain a novel coding exon.* Eur J Hum Genet, 2003. **11**(12): p. 951-954.
93. Orban, T.I. and E. Olah, *Emerging roles of BRCA1 alternative splicing.* Mol Pathol, 2003. **56**(4): p. 191-197.
94. Radice, P., et al., *Unclassified variants in BRCA genes: guidelines for interpretation.* Ann Oncol, 2011. **22**(suppl 1): p. i18-i23.
95. Millot, G.A., et al., *A guide for functional analysis of BRCA1 variants of uncertain significance (VUS).* Hum Mutat, 2012.
96. Kote-Jarai, Z., et al., *Gene Expression Profiling after Radiation-Induced DNA Damage Is Strongly Predictive of BRCA1 Mutation Carrier Status.* Clin Cancer Res, 2004. **10**(3): p. 958-963.
97. Kote-Jarai, Z., et al., *Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage.* Clin Cancer Res, 2006. **12**(13): p. 3896-3901.

98. Waddell, N., et al., *BRCA1 and BRCA2 missense variants of high and low clinical significance influence lymphoblastoid cell line post-irradiation gene expression.* PLoS Genet, 2008. **4**(5): p. e1000080.
99. Walker, L.C., et al., *Use of DNA-damaging agents and RNA pooling to assess expression profiles associated with BRCA1 and BRCA2 mutation status in familial breast cancer patients.* PLoS Genet, 2010. **6**(2): p. e1000850.
100. Michiels, S., S. Koscielny, and C. Hill, *Prediction of cancer outcome with microarrays: a multiple random validation strategy.* Lancet, 2005. **365**(9458): p. 488-492.
101. Ein-Dor, L., O. Zuk, and E. Domany, *Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer.* Proc Natl Acad Sci USA, 2006. **103**(15): p. 5923-5928.
102. Gatza, M.L., et al., *A pathway-based classification of human breast cancer.* Proc Natl Acad Sci USA, 2010. **107**(15): p. 6994-6999.
103. Chuang, H.-Y., et al., *Network-based classification of breast cancer metastasis.* Mol Syst Biol, 2007. **3**.
104. Lee, E., et al., *Inferring pathway activity toward precise disease classification.* PLoS Comput Biol, 2008. **4**(11): p. e1000217.
105. Taylor, I.W., et al., *Dynamic modularity in protein interaction networks predicts breast cancer outcome.* Nat Biotech, 2009. **27**(2): p. 199-204.
106. Abraham, G., et al., *Prediction of breast cancer prognosis using gene set statistics provides signature stability and biological context.* BMC Bioinformatics, 2010. **11**(1): p. 277.
107. Staiger, C., et al., *A critical evaluation of network and pathway-based classifiers for outcome prediction in breast cancer.* PLoS One, 2012. **7**(4): p. e34796.
108. Xu, B., S.-t. Kim, and M.B. Kastan, *Involvement of Brca1 in S-Phase and G2-Phase Checkpoints after Ionizing Irradiation.* Mol Cell Biol, 2001. **21**(10): p. 3445-3450.
109. Shorrocks, J., et al., *Primary fibroblasts from BRCA1 heterozygotes display an abnormal G1/S cell cycle checkpoint following UVA irradiation but show normal levels of micronuclei following oxidative stress or mitomycin C treatment.* Int J Radiat Oncol Biol Phys, 2004. **58**(2): p. 470-478.
110. Xu, X., *Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells.* Mol Cell, 1999. **3**: p. 389-395.
111. Su, F., et al., *Hemizygosity for Atm and Brca1 influence the balance between cell transformation and apoptosis.* Radiat Oncol, 2010. **5**(1): p. 15.
112. Friedenson, B., *Inflammation targets specific organs for cancer in carriers of BRCA1/2 pathway mutations*. 2011, Available from Nature Precedings.
113. Kote-Jarai, Z., *Increased level of chromosomal damage after irradiation of lymphocytes from BRCA1 mutation carriers.* British Journal of Cancer, 2006. **94**: p. 308-310.
114. Rothfuβ, A., et al., *Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families.* Cancer Res, 2000. **60**(2): p. 390-394.
115. Spurdle, A.B., et al., *ENIGMA—Evidence-based network for the interpretation of germline mutant alleles: An international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes.* Hum Mutat, 2012. **33**(1): p. 2-7.
116. Plon, S.E., et al., *Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results.* Hum Mutat, 2008. **29**(11): p. 1282-1291.
117. Lovelock, P., et al., *Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants?* Breast Cancer Res, 2007. **9**(6): p. R82.
118. Spurdle, A.B., et al., *Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis - a report from the kConFab Investigators.* J Clin Oncol, 2008. **26**(10): p. 1657-1663.
119. Farrugia, D.J., et al., *Functional assays for classification of BRCA2 variants of uncertain significance.* Cancer Res, 2008. **68**(9): p. 3523-3531.
120. Spearman, A.D., et al., *Clinically applicable models to characterize BRCA1 and BRCA2 variants of uncertain significance.* J Clin Oncol, 2008. **26**(33): p. 5393-5400.
121. Gersen, S.L. and M.B. Keagle, *The Principals of Clinical Cytogenetics*. 2 ed. 2005, New Jersey, USA: Humana Press.

122. Kaczmarek, L., B. Calabretta, and R. Baserga, *Expression of cell-cycle-dependent genes in phytohemagglutinin-stimulated human lymphocytes.* Proc Natl Acad Sci U S A, 1985. **82**(16): p. 5375-5379.
123. Vaughn, J.P., et al., *Cell cycle control of BRCA2.* Cancer Research, 1996. **56**(20): p. 4590-4594.
124. MacLachlan, T.K. and W. El-Deiry, *Functional Interactions Between BRCA1 and the Cell Cycle*. 2005, Austin, USA: Landes Bioscience.
125. Liu, Y., et al., *PTEN transcript variants caused by illegitimate splicing in "aged" blood samples and EBV-transformed cell lines.* Hum Genet, 2010. **128**(6): p. 609-614.
126. Wimmer, K., et al., *Illegitimate splicing of the NF1 gene in healthy individuals mimics mutation-induced splicing alterations in NF1 patients.* Hum Genet, 2000. **106**: p. 311-313.
127. Heikkinen, K., et al., *RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability.* Carcinogenesis, 2006. **27**(8): p. 1593-1599.
128. Renwick, A., et al., *ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(8): p. 873-875.
129. Erkko, H., et al., *A recurrent mutation in PALB2 in Finnish cancer families.* Nature, 2007. **446**(7133): p. 316-319.
130. Shen, J., C.B. Ambrosone, and H. Zhao, *Novel genetic variants in microRNA genes and familial breast cancer.* Int J Cancer, 2009. **124**(5): p. 1178-1182.
131. Meindl, A., et al., *Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene.* Nat Genet, 2010. **42**(5): p. 410-414.
132. Ratajska, M., et al., *Cancer predisposing BARD1 mutations in breast–ovarian cancer families.* Breast Cancer ResTreat, 2011: p. 1-9.
133. Lheureux, S., et al., *Two novel variants in the 3'UTR of the BRCA1 gene in familial breast and/or ovarian cancer.* Breast Cancer ResTreat, 2011. **125**(3): p. 885-891.
134. Loveday, C., et al., *Germline mutations in RAD51D confer susceptibility to ovarian cancer.* Nat Genet, 2011. **43**(9): p. 879-882.
135. Morgan, J.E., et al., *Genetic diagnosis of familial breast cancer using clonal sequencing.* Hum Mutat, 2010. **31**(4): p. 484-491.
136. Walsh, T., et al., *Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing.* Proc Natl Acad Sci U S A, 2010. **107**(28): p. 12629-12633.
137. De Leeneer, K., et al., *Massive parallel amplicon sequencing of the breast cancer genes BRCA1 and BRCA2: opportunities, challenges, and limitations.* Hum Mutat, 2011. **32**(3): p. 335-344.
138. Snape, K., et al. (2012) *Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer*. Breast Cancer Res Treat, 1-5 DOI: 10.1007/s10549-012-2057-x.
139. Wang, J.L., et al., *TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing.* Brain, 2010. **133**(12): p. 3510-3518.
140. Johnson, J.O., et al., *Exome sequencing reveals VCP mutations as a cause of familial ALS.* Neuron, 2010. **68**(5): p. 857-864.
141. Park, D.J., et al., *Rare mutations in XRCC2 increase the risk of breast cancer.* Am J Hum Genet, 2012. **90**(4): p. 734-739.
142. Iorns, E., et al. (2012) *Whole genome in vivo RNAi screening identifies the leukemia inhibitory factor receptor as a novel breast tumor suppressor*. Breast Cancer Res Treat, 1-13 DOI: 10.1007/s10549-012-2068-7.

SUMMARY

The challenge of genetic counselling for persons at high risk of breast and/or ovarian cancer is individualized risk estimation and management. Although other factors may be involved, personalized counselling will obviously not be possible until identification of the genetic cause of the disease for all the HBOC families. During the genetic screening of the major susceptibility *BRCA1* and *BRCA2* genes, several variants of undetermined clinical significance (VUS) are identified. Their clinical relevance must be assessed. Secondly, risk factors other than the mutation in the *BRCA1* or *BRCA2* genes should be integrated in the risk estimation. This thesis compiles several studies aiming at improving the risk assessment of individuals from breast and/or ovarian cancer families with BRCA1/2 genetic variants using genomic technology and transcriptional studies.

Chapter 2 describes the mutations identified during the genetic screening of Portuguese families, who attended the Oncology Centre of Lisbon Francisco Gentil for genetic counselling. The identification and characterization of a Portuguese founder mutation, *BRCA2* c.156_157insAlu, is reported in detail. This mutation, present in 8% (17/210) of the families, was estimated to have occurred many centuries ago. Additionally, in a group of 53 patients, of which 3 carried the founder mutation, other 15 different mutations were identified in 16 individuals. Screening Portuguese high-risk families and their descendants for the founder mutation, prior to the screening of the complete *BRCA1/2*-coding regions is rapid and cost-effective.

Prediction of the cancer-site risk is not possible, even for individuals with the same mutation, irrespective of whether they belong to the same family, although a certain tendency could be observed within each family. Several studies have suggested the existence of additional genetic modifying factors. Since cancer sites affected by *BRCA1/2* mutations mainly involve organs affected by steroid hormone metabolism, we hypothesised that functional polymorphisms in genes involved in steroid hormone metabolism could influence cancer-site risk. The pilot study described in **chapter 3.1** included women from families with a *BRCA1* or *BRCA2* mutation diagnosed at our centre. Two of those genes, the progesterone receptor (*PR*) gene and *FGFR2*, a PR activator, were analysed as putative genetic risk modifiers. Additionally, three polymorphisms in *TNRC9* (*TOX3*) and *CASP8*, previously described to modify the risk of sporadic and inherited breast cancer, were also investigated. The results of this study indicated that the minor allele of *FGFR2* (rs2981582) is protective against ovarian cancer, but increases the risk of bilateral breast cancer. The other polymorphisms did not show or confirm an association with breast cancer or ovarian cancer, probably due to lack of power. In **chapter 3.2** we have analyzed the risk of *FGFR2* (rs2981582) for ovarian cancer in an international and much larger population of more than 20.000 *BRCA1/2* female carriers. This study confirmed that *FGFR2* is indeed protective for ovarian cancer (HR= 0.86, p= 0.090 and HR= 0.67, p= 0.005 for BRCA1 and BRCA2,

respectively), especially among *BRCA2* carriers. These results may have a clinical impact in the risk-reducing strategies of women with a *BRCA1/2* mutation.

A large proportion of the genetic variants identified during genetic screening are variants of unclear clinical relevance (VUS), which pose problems to pre-symptomatic DNA-testing and accurate genetic counselling. **Chapters 4.1 and 4.2** evaluate and characterise the effect of a set of *BRCA1/2* VUS predicted by *in silico* models to have an effect on mRNA splicing, using a combination of allele- and transcript-specific primers. The results determined conclusively the pathogenicity of six (out of 14) variants: three were considered pathogenic and three non-pathogenic. From the remaining variants: one variant was determined to be likely pathogenic; three other variants were found to increase the expression of naturally occurring isoforms, without decreasing the expression of the full-length transcript, and therefore remain unclassified; four variants had no effect on splicing, however their pathogenicity could not be excluded, since they lead to missense changes that might affect the protein function. Furthermore, this work allowed to suggest a likelihood threshold that may be used for the software models for future selection of variants with a putative effect on splicing. In summary, except for variants affecting a non-canonical splice site, three algorithms must predict at least 10% decrease or two algorithms must predict 20% decrease of the splice site score. In **Chapter 4.2** an *ex vivo* minigene splicing assay was used as a complementary technique to those used in the previous study. This approach might substitute the allele-specific analysis when allele-specific amplification cannot be used. The study provided definite evidence that the full-length transcript was not expressed from the allele with the variant c.4987-3C>G and enabled to conclude that it is a pathogenic variant.

Chapter 5 evaluates the potential of MLPA (Multiplex Ligation-dependent Probe Amplification) to detect complete exon skipping events in mRNA from *BRCA1*. The intended use of the commercial *BRCA1* MLPA kit is the detection of genomic rearrangements. Currently, there is no routine technique available to detect complete exon skipping events in mRNA without using specific primers. The availability of such a technique would be very useful in the diagnostic setting. Although after several optimization steps the technique could detect exon skipping events, the results were not fully satisfactory for routine RNA analysis in a diagnostic setting since it was not always reproducible.

Chapter 6 explores the transcriptome of cultured $BRCA1^{+/-}$ lymphocytes after irradiation, aiming at the identification of biologically relevant genes that could be incorporated in a genetic classifier able to distinguish between *BRCA1* mutation-carrier and non-carrier samples. We identified many differentially expressed genes associated with biological processes in which *BRCA1* is known to play important roles. More specifically, we showed that $BRCA1^{+/-}$ lymphocytes, in response to irradiation, had deficient cell cycle arrest, decreased apoptosis, decreased immune response, increased chromosomal instability, and decreased mitotic spindle assembly and chromosome separation, leading to increased micronuclei induction. A gene signature including 25 genes from these pathways

and from 141 overlapping genes between two independent groups of samples allowed to cluster separately samples from *BRCA1*-mutated and normal lymphocytes. This approach was more robust than using just the most statistically different expressed genes, as generally used in class prediction approaches. The hierarchical clustering results obtained with this gene signature in datasets from independent studies, which used different cell types and DNA-damage agents, supported this observation.

In conclusion, the studies compiled in this thesis lead to the identification and characterization of several *BRCA1/2* pathogenic variants. A polymorphism in *FGFR2* was found to modify the cancer-site risk among *BRCA1/2*-carriers. However, the result of the combination with other risk and protective factors remains unclear. It is necessary to understand this before the test can be used in a clinical setting. A genetic classifier able to determine the clinical relevance of *BRCA1*-VUS was developed and the results obtained in independent datasets were very promising. Further analysis of different datasets would allow refining the gene list and test its reproducibility. The studies in this thesis contributed to the long and challenging process to get to personalized risk assessment for *BRCA1/2*-mutation carriers.

SAMENVATTING

We staan voor de grote uitdaging om de risicoschattingen en het klinisch management verder te personaliseren voor individuen met een hoog risico op borst- en/of eierstokkanker (HBOC) in de genetische counseling. Om dit mogelijk te maken zullen in eerste instantie alle genetische factoren die een rol spelen in HBOC-families geïdentificeerd moeten worden. Mutaties in de *BRCA1* en *BRCA2* genen zijn tot nu toe de belangrijkste oorzaak van HBOC gebleken. Maar er worden in deze genen ook veel genetische varianten met een onbekende klinische relevantie gevonden, zogenaamde "*variants of undetermined significance"* (VUS). Hun klinische relevantie moet duidelijk worden, en daarnaast moeten ook andere (genetische) risicofactoren in de risicoschattingen meegenomen worden. Dit proefschrift bevat een aantal studies die gericht zijn op het verbeteren van de risicoschattingen voor personen uit HBOC families met afwijkingen in de *BRCA1/2*-genen, met behulp van DNA-, en RNA- technologieën.

Hoofdstuk 2 beschrijft de mutaties die gevonden zijn tijdens de genetische screening van Portugese families die het Oncologisch Centrum Francisco Gentil van Lissabon bezochten voor genetische counseling. Met name de identificatie en karakterisering van de Portugese "founder" mutatie c.156_157insAlu in het *BRCA2*-gen wordt gedetailleerd beschreven. Dergelijke mutaties worden generatie op generatie doorgegeven en raken daardoor wijdverspreid in een populatie. Voor deze Portugese mutatie, gevonden in 8% (17/210) van de families, wordt geschat dat deze vele eeuwen geleden in de populatie geïntroduceerd werd. In een additionele groep van 53 patiënten, werd naast 3 personen met de "founder" mutatie, 15 andere mutaties gedetecteerd in 16 personen. Het screenen van Portugese hoog-risico families voor de "founder" mutatie, voorafgaand aan de screening van de complete coderende regio's van de *BRCA1/2*-genen, is een snelle en kosteneffectieve methode om een mutatie te detecteren.

Voorspellen of iemand meer risico heeft op borst-, of eierstokkanker is momenteel niet mogelijk, ook niet voor individuen met dezelfde mutatie en zelfs niet voor individuen die tot dezelfde familie behoren, alhoewel er binnen een familie wel een bepaalde tendens waar te nemen kan zijn. In verschillende studies is het bestaan van additionele genetische factoren die het kankerrisico modificeren gesuggereerd. De organen waarin tumoren ontstaan ten gevolge van *BRCA1/2* mutaties staan onder invloed van steroïde hormonen. Dit heeft geleid tot het formuleren van de hypothese dat functionele polymorfismen, in genen die betrokken zijn bij het steroïde hormoon metabolisme de plaats waar de kanker optreedt kunnen beïnvloeden. In de pilotstudie beschreven in hoofdstuk 3.1 zijn vrouwen opgenomen uit families met een *BRCA1* of *BRCA2* mutatie, zoals vastgesteld in het klinisch genetisch centrum te Maastricht. Polymorfismen in het progesteron receptor (*PR*)-, en *FGFR2*-gen (een PR-activator) werden geanalyseerd als mogelijke additionele genetische risicofactoren. Tevens, werden drie polymorfismen in *TNRC9* (*TOX3*) en *CASP8* onderzocht. Voor deze polymorfismen werd eerder beschreven dat ze het risico op

sporadische en erfelijke borstkanker wijzigen. De resultaten van deze studie toonden aan dat het "minor" allel van het *FGFR2* polymorfisme rs2981582, beschermt tegen eierstokkanker, maar het risico op tweezijdige borstkanker verhoogt. De andere polymorfismen vertoonden geen associatie met borst-, of eierstokkanker, waarschijnlijk door een gebrek aan voldoende statistisch onderscheidend vermogen.

In hoofdstuk 3.2 hebben we het risico van het *FGFR2* polymorfisme rs2981582 op eierstokkanker in een veel grotere, internationale groep van meer dan 20.000 *BRCA1/2* draagsters bestudeerd. Deze studie bevestigt dat het *FGFR2* polymorfisme inderdaad beschermt tegen eierstokkanker, met name bij *BRCA2*-mutatiedraagsters (HR = 0,86, p = 0,090 en HR = 0,67, p = 0,005 voor *BRCA1* en *BRCA2*, respectievelijk). Deze resultaten kunnen klinische gevolgen hebben met betrekking tot kankerrisico-reducerende strategieën voor vrouwen met een *BRCA1/2* mutatie.

Een groot deel van de genetische varianten die geïdentificeerd worden tijdens de screening van genen, zijn varianten met een onduidelijke klinische relevantie (VUS). Dit maakt het moeilijk om niet-aangedane familieleden gericht genetisch te counselen op basis van DNA testen. In **Hoofdstukken 4.1** en **4.2** wordt het effect van een set *BRCA1/2* varianten op mRNA splicing geëvalueerd en gekarakteriseerd met behulp van allel-, en transcript-specifieke primers. Voor deze varianten werd vooraf een mogelijk effect op mRNA voorspeld door computerprogramma's, hetgeen vervolgens experimenteel werd getest. Uit de resultaten kon het pathogene karakter definitief worden vastgesteld voor zes van de 14 varianten: drie van de zes zijn daadwerkelijk pathogeen, tegenover drie varianten waarvoor toch geen afwijking in splicing werd gevonden en daarmee niet-pathogeen zijn. Voor één van de overige varianten werd geconcludeerd dat het waarschijnlijk een pathogene variant betreft, maar dat verder onderzoek nodig is om dit te bevestigen. Voor drie andere varianten werd gevonden dat de expressie van natuurlijke mRNA isovormen verhoogd is, terwijl de expressie van het normale transcript met de volledige lengte gelijk blijft. De biologische relevantie hiervan is onduidelijk en de varianten blijven daarom in de categorie VUS. Vier varianten hadden geen effect op splicing, maar hun pathogeniciteit kon niet worden uitgesloten, omdat ze wel leiden tot de verandering van één aminozuur in het eiwit, hetgeen de functie van het eiwit kan beïnvloeden. Ook deze varianten blijven voorlopig VUS. Op basis van deze studie kon ook een voorstel worden gedaan met betrekking tot de drempelwaarden die gehanteerd kunnen worden voor een betere selectie van varianten die met grote waarschijnlijkheid een effect op splicing zullen hebben. Ten minste drie voorspellende computeralgoritmen moeten een minimale afname van 10% laten zien in hun voorspellende score, of 2 algoritmen die een minimale afname van 20% laten zien. Dit geldt echter niet voor varianten in zogenaamde "non-canonical splice sites". In **hoofdstuk 4.2** werd een "*ex vivo* minigen splicing assay" gebruikt ter aanvulling van de experimenten zoals beschreven in het vorige hoofdstuk. Met deze aanpak kan een allelspecifieke analyse worden uitgevoerd, indien een normale allelspecifieke amplificatie in aanwezigheid van twee allelen niet mogelijk is. Het onderzoek leverde het definitieve bewijs dat het allel met de *BRCA1* variant c.4987-3C>G geen normaal transcript van de

volledige lengte meer kan maken. Hieruit kon geconcludeerd worden dat het een pathogene variant betreft.

In **Hoofdstuk 5** wordt de bruikbaarheid van de MLPA (Multiplex Ligatie-afhankelijke Probe Amplification) techniek geëvalueerd voor de detectie van afwijkend *BRCA1* mRNA waarin exonen volledig worden overgeslagen ("skipping"). Normaal gesproken wordt de commerciële BRCA1 MLPA-kit gebruikt voor de detectie van deleties of duplicaties op DNA-nivo. Momenteel is er geen techniek beschikbaar die routinematig gebruikt kan worden voor de universele detectie van de skipping van welk exon dan ook in een transcript in één test. De beschikbaarheid daarvan zou erg bruikbaar zijn in een diagnostische setting. Uit de gedane experimenten bleek dat het met MLPA mogelijk was om, na een aantal optimalisatie stappen, exon-skipping te detecteren. De reproduceerbaarheid was echter niet voldoende om de techniek routinematig en diagnostisch toe te passen voor RNA analyse.

In **Hoofdstuk 6** wordt het transcriptoom van gekweekte *BRCA1*$^{+/-}$ lymfocyten na radioactieve bestraling geanalyseerd. De bedoeling hiervan was om biologisch relevante genen te identificeren die het onderscheid kunnen maken tussen *BRCA1*-mutatiedragers en niet-dragers, door deze genen te includeren in een genetische "classifier". Er werden veel van dergelijke genen gevonden waarvoor ook bekend is dat ze een rol spelen in biologische processen waar ook *BRCA1* bij betrokken is. Om meer precies te zijn hebben we gevonden dat *BRCA1*$^{+/-}$ lymfocyten, in reactie op bestraling, een aantal genen anders tot expressie brengen die duiden op een deficiënte celcyclus, verminderde apoptose en immuunrespons, meer chromosomale instabiliteit en verminderde assemblage van de mitotische spoel en scheiding van de chromosomen, hetgeen leidt tot een verhoogd aantal micronuclei. Een lijst van genen werd opgesteld, waaronder 25 genen uit deze processen en 141 genen die in twee onafhankelijke experimenten gemeenschappelijk anders tot expressie bleken te komen in vergelijking met de controles. Met deze genen konden monsters van *BRCA1*-mutatiedraagster onderscheiden worden van normale lymfocyten. Deze aanpak was robuuster dan het gebruik van alleen de genen die statistisch gezien het meest verschillend tot expressie komen, zoals gewoonlijk wordt gedaan bij benaderingen om monsters toe te wijzen aan een bepaalde groep. Deze observatie wordt ondersteund door het feit dat met behulp van de opgestelde lijst van genen ook mutatie-, en controlemonsters van elkaar konden worden onderscheiden op basis van data uit onafhankelijke, gepubliceerde data door andere groepen die tevens andere celtypen en manieren om DNA-schade toe te brengen gebruikten.

Kortom, de studies gebundeld in dit proefschrift hebben geleid tot de identificatie en karakterisering van verschillende *BRCA1/2* pathogene varianten. Verder werd gevonden dat een polymorfisme in het *FGFR2*-gen effect heeft op de lokalisatie van een tumor in *BRCA1/2*-dragers. Hoe dit precies werkt, in combinatie met andere risico verhogende of beschermende factoren, is momenteel onduidelijk. Meer inzicht hierin is cruciaal voordat de genetische testen hiervoor toegepast kunnen worden in de klinische praktijk. Tevens, werd een klassificatie methode ontwikkeld op basis van gen-expressie data, die mogelijk

gebruikt kan worden om de klinische relevantie van *BRCA1* VUS te bepalen. De resultaten met onafhankelijke datasets waren veelbelovend. Verdere studies zijn nodig om de methode verder te verfijnen en de reproduceerbaarheid te vergroten. De studies in dit proefschrift hebben in ieder geval bijgedragen aan het nog lange en uitdagende traject op weg naar gepersonaliseerde risicoschattingen voor *BRCA1/2*-mutatiedragers.

RESUMO

O cancro da mama pode ser esporádico ou hereditário. No caso de cancro da mama hereditário há, geralmente, vários elementos da família diagnosticados com tumores – essencialmente mamários, mas também do ovário, pâncreas, estômago, colon, e outros. Famílias com cancro da mama hereditário podem receber aconselhamento e rastreio genético. O rastreio genético permite identificar a mutação causadora da doença, a qual pode ser transmitida de geração em geração. O desafio atual do aconselhamento genético a pessoas com risco elevado de cancro da mama e/ou ovário é estimar e gerir o risco individual. Para que tal seja possível, há que primeiro conseguir identificar as causas genéticas da doença para todas as famílias com síndrome hereditário de cancro da mama e/ou ovário, o que ainda não é possível. Durante o rastreio genético dos genes *BRCA1* e *BRCA2*, são identificadas várias variantes de significado clínico desconhecido. É necessário investigar a relevância clínica destas variantes para um correto aconselhamento dos indivíduos portadores destas variantes. Adicionalmente, devem ser integrados fatores de risco para além das mutações nos genes *BRCA1/2* na avaliação do risco de cancro para permitir estimar com maior precisão os riscos de cada indivíduo. Esta tese consiste em vários estudos efetuados com o objetivo de melhorar a avaliação do risco de indivíduos de famílias com síndrome de cancro da mama e/ou ovário com variantes genéticas nos genes *BRCA1/2*.

O **capítulo 2** desta tese descreve as mutações identificadas durante o rastreio genético de famílias Portuguesas, selecionadas na Consulta de Risco de Cancro da Mama e Ovário do Instituto Português de Oncologia Francisco Gentil de Lisboa, E.P.E.. O trabalho descreve com maior detalhe uma mutação fundadora Portuguesa que foi identificada em 8% (17/210) das famílias rastreadas e que se estimou ter surgido há alguns séculos atrás. O rastreio desta mutação antes do rastreio genético das regiões codificadoras dos genes *BRCA1/2* é mais rápido e barato.

Indivíduos com a mesma mutação, sendo ou não da mesma família, podem desenvolver cancro da mama ou do ovário e, embora haja uma certa tendência em cada família, actualmente não é possível prever a localização da sua ocorrência. Vários estudos sugeriram a existência de fatores genéticos adicionais que modificam o risco de cancro dos indivíduos com mutação num dos genes *BRCA*. Como os órgãos mais afetados por cancro na presença de mutações nos genes *BRCA1/2* estão envolvidos no metabolismo de hormonas esteróides, colocámos a hipótese de que polimorfismos em genes envolvidos no metabolismo das hormonas esteróides poderiam influenciar o local onde o cancro se desenvolve. O estudo-piloto descrito no **capítulo 3.1** incluiu mulheres de famílias com mutações nos genes *BRCA1* ou *BRCA2* aconselhadas no Hospital de Maastricht. Foram analisados Polimorphismos em dois genes, o recetor da progesterona (*PR*) e um activador do PR, *FGFR2*. Adicionalmente, Foram também investigados três polimorfismos nos genes *TNRC9* (*TOX3*) e *CASP8*, que influenciam o risco de cancro da mama esporádico e familiar. Os resultados obtidos demonstram que as mulheres portadoras do alelo de menor

frequência do gene *FGFR2* (rs2981582) têm menor risco de cancro do ovário, embora tenham maior risco de cancro da mama bilateral. O **capítulo 3.2** confirmou, numa população de mais de 20 mil mulheres portadoras de mutações nos genes *BRCA1/2*, que a variante *FGFR2* (rs2981582) protege do cancro do ovário, especialmente nas mulheres portadoras de mutações no gene *BRCA2*.

Durante o rastreio genético, são identificadas várias variantes não classificadas, de relevância clínica ainda indeterminada, que dificultam o correto aconselhamento genético. Os **capítulos 4.1 e 4.2** desta tese descrevem a avaliação e caracterização do efeito de variantes indeterminadas – aquelas que podem ou não ser a causa dos cancros familiares – no mecanismo de excisão (ou slicing), que é um processo importante para o processamento de ARN mensageiro, que ocorre durante a fabricação de proteínas. As variantes que afetam este processo podem dar origem a proteínas defeituosas e daí ser possível concluir que são patogénicas, ou seja, são a causa do cancro familiar. Estes estudos permitiram determinar a patogenicidade de 7 (de 14) variantes, sendo que 4 variantes foram classificadas como patogénicas e 3 como não patogénicas. Três outras variantes afetam o mecanismo de excisão de uma forma que não nos permite ainda saber se pode ser a causa do cancro. As restantes 4 variantes não afetam o mecanismo de excisão, mas como originam alteração de um aminoácido na sequência da proteína, podem ainda afetar a função da proteína. O capítulo **4.2** descreve os resultados obtidos através do uso de um plasmídeo para avaliação do processo de excisão o qual permitiu concluir, sem margem para dúvidas, que a variante c.4987-3C>G é patogénica.

No **capítulo 5** avaliámos a capacidade do uso do *kit* MLPA (Multiplex Ligation Probe Amplification), normalmente comercializado para diagnóstico de delecções e duplicações em ADN, para a deteção de deleções de exões em ARN mensageiro. No entanto, apesar de várias tentativas de otimização, os resultados não foram suficientemente satisfatórios para avaliar defeitos no ARN como teste de rotina.

O **capítulo 6** explorou o transcriptoma de linfócitos de portadores de mutações no gene *BRCA1*, após terem sido sujeitos a cultura e irradiação. Comparando os resultados obtidos com os resultados de linfócitos de pessoas sem mutações nos genes *BRCA1* ou *BRCA2*, foi possível encontrar uma lista de genes que pode ser usada para distinguir portadores de mutações no gene *BRCA1*. Esta observação foi confirmada em amostras de estudos independentes, demonstrando assim a eficácia do método utilizado. É provável que a partir desta experiência seja possível obter um teste robusto para a classificação de variantes indeterminadas como causadoras do cancro ou não. No entanto mais estudos são necessários antes do teste poder ser usado para diagnóstico.

Resumindo, os estudos nesta tese permitiram a identificação e caracterização de várias variantes patogénicas dos genes *BRCA1/2*. Alguns destes estudos contribuíram, assim, para melhorar o aconselhamento dos indivíduos provenientes de famílias com cancro hereditário da mama e do ovário. Outros estudos permitiram melhorar o conhecimento actual e é possível que, no futuro, venham a ser usados com sucesso testes de diagnóstico baseados nos métodos aqui utilizados para classificar variantes indeterminadas.

CURRICULUM VITAE

AND

LIST OF PUBLICATIONS

CURRICULUM VITAE

Rita Dias Brandão was born in Lisbon, Portugal, on 15th November, 1980. She studied in Coimbra and then in Faro, where she completed her secondary school education with grade 16.5 (out of 20).

From 1998 to 2004 she took a five-year degree in Applied Chemistry – Biotechnology (pre-Bologna University course) in the Faculty of Science and Technology of the Universidade Nova de Lisboa (final grade: 14 out of 20). The curricular thesis was on hereditary breast and ovarian cancer syndrome, to which was attributed grade 19 (out of 20). From 2002 to 2006 she conducted research at the Research Centre of Molecular Pathobiology in the Portuguese Oncology Institute of Lisbon Francisco Gentil, where she started as an intern, researching for the aforesaid curricular thesis. The work conducted after the internship was supported by a scholarship for initiation to research from the Science and Technology Foundation.

In 2005 she followed a post-graduation course entitled "Genetics, Genome and Genomic: from Clinic to Public Health" in the National School of Public Health of the Universidade Nova de Lisboa. In 2006 she was one of the 4 researchers awarded with the prize "Investigation in Oncology - NRS/LPCC-Terry Fox for Young Researchers" and she won the prize for best poster presentation in the X Jornadas of Hospitais Cuf. She is also co-inventor of the patent nr. PT103726 (International classification: C12Q 1/68 (2006.01).

She moved to the Netherlands in 2007, after being awarded with a ten-month travel from Calouste Gulbenkian Foundation and a PhD-fellowship from the Science and Technology Foundation to conduct studies at Maastricht University with Dr. ir. Marinus J. Blok and Dr. Encarna Goméz García in hereditary breast and/or ovarian cancer. She has received several travel grants to attend conferences from private, non-profit organizations and GROW − School for Oncology and Developmental Biology from Maastricht University. During these last years she has also collaborated with other colleagues from the departments of Pathology and Internal Medicine of Maastricht University. Additionally, she has been a member of the international consortium ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles) since 2011.

LIST OF PUBLICATIONS

GA Millot, MA Carvalho, SM Caputo, MPG. Vreeswijk, MA Brown, M Webb, E Rouleau, SL Neuhausen, T v O Hansen, A Galli, **RD Brandão**, MJ Blok, A Velkova, FJ Couch, ANA Monteiro, on behalf of the ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles Consortium) Functional Assay Working Group. "A Guide for functional analysis of *BRCA1* variants of uncertain significance (VUS)". Human Mutation, doi: 10.1002/humu.22150

RD Brandão, D Tserpelis, E Gómez García, MJ Blok. "Detection of exon skipping events in *BRCA1* RNA using MLPA kit P002". Molecular Biology Reports 2012; 39 (7): 7429-33

AB Spurdle, S Healey, A Devereau, FB Hogervorst, AN Monteiro, KL Nathanson, P Radice, D Stoppa-Lyonnet, S Tavtigian, B Wappenschmidt, FJ Couch, DE Goldgar; on behalf of ENIGMA (including **R Brandão**). ENIGMA-Evidence-based network for the interpretation of germline mutant alleles: An international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. Human Mutation 2012;33(1):2-7

RD Brandão, K van Roozendaal, D Tserpelis, E. Gómez García, MJ Blok. *"BRCA1* c.4987-3C>G is a pathogenic mutation" Breast Cancer Research and Treatment 2011; 131(2):723-5, Letter to the Editor

RD Brandão, K van Roozendaal, D Tserpelis, E Gómez García, MJ Blok. "Characterisation of unclassified variants in the *BRCA1/2* genes with a putative effect on splicing" Breast Cancer Research and Treatment 2011, 129(3):971-82

JJ van Harssel, CE van Roozendaal, Y Detisch, **RD Brandão**, AD Paulussen, M Zeegers, MJ Blok, EB Gómez García. "Efficiency of BRCAPRO and Myriad II mutation probability thresholds versus cancer history criteria alone for *BRCA1/2* mutation detection". Familial Cancer 2010, 9(2):193-201

NY Souren, AD Paulussen, A Steyls, RJ Loos, **RD Brandão**, M Gielen, HJ Smeets, G Beunen, R Fagard, C Derom, R Vlietinck, JP Geraedts, MP Zeegers. "Parent-of-origin specific linkage and association of the IGF2 gene region with birth weight and adult metabolic risk factors" International Journal of Obesity 2009, 33(9):962-70

A Romano, M Baars, H Martens, **R Brandão**, Y Detisch, E Jongen, MJ Blok, P Lindsey, D-C Fischer, EB Gomez Garcia. "Impact of two functional *progesterone receptor*

polymorphisms (PRP): +331G/A and PROGINS on the cancer risks in familial breast/ovarian cancer". The Open Cancer Journal, 2007, 1(1):1-8

PM Machado, BM Cavaco, **R Brandão**, J Eugenio, S Santos, A Opinião, F Vaz. "In Reply". Journal of Clinical Oncology 2007;25(31):5036-8. Answer to a Letter to the Editor

FH Vaz, PM Machado, **RD Brandão**, CT Laranjeira, JS Eugénio, AH Fernandes, SP André. "Familial breast/ovarian cancer and BRCA1/2 genetic screening: the role of immunohistochemistry as an additional method in the selection of patients" The Journal of Histochemistry and Cytochemistry 2007, 55(11):1105-13

P Machado*, **RD Brandão***, BM. Cavaco, J Eugenio, S Bento, M Nave, A Fernandes, F Vaz. "Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes" Journal of Clinical Oncology 2007, 25(15):2027-34, *- both authors contributed equally for this work

Articles submitted or in preparation

RD Brandão*, J Veeck*, K.K. Van de Vijver, P. Lindsey, B. de Vries, C.H.M.J. van Elssen, M.J. Blok, K. Keymeulen, J. Allemeersch, T. Ayoubi, H.J.M. Smeets, V.C. Tjan-Heijnen, P.S. Hupperets. "A randomised controlled phase II trial of pre-operative celecoxib treatment reveals anti-tumour transcriptional response in primary breast cancer". Submitted, *- both authors contributed equally for this work

RD Brandão, L Eijssen, D Tserpelis, DGAJ Hebels, HJM Smeets, E Gómez García, MJ Blok. "Exploring transcriptional changes in IL2/PHA stimulated *BRCA1*+/- lymphocytes after gamma-irradiation to construct a robust classifier". Submitted

RD Brandão, MJ Blok, AC Antonis, D Barrowdale, G Chenevix-Trench, FJ Couch, EB Gómez García on behalf of the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). "*FGFR2* is protective for ovarian cancer in *BRCA1/2*-mutation carriers". In preparation

ACKOWLEDGMENTS

This thesis would not have been possible without the support of many people. And after all these years, it is with pleasure that I write these words of thanks and appreciation.

I would like to start by thanking my co-supervisors Encarna and Rien for their trust, faith and ever lasting support to me and the project during these past years. Encarna and Rien, definitely it would not have been the same without your guidance and encouragement! Since, you have other obligations, very often you had to read my manuscripts/projects in the evenings and weekends. I am truly indebted and thankful to you.

I am also very grateful to Bert Smeets for all the support provided to enable me to conduct my research and finalize my thesis.

Demis was a great help for my project. Without his precious help I would still be culturing and pipetting. Hartelijke bedankt, Demis! Kees, thank you for all the constructive discussions we had. They have contributed to this thesis too. Roel and Carolien, from the breast cancer group, also helped me in some experiments and around the lab. They were always available to help me and I thank them for that!

Anja, Alexandra and Eveline supported me with the pyrosequencing technology and Kim prepared several DNA plates, which made my work so much easier. I am also thankful for their contribution. Aiad and Frank, thanks for being always available when I needed to deliver samples for sequencing. I know that sometimes I arrived there at break time or too late in the afternoon… Sorry for that! Bieke, thank you very much for hybridizing the microarrays and the discussions about the protocol. I also thank Torik Ayoubi and Prof. dr. Brad Wouters for the helpful discussions we had about the preparation of samples for the microarrays study. I want to thank also to Andrea, Aimee, Roselie, Arthur, Marij, Bart, Jos, and Marion for the scientific discussions during my PhD time.

Fons, thanks for keeping online my isoforms website each time I update it and providing the necessary passwords. Jo, thanks for saving my computer (over and over and over) from being thrown out of the window. Indirectly this has also saved my mental sanity! The discussions about cloning with Guillaume and Barbie and those about immunostaining with Erika and Willem Voncken were very useful for my experiments. Iris and Rick, thank you both for the help with the preparation of my future experiment. Rosy, I appreciate very much the help you gave me on these last months to arrange everything for the final phase of my PhD! Sometimes things go wrong… Bianca once had to help me finding the long oligos (with modifications and special purification – really expensive ones) in the garbage!! Thanks Bianca!

Ellen, Wanwisa, Sabine, and recently arrived Jo and Geert, although we didn't have the opportunity to work together, I thank you for the "gezelligheid"! I also would like to thank those who helped me in one or another way, such as finding things around or reaching shelves.

To help me with the bioinformatics analysis of my microarray study I had the help of Lars Eijssen from the BigCat department. The discussions we had about the elaboration

of a genetic classifier were very important. Lars, thank you (also for the nice moments we shared outside work)!

Maurice, thank you very much for all the analysis that you had to do for our project. Thank for doing this in the weekends and evening. I won't forget that once, in a very important day for you, we had an appointment in the evening at Café Zuid!!

During my PhD time I also had some students who helped me with some experiments. These were Hakima Bari, Chris Wijnen, and Laura van Zon. They have made their best when they were around and I am very grateful for that! Wish you all lots of success!

Before moving to Maastricht I had already worked on hereditary breast and/or ovarian cancer syndrome in Lisbon. So I would like to thank my colleagues from the Centro de Investigação de Patobiologia Molecular (CIPM) for their help around the lab. I specially would like to thank Dra. Fátima Vaz, who provided me the opportunity to start working on the inherited breast cancer topic. It was thanks to her that I developed my enthusiasm on this topic.

During these years I had the opportunity to work on some other projects that are not part of this thesis. I greatly appreciate that opportunity, as it allowed me to increase my knowledge. One of these projects was the randomized clinical trial with celecoxib on primary breast cancer patients. During this project I worked with Pierre Hupperets, Jürgen Veeck, Koen Van de Vijver, and Patrick Lindsey. Their enthusiasm and dedication to this study made them a very good team to work with and it was a great pleasure to be part of the study.

But life is not only about work! And I was very lucky to have the opportunity to work in a very nice department that even organizes a "borrel" at the end of each month, carnival parties, etc! It was "gezellig"!!

And of course that I would like to thank the people in the PhD-students room who provided me with very funny moments during the past years. I won't forget the "uga-uga" in the lab, used to pray to the pyrosequencing "gods", that I and Nicole performed to get our pyrosequencing experiments working. Florence, thank you for sharing protocols, a coffee mug, and the toaster! I have to say that I really missed Ruben's sense of humour, when he left us to go to USA. I thank him for the real Indonesian "sate" recipe, his cocktails, his honesty, even the fun times at badminton (not so fun was to be used as a target, which even happened when we played in the same team!), and, of course, his friendship. Auke brought to the room his hyperactive attitude and humour. He thought of keeping "Double Bliss" bottles so that we could play bowling. That was great! Don't want to sound cocky, but I was actually much better than at real bowling. I was actually winning to Auke and Mike! ☺ Besides bowling in the department, Mike was also a great company for fitness. I am so happy for that, since fitness would otherwise have been so boring that I would have probably just learned to live with my shoulder injury. But with you it was fun! Abishek, it was a pity that you decided to change your PhD project. Your Indian accent would make everything sound even funnier! Wish you good luck with your new project.

Minh, Tom, and Romy have recently joined the room and I thank them for the fun moments that we've already shared and wish them lots of success!

As I mentioned before, life is not only about work. And definitely it is not only within the Maastricht University walls! So, I would like to express my gratitude to all those who made my stay in Maastricht so much fun! In times, they were very important to distract me from work or to boost my motivation. Erik, thanks for the badminton games, dinners, cinema nights, BBQs, and drinks. Mirjam, thank you for your kindness, friendship, and advices! Good luck in improving your English skills in Ireland (irony. blink blink)! Nuria, your company at capoeira, acrobatics, badminton, and parties was missed. Wish you nothing but the best wherever you go! Hyperactive Esther also organized nice cocktail evenings!! Thanks Nuria, Mirjam and Esther for the advices on the wedding dress! You were good "sisters"! Raffaele, your contagious good mood was welcome. Even if it doesn't solve my problems, then I can forget about them (at least for a while)! The lunches you organized on Sundays really helped going through the long and cold winters in Maastricht. Tsilla, thank you for organizing nice parties, trips, and for the many stories you have to tell, and thank you both for the company to the salsa evenings! Bea and Hilde, you were both always very busy. Still, we had some fun moments during dinners and going-outs! That was nice. Yeliz, thanks for my first Turkish meal and for showing me the way to the best Turkish pizza in Maastricht! Alejandro, thank you for all the fun moments and for organizing the dinners at the Blind restaurant and in Budapeste! Ariane, you were only in Maastricht for a short period of time, but I remember very well your friendly and cheerful attitude. Chiara, you're too nice!! More than once you woke up before sunrise to be able to bring us the extremely delicious chocolate "bigne" from Florence!! Mo, you must stay around Maastricht. There's a comedy night coming up soon and we need to book all their tickets even before we send an email around. Oh no! It seems they didn't like that so much, because now they have a limit and we have to pay them in advance!! But we should go anyway! And when you're away and you finally get your brevet, you just come on you private jet! Maria (or Mária, as Raf calls you), nice that you joined for the trip to the castles along the Rhine river. It was very funny! Mª João, obrigada pela tua amizade e companheirismo. Diverti-me imenso e sinto a falta dos nossos almoços, cafés, de irmos para a relva apanhar um solinho enquanto fazemos um reality-check, etc. Ah, e ainda bem que não sou a única que não se lembra das palavras em português e que mistura umas palavras de inglês pelo meio das frases. I have to thank Mª João and Judith for the nice picture from my brain! It was worth to be for 2h in the MRI! Zoe and Joost, thank you for all the cheerful moments during our coffees, drinks/cocktails, and shows. Lyzel, you're always so busy travelling through Holland everyday, doing your research, and then you still have to see patients… Luckily, once in a while you could find some to meet me for catching up!

A mudança para Maastricht afastou-me dos amigos que tinha em Portugal. Embora os veja menos, eles continuam presentes! Alguns tiveram oportunidade de me

virem visitar. Outros fui eu visitar. Encontramo-nos quando é possível e é como se nunca nos tivéssemos separado. Todos eles acabaram por me ajudar em alguns momentos da minha vida. Portanto ficam aqui beijinhos grandes para a Marta e Tó de Coimbra; Susana, Sílvia, Pedro, Sabu e Rita de Faro. Durante e depois da universidade tive o apoio e companhia para festas de muita gente também: Tânia, Sofia, David, Verinha (que ficou muitas noitinhas comigo no CIPM a ajudar-me), Susaninha, João, Leo, Xana, Sónia C., Ruben, Sónia B., Elsa D., Ricardo, Rui e Joana, Dora, David, Tânia, Sofia, Sandra, Vald, Rui Dias e tantos outros!! Abraços grandes malta!

Durante a minha estadia em Maastricht ganhei nova família. E os meus sogros e cunhados também sempre me apoiaram, pelo que lhes estou muito grata pelo carinho dado, bem como pelas iguarias portuguesas que ofereciam para poder ter nos Países Baixos alguns sabores de Portugal.

Pai, mãe, obrigada por todo o apoio que me deram durante este periodo. Nem sempre foi fácil, mas vocês estavam sempre disponíveis para me ouvir independentemente da distância. O vosso apoio e carinho não vem só de agora, é claro. Estou-vos grata por todo o apoio que tive durante toda a minha vida e lamento até as discussões que tivémos quando estudar não fazia parte dos meus planos. Agora que estou prestes a ser mãe, só desejo poder dar aos meus filhos as mesmas condições que me deram a mim. Beijos grandes!

Agradeço também à minha mana Susana a força e incentivo que me deu para prosseguir o doutoramento no estrangeiro e ao meu mano Pedro o apoio que me prestou durante este período, em especial quando tive de organizar festas!

Os amigos contribuiram bastante para a minha felicidade e conforto em Maastricht. Ainda assim, não teria sido o mesmo se não tivésse o Luís por perto. Luís, agradeço-te por te teres mudado para a Holanda (que sorte que foi teres encontrado emprego na Bélgica!), pelo apoio constante e incondicional que me dás, por me obrigares a "desligar" do trabalho e pela companhia para as aventuras por este mundo fora.

Rita