Identification of novel \textit{BRCA1} large genomic rearrangements by a computational algorithm of amplicon-based Next-Generation Sequencing data

Arianna Nicolussi $^1$, Francesca Belardinilli $^2$, Valentina Silvestri $^2$, Yasaman Mahdavian $^2$, Virginia Valentini $^2$, Sonia D’Inzeo $^3$, Marialaura Petroni $^4$, Massimo Zani $^3$, Sergio Ferraro $^2$, Stefano Di Giulio $^2$, Francesca Fabretti $^2$, Beatrice Fratini $^1$, Angela Gradilone $^2$, Laura Ottini $^1$, Giuseppe Giannini $^2$, Anna Coppa $^1$, Carlo Capalbo $^2$

$^1$ Department of Experimental Medicine, University of Roma “La Sapienza”, Roma, Italy
$^2$ Department of Molecular Medicine, University of Roma “La Sapienza”, Roma, Italy
$^3$ U.O.C. Microbiology and Virology Laboratory, A.O. San Camillo Forlanini, Roma, Italy
$^4$ Istituto Italiano di Tecnologia, Center for Life Nano Science@Sapienza, Roma, Italy
$^5$ Istituto Pasteur-Fondazione Cenci Bolognetti, Roma, Italy

Corresponding Authors: Giuseppe Giannini, Anna Coppa
Email address: giuseppe.giannini@uniroma1.it, anna.coppa@uniroma1.it

\textbf{Background}: Genetic testing for \textit{BRCA1/2} germline mutations in hereditary breast/ovarian cancer patients requires screening for single nucleotide variants, small insertions/deletions and large genomic rearrangements (LGRs). These studies have long been run by Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA). The recent introduction of next-generation sequencing (NGS) platforms dramatically improved the speed and the efficiency of DNA testing for nucleotide variants, while the possibility to correctly detect LGRs by this mean is still debated. The purpose of this study was to establish whether and to which extent the development of an analytical algorithm could help us translating NGS sequencing via an Ion Torrent PGM platform into a tool suitable to identify LGRs in hereditary breast-ovarian cancer patients. \textbf{Methods}: We first used NGS data of a group of 3 patients (training set), previously screened in our laboratory by conventional methods, to develop an algorithm for the calculation of the dosage quotient (DQ) to be compared with the Ion Reporter (IR) analysis. Then, we tested the optimized pipeline with a consecutive cohort of 85 uncharacterized probands (validation set) also subjected to MLPA analysis. Characterization of the breakpoints of three novel \textit{BRCA1} LGRs was obtained via long-range PCR and direct sequencing of the DNA products. \textbf{Results}: In our cohort, the newly defined DQ-based algorithm detected 3/3 \textit{BRCA1} LGRs, demonstrating 100% sensitivity and 100% negative predictive value (NPV) [95\% CI:87.6-99.9]) compared to 2/3 cases detected by IR (66.7\% sensitivity and 98.2\% NPV [95\% CI:85.6-99.9]). Interestingly, DQ and IR shared 12 positive results, but exons deletion calls matched only in 5 cases, two of which confirmed by MLPA. The breakpoints of the 3
novel BRCA1 deletions, involving exons 16-17, 21-22 and 20, have been characterized. **Conclusions:** Our study defined a DQ-based algorithm to identify BRCA1 LGRs using NGS data. Whether confirmed on larger data sets, this tool could guide the selection of samples to be subjected to MLPA analysis, leading to significant savings in time and money.
Identification of novel BRCA1 large genomic rearrangements by a computational algorithm of amplicon-based Next-Generation Sequencing data

Arianna Nicolussi1; Francesca Belardinilli2; Valentina Silvestri2; Yasaman Mahdavian2; Virginia Valentini2; Sonia D’Inzeo3; Marialaura Petroni4; Massimo Zani2; Sergio Ferraro2; Stefano Di Giulio2; Francesca Fabretti2; Beatrice Fratini1; Angela Gradilone2; Laura Ottini2; Giuseppe Giannini2,5; Anna Coppa1* and Carlo Capalbo2*

1 Department of Experimental Medicine, University of Roma “La Sapienza”, Italy
2 Department of Molecular Medicine, University of Roma “La Sapienza”, Italy
3 A.O. San Camillo Forlanini, U.O.C. Microbiology and Virology Laboratory, Roma, Italy
4 Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, Roma, Italy
5 Istituto Pasteur-Fondazione Cenci Bolognetti, Roma, Italy

* These authors equally contributed to this work

Corresponding authors:
Dr. Anna Coppa, Department of Experimental Medicine, University of Roma “La Sapienza”, Italy. Email address: anna.coppa@uniroma1.it
Prof. Giuseppe Giannini, Department of Molecular Medicine, University of Roma “La Sapienza”, Italy Email address: giuseppe.giannini@uniroma1.it
Abstract

Background: Genetic testing for BRCA1/2 germline mutations in hereditary breast/ovarian cancer patients requires screening for single nucleotide variants, small insertions/deletions and large genomic rearrangements (LGRs). These studies have long been run by Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA). The recent introduction of next-generation sequencing (NGS) platforms dramatically improved the speed and the efficiency of DNA testing for nucleotide variants, while the possibility to correctly detect LGRs by this mean is still debated. The purpose of this study was to establish whether and to which extent the development of an analytical algorithm could help us translating NGS sequencing via an Ion Torrent PGM platform into a tool suitable to identify LGRs in hereditary breast-ovarian cancer patients.

Methods: We first used NGS data of a group of 3 patients (training set), previously screened in our laboratory by conventional methods, to develop an algorithm for the calculation of the dosage quotient (DQ) to be compared with the Ion Reporter (IR) analysis. Then, we tested the optimized pipeline with a consecutive cohort of 85 uncharacterized probands (validation set) also subjected to MLPA analysis. Characterization of the breakpoints of three novel BRCA1 LGRs was obtained via long-range PCR and direct sequencing of the DNA products.

Results: In our cohort, the newly defined DQ-based algorithm detected 3/3 BRCA1 LGRs, demonstrating 100% sensitivity and 100% negative predictive value (NPV) [95% CI:87.6-99.9]) compared to 2/3 cases detected by IR (66.7% sensitivity and 98.2% NPV [95% CI:85.6-99.9]). Interestingly, DQ and IR shared 12 positive results, but exons deletion calls matched only in 5
cases, two of which confirmed by MLPA. The breakpoints of the 3 novel \textit{BRCA1} deletions, involving exons 16-17, 21-22 and 20, have been characterized.

**Conclusions:** Our study defined a DQ-based algorithm to identify \textit{BRCA1} LGRs using NGS data. Whether confirmed on larger data sets, this tool could guide the selection of samples to be subjected to MLPA analysis, leading to significant savings in time and money.

**Background**

Hereditary breast and ovarian cancer syndrome, caused by germline pathogenic mutations in the \textit{BRCA1} (MIM#113705) or \textit{BRCA2} (MIM#600185) genes, is characterized by an increased risk for breast, ovarian, pancreatic and other cancers (Palma et al., 2006). It has been recently estimated that the cumulative risks of breast cancer to age 80 years was 72% for \textit{BRCA1} and 69% for \textit{BRCA2} carriers (Kuchenbaecker et al., 2017). Differences in mutation type and site may at least partially impact on cancer risk definition (Rebbeck et al., 2015; Coppa et al., 2018; Rebbeck et al., 2018). \textit{BRCA1} and \textit{BRCA2} gene mutations are typically found in 25–30% of the breast cancer families subjected to genetic testing (Giannini et al., 2006; Economopoulou, Dimitriadis & Psyrri, 2015). The relatively low rate of success in finding relevant pathogenic mutations in this settings is likely due to the contribution of other moderate-to-high penetrance breast cancer susceptibility genes (i.e., \textit{PALB2}, \textit{ATM}, \textit{CHK2}) (Economopoulou, Dimitriadis & Psyrri, 2015; Coppa et al., 2018), or to the influence of low penetrance and risk-modifying alleles (Couch et al., 2012; Ottini et al., 2013; Kuchenbaecker et al., 2014; Peterlongo et al., 2015), all of which needs to be taken into account for a more appropriate assessment of individual cancer risk. For quite some time, the use of classical qualitative PCR-based techniques incapable of detecting large genomic rearrangements (LGRs) also contributed to failures in the
identification of BRCA mutation carriers. Interestingly, the prevalence of BRCA1/BRCA2 LGRs varies greatly among different populations ranging from 0 to 27% of mutation positive families in Iranian/French, Canadian, Dutch, Spanish, German, French and South Africa populations (Gad et al., 2002; Hogervorst et al., 2003; Hartmann et al., 2004; Pietschmann et al., 2005; Moisan et al., 2006; la Hoya et al., 2006; Sluiter & van Rensburg, 2011). Relevant differences in the frequency of BRCA1 LGRs have also been reported within the Italian population (Montagna et al., 2003; Buffone et al., 2007). In general, BRCA2 LGRs are less frequent (Woodward et al., 2005; Agata et al., 2005; Buffone et al., 2007), probably due to the lower density of Alu sequences compared to BRCA1, which are involved in the genesis of LGRs (Smith et al., 1996).

Multiplex ligation-dependent probe amplification (MLPA) is the most commonly used technique for the detection of large deletions/duplications in BRCA1/2 genes. The recent advances in sequencing technologies have increased the speed and efficiency of DNA testing and the emergence of benchtop next-generation sequencing (NGS) instruments are becoming the standard in molecular genetic diagnosis (Feliubadalo et al., 2013; Trujillano et al., 2015). NGS is capable of sensitive detection of sequence variants, but may also be used for detection of LGRs by the evaluation of Copy Number Variations (CNVs) (Tarabeux et al., 2014; Enyedi et al., 2016; Schenkel et al., 2016; Schmidt et al., 2017). The CNVs assessment is mainly performed using the sequencing read depth (RD) assessment approach, whose assumption is that the RD signal is proportional to the number of copies of chromosomal segments present in that specimen (Tan et al., 2014). The ability to detect CNVs from NGS multigene panel largely, but not uniquely, depends on the library preparation, and target enrichment approaches based on hybridization and capture seem to have better performances compared to amplicon-based methods. In general, NGS data are not routinely used for CNVs detection in clinical settings for...
BRCA mutation screenings, due to concerns related to library preparation protocols, normalization procedures and employed software (Feliubadalo et al., 2013; Wallace, 2016).

Recently, we adopted the NGS Ion AmpliSeq™ BRCA1 and BRCA2 Panel to perform routine BRCA1/2 mutation screening on the Ion PGM platform (Nicolussi et al., 2019). Here, we aimed at establishing whether sequencing data generated by this approach could be processed by a computational algorithm to efficiently predict the presence of LGRs, based on the dosage quotient (DQ) calculation and the Ion Reporter (IR) analysis.

**Methods**

**Patients and DNA**

Families putatively affected by hereditary breast/ovarian cancer syndrome were recruited at the Hereditary Tumors section of Policlinico Umberto I, University La Sapienza, between July 2015 and September 2017 and selected as previously described (Capalbo et al., 2006b,a; Coppa et al., 2014). Comprehensive pre-test counseling was offered to all probands and their family members and informed consent was obtained. For each study participant, samples of blood or DNA from peripheral blood leukocytes were collected. DNA from blood samples was extracted and quantified as described by Nicolussi et al. (Nicolussi et al., 2019). All investigations were approved by Ethics Committee of the University of Roma “La Sapienza” (Prot.: 88/18; RIF.CE:4903, 31-01-2018) and conducted according to the principles outlined in the declaration of Helsinki.

A retrospective group of 3 DNA samples, previously found positive for BRCA1 LGRs by MLPA was used as a training set (TS). LGRs in the TS were as follows: sample BR59, BRCA1 exon 23-24 deletion (c.5407-?*(1_?)del); sample BR328, BRCA1 exon 18-19 deletion (c.5075-
(Buffone et al., 2007) and sample BR409, NBR2 exon1 and BRCA1 exon 1-2 deletion (NBR2del EX1_BRCA1 delEX1-2) (Coppa et al., 2018) (Table 1).

For NGS-based LGR analysis, a consecutive group of 127 NGS/MPLA negative samples have been used to create a baseline and a prospective consecutive cohort of 85 uncharacterized probands, validation set (VS), was studied.

119 **Ion Torrent PGM sequencing**

The target regions in the *BRCA1* and *BRCA2* genes were amplified using the Ion AmpliSeq™ BRCA1 and BRCA2 Panel (Life Technologies) according to the manufacturers’ procedures and processed as previously described (Belardinilli et al., 2015; Nicolussi et al., 2019).

124 **Sanger sequencing**

All clinical samples were sequenced for the entire coding regions by Sanger sequencing, using an ABI PRISM DyeDeoxy Terminator Cycle Sequencing Kit and an ABI 3100 Genetic Analyzer (Applied Biosystems, Warrington, UK). Reference sequence for *BRCA1* was Genbank, NM_007294.3, and reference sequence for *BRCA2* was Genebank, NM_000059.3.

130 **MLPA analyses**

MLPA methodology (Schouten et al., 2002) was performed, according to the manufacturer’s instructions (MRC–Holland, Amsterdam, the Netherlands), to identify *BRCA1/2* genomic rearrangements. For the statistical analysis we transferred the size and the peak areas of each sample to an Excel file. The peak areas of the expected MLPA products were evaluated by
comparison with a normal control and by cumulative comparison of all samples within the same experiment (Buffone et al., 2007; Coppa et al., 2018).

NGS-based LGRs Analysis

LGRs in BRCA1 gene were studied by two distinct approaches: the manual calculation of the DQ and the IR platform. In the manual approach, DQ for each sample was calculated as follows:

- Amplicon read count normalized on the BRCA1 and BRCA2 total reads/average of normalized amplicon read counts obtained from all samples. Specifically, we referred to DQA when amplicon counts were normalized vs the coverage data of all samples run on the same single chip, and to DQB when amplicon counts were normalized vs coverage data obtained from a baseline built from 127 LGRs negative samples. In addition, DQB has been alternatively obtained either considering together all amplicons of the Ion AmpliSeq™ BRCA1 and BRCA2 Panel (DQB1) or by separately considering the three different pools of amplicons (DQB2). DQ value higher than mean plus 2 standard deviations (SD) was considered indicative of a duplication; DQ value lower than mean minus 2 SD was considered indicative of a deletion.

- Particular attention has been also payed to reduction of multiple consecutive amplicons, even when they failed to trespass the above defined thresholds.

In the IR approach, we create a user-defined CNV detection workflow by a tunable Ion Reporter™ Software algorithm based on Hidden Markov Model (HMM), that utilize normalized read coverage across amplicons to predict the copy number or ploidy (https://assets.thermofisher.com/TFS-Assets/LSG/brochures/CNV-Detection-by-Ion.pdf). The data coverage of 20 mutation-negative patients has been used as CNV baseline to analyze the
samples of both TS and VS. We detected no BRCA2 LGR in both the TS and VS. Thus, our
analysis is necessarily limited to BRCA1 LGRs.

DNA breakpoint analysis

Newly discovered BRCA1 large deletions were validated by characterization of the genomic
breakpoints. Long-range PCR was performed according to the manufacturer’s instructions using
the kit Platinum Taq DNA polymerases High Fidelity (Thermo Fisher) with the primers sitting
on closer undeleted exons as described in Table S2. PCR products were purified with ExoSAP-
IT (USB Corp., Cleveland, USA) according to the manufacturer’s instructions and sequenced
using the ABI PRISM DyeDeoxy Terminator Cycle Sequencing Kit and an ABI 3100 Genetic
Analyzer (Applied Biosystems, Warrington, UK). Reference sequences for BRCA1 and BRCA2
were Genebank NM_007294.3 and NM_000059.3, respectively.

Statistical analysis

Validation metrics were defined as: Accuracy = (TP + TN)/(TP + FP + TN + FN); Sensitivity =
TP/(TP + FN); Specificity = TN/(TN+FP); FDR = FP/(TP + FP); Negative Predictive Value =
TN/(TN + FN), where TP = true positives, TN = true negatives, FP = false positives, FN =
false negatives. The confidence intervals (CIs) were calculated by the method of Wilson (1927)
(EB, 2019).

Results

NGS-dependent LGR analyses

To establish whether the data obtained by NGS via Ion AmpliSeq™ BRCA1 and BRCA2 Panel
were suitable to identify copy number alterations in *BRCA1*, we used data from three samples (TS), already characterized in our laboratory for the presence of *BRCA1* LGRs by MLPA (Table 1). The sequencing data of the TS were analyzed by a locally devised algorithm for the calculation of the DQ and by our custom modified IR analysis, as described in materials and methods. The intrarun DQ calculation (DQA), which includes normalization based on the coverage data of the samples sequenced in the same chip, was always included to monitor the variability eventually due to different batches of reagents or to time-related variables. In general, however, we thought we could get improved resolution and reduced numbers of CNV false calls by normalizing the coverage data of all amplicons of each sample vs those obtained from a reference set of 127 MLPA negative samples selected on the basis of their quality and uniformity of the coverage (DQB analysis). This baseline has been used to perform two DQB calculations, considering either all amplicons contained in the Ion AmpliSeq™ BRCA1 and BRCA2 Panel (DQB1) or dividing them into the three subsets identified by the amplification primer pools (DQB2).

As shown in Fig 1A, the DQA plot of the TS samples revealed the presence of peaks below the thresholds, in samples BR328 and BR409 (corresponding to deletions of *BRCA1* exons 18-19 and 1-2, respectively, in agreement with MLPA results). The DNA quality of BR59 sample was rather low, as evidenced by the many peaks out of the threshold. Nevertheless, the DQB1 analysis evidenced values below the threshold for 3 consecutive amplicons (AMPL223551867, 223530147 and 223954665), identifying *BRCA1* exon 23-24 deletion (Fig.1B), already discovered by MLPA analysis. Although they fail to trespass the threshold, the same consecutive amplicons showed strongly reduced values also at DQA evaluation (Fig. 1A). Hence, the careful examination of the two DQ calculations allowed us to identify all three *BRCA1* LGRs in the TS.
Also, the analysis performed by IR software detected the presence of CNV (CNV=1) in the proper regions in all three TS samples (Table 2). On this basis, we extended DQA, DQB and IR analysis to a group of 85 consecutive samples (VS) negative for BRCA1/2 pathogenic variants at NGS analysis and compared it with MLPA results. Overall, DQA and DQB analysis resulted in detection of positive calls in 33/85 (39%) samples, while IR analysis detected CNVs in 29/85 (34%) (Table 3). Interestingly, DQ and IR evaluation only shared 12 positive results, with exon calls being not coincident in 7 of them and with a rather precise, although imperfect, indication of the exons involved in the remaining 5 (Table S1). MLPA confirmed BRCA1 LGRs in 3/85 samples (Fig. 2): BR963 and BR1379, belonging to the small group of 5 DQ/IR double positive samples, and BR1154 resulted DQ positive-IR negative. Therefore, DQ calculation resulted 100% sensitive and displayed a 100% NPV (95% CI:87.6-99.9) (Table 3) in our VS, values not reached by IR analysis, which failed in the identification of BR1154 (Table 2). Within DQ analysis, the correct calls were more clearly defined by the DQB2 calculation (Fig. 3A, B, C). The appropriateness of the deletions calls of DQ, IR and MLPA evaluations were confirmed by the molecular characterization of the breakpoints, as described below.

Characterization of LGRs

Identification of the breakpoints characterizing the LGRs is important for several reasons, including the possibility to develop diagnostic assays for segregation analyses in relatives. For different reasons DQ, IR and MLPA analyses are not able to provide such detailed molecular characterization of LGR. To define the breakpoints of the newly identified BRCA1 LGRs, PCR amplification of genomic DNA from the three samples and direct sequencing were performed.
As shown in Fig. 4A, PCR amplification of genomic DNA from the BR963 patient resulted in an aberrant fragment of approximately 1353 bp, whose direct sequencing confirmed loss of *BRCA1* exons 21 and 22, possibly originating from an erroneous homologous recombination process between an AluSq2 (Alu family, SINE class; chr17:41206762-41207066) and an AluSz (Alu family, SINE class; chr17:4120521-41200834) motifs. The rearrangement involved a perfectly repeated stretch of 24 bases and resulted in the deletion of 6228 nucleotides encompassing part of IVS20, exons 21-22 and IVS22 (Fig. 4B, C). The BR963 proband was affected with breast cancer at age 40 and belonged to HBC family. Segregation analysis demonstrated that the mutation came from the maternal lineage (Fig. 5A). PCR amplification of genomic DNA from BR1154 patient resulted in an aberrant fragment of approximately 872 bp (also present in her mother, sample BR1148), whose direct sequencing confirmed loss of *BRCA1* exons 20, possibly originating from an erroneous homologous recombination process between an AluY (Alu family, SINE class; chr17:41205398-41205698) and an AluY (Alu family, SINE class; chr17:41205398-41205698) motifs. The rearrangement involved a perfectly repeated stretch of 11 bases and resulted in the deletion of 4173 nucleotides encompassing part of IVS19, exon 20 and IVS20 (Fig. 4D, E, F). The BR1154 proband was affected with ovarian cancer at age 52 and belonged to a HBOC family (Fig. 5B). The segregation analysis demonstrated that the mutation originating from the maternal lineage segregated in three individuals (Fig. 5B). Finally, PCR amplification of genomic DNA from BR1379 patient, resulted in an aberrant fragment of approximately 2027 bp, whose direct sequencing confirmed loss of *BRCA1* exons 16 and 17, possibly originated from an erroneous homologous recombination process between an AluSp (Alu family, SINE class; chr17:41224585-41224884) and an AluSg (Alu family, SINE class; chr17:41218424-41218724) motif. The rearrangement involved a perfectly repeated stretch of 16 bases and resulted in the
deletion of 6155 nucleotides encompassing part of IVS15, exons 16-17 and IVS17 (Fig. 4G, H, I). The BR1379 proband was affected with bilateral breast cancer at age 42 and 58 and belonged to a family with colon cancer and hepatomas cases (Fig. 5C).

In conclusion, our results in the VS allow us to propose an operative algorithm which uses DQ calculation and IR analysis to select samples to be subjected to MLPA analysis, as indicated in Fig. 6. Indeed, all DQ positive samples should be subjected to MLPA, while DQ and IR double positive samples, sharing calls in the same regions, could be directly subjected to second level confirmation assay or directly to breakpoint characterization. In principle, all DQ negative samples (52 sample out of 85 in our VS) could be considered negative for LGRs, thus completing the analysis at this step.

Discussion

A complete clinical level analysis of BRCA1 and BRCA2 in hereditary breast/ovarian cancer includes the study of LGRs. Many methods have been used to identify LGRs, such as fluorescent in-situ hybridization (FISH) and microarrays (Xia et al., 2018), Southern blot, long-range PCR, quantitative multiplex PCR of short fragments (QMPSF) (Ewald et al., 2009), semiquantitative multiplex PCR, real-time PCR, restriction analysis and sequencing (Armour et al., 2002). All these methods are limited by their low throughput, time consuming, large amounts of high molecular weight DNA request and several false negative results (Ewald et al., 2009). More recently a multiplex PCR-based method that allows the determination of copy number status of multiple loci in a single assay, has been developed by Multiplicom (http://www.multiplicom.com) and described as a valid method (Concolino et al., 2014). However, the MLPA represents the most widely used approach to scan for LGRs in BRCA1/2 genes (Ruiz
The simultaneous detection of mutations and copy number alterations is an attractive and useful prospect for clinical settings. In the last years the NGS-based approaches for genetic testing offered a powerful alternative for BRCA1/2 mutation detection. However, the specificity of this approach is still considered not completely satisfactory for a correct LGRs detection. One of the most relevant aspects concerns the library preparation method, with the amplicon-based approach having a lower specificity compared to target enrichment approaches (Apessos et al., 2018). Here we reported the definition of an operative algorithm to use amplicon-based Ion-PGM/Ampliseq BRCA1/BRCA2 sequencing data to efficiently predict the occurrence of BRCA1 LGRs. By comparison of the results obtained with DQ and IR analyses, we demonstrate that DQ had 100% sensitivity and 100% NPV, at variance with IR analysis, which failed in the identification of a BRCA1 exon 20 deletion. This result is consistent with one known limitation of the IR software, able to detect CNVs only if the region of interest is covered by more than one amplicon (https://assets.thermofisher.com/TFS-Assets/LSG/brochures/CNV-Detection-by-Ion.pdf). Indeed, BRCA1 exon 20, deleted in BR1154 sample, is covered by only one amplicon in the Ion AmpliSeq™ BRCA1 and BRCA2 Panel, making IR incapable of calling this CNV. Of course, a major caveat deals with the limited specificity and accuracy of our approach, which could not overcome the limitations also reported by other groups (Feliubadalo et al., 2013; Pilato et al., 2016). Thus, although our operative algorithm cannot fully substitute for MLPA analysis, and if our data will be confirmed in larger data sets, we suggest that combined DQ and IR analyses could be used for selecting samples to be subjected to MLPA analysis following the flow chart depicted in Fig. 4, with significant savings in time and money. Another important contribution of this paper is the molecular characterization of the three novel BRCA1 rearrangements up to providing their unique breakpoint coordinates. Deletion of exons
21 and 22 causing damage to the C-terminal BRCT domain of the BRCA1 protein has been reported and characterized in Czech (Vasickova et al., 2007; Ticha et al., 2010) and Malay population (Hasmad et al., 2015), but with different breakpoints. BRCA1 exon 20 deletion has been described in Italian and Greek population (Montagna et al., 2003; Belogianni et al., 2004; Armaou et al., 2007) but all different from each other and from our own, with respect to their breakpoints. The BRCA1 exons 16-17 deletion, responsible of BRCA1 loss of function (Carvalho et al., 2009), has been reported in Latin America/Caribbean population, but the breakpoints were not provided by the authors (Judkins et al., 2012). Similar to many other cases (Mazoyer, 2005; Buffone et al., 2007; Ewald et al., 2009), all three novel rearrangements described here, are likely to be due to an erroneous homologous recombination event between perfectly matching Alu repeats.

**Conclusion**

In conclusion, here we described a simple approach that require the use of a basic statistical package such as Microsoft Excel, to predict the occurrence of LGRs by the analysis of NGS data designed for Ion AmpliSeq™ BRCA1 and BRCA2 Panel/IT-PGM platform, applicable to all NGS platforms in use to reduce the number of samples to be subjected to MLPA analysis. We also characterized for the first time the breakpoints of three novel BRCA1 LGRs.

**Acknowledgements**

The authors would like to thank Dr. Alessandro Albiero for advice on data analysis.

**Reference**
1. Agata S, Dalla Palma M, Callegaro M, Scaini MC, Menin C, Ghiotto C, Nicoletto O, Zavagno G, Chieco-Bianchi L, D'Andrea E, Montagna M 2005. Large genomic deletions inactivate the BRCA2 gene in breast cancer families. *Journal of medical genetics* 42:e64. DOI: 10.1136/jmg.2005.032789.

2. Apessos A, Agiannitopoulos K, Pepe G, Tsaousis GN, Papadopoulou E, Metaxa-Mariatou V, Tsirigoti A, Efstathiadou C, Markopoulos C, Xepapadakis G, Venizelos V, Tsiftsoglou A, Natsiopoulos I, Nasioulas G 2018. Comprehensive BRCA mutation analysis in the Greek population. Experience from a single clinical diagnostic center. *Cancer genetics* 220:1–12. DOI: 10.1016/j.cancergen.2017.10.002.

3. Armaou S, Konstantopoulou I, Anagnostopoulos T, Razis E, Boukouinas I, Xenidis N, Fountzilas G, Yannoukakos D 2007. Novel genomic rearrangements in the BRCA1 gene detected in Greek breast/ovarian cancer patients. *European Journal of Cancer* 43:443–453. DOI: 10.1016/j.ejca.2006.09.021.

4. Armour JAL, Barton DE, Cockburn DJ, Taylor GR 2002. The detection of large deletions or duplications in genomic DNA. *Human Mutation* 20:325–337. DOI: 10.1002/humu.10133.

5. Belardinilli F, Capalbo C, Buffone A, Petroni M, Colicchia V, Ferraro S, Zani M, Nicolussi A, D'Inzeo S, Coppa A, Screpanti I, Gulino A, Giannini G 2015. Validation of the Ion Torrent PGM sequencing for the prospective routine molecular diagnostic of colorectal cancer. *Clinical biochemistry* 48:908–910. DOI: 10.1016/j.clinbiochem.2015.04.003.

6. Belogianni I, Apessos A, Mihalatos M, Razi E, Labropoulos S, Petounis A, Gaki V, Keramopoulos A, Pandis N, Kyriacou K, Hadjisavvas A, Kosmidis P, Yannoukakos D, Nasioulas G 2004. Characterization of a novel large deletion and single point mutations in the
BRCA1 gene in a Greek cohort of families with suspected hereditary breast cancer. *BMC Cancer* 4:61. DOI: 10.1186/1471-2407-4-61.

7. Buffone A, Capalbo C, Ricevuto E, Sidoni T, Ottini L, Falchetti M, Cortesi E, Marchetti P, Scambia G, Tomao S, Rinaldi C, Zani M, Ferraro S, Frati L, Screpanti I, Gulino A, Giannini G 2007. Prevalence of BRCA1 and BRCA2 genomic rearrangements in a cohort of consecutive Italian breast and/or ovarian cancer families. *Breast Cancer Research and Treatment* 106:289–296. DOI: 10.1007/s10549-007-9499-6.

8. Capalbo C, Ricevuto E, Vestri A, Ristori E, Sidoni T, Buffone O, Adamo B, Cortesi E, Marchetti P, Scambia G, Tomao S, Rinaldi C, Zani M, Ferraro S, Frati L, Screpanti I, Gulino A, Giannini G 2006a. BRCA1 and BRCA2 genetic testing in Italian breast and/or ovarian cancer families: mutation spectrum and prevalence and analysis of mutation prediction models. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 17 Suppl 7:vii34–40. DOI: 10.1093/annonc/mdl947.

9. Capalbo C, Ricevuto E, Vestri A, Sidoni T, Buffone A, Cortesi E, Marchetti P, Scambia G, Tomao S, Rinaldi C, Zani M, Ferraro S, Frati L, Screpanti I, Gulino A, Giannini G 2006b. Improving the accuracy of BRCA1/2 mutation prediction: validation of the novel country-customized IC software. *European journal of human genetics : EJHG* 14:49–54. DOI: 10.1038/sj.ejhg.5201511.

10. Carvalho M, Pino MA, Karchin R, Beddor J, Godinho-Netto M, Mesquita RD, Rodarte RS, Vaz DC, Monteiro VA, Manoukian S, Colombo M, Ripamonti CB, Rosenquist R, Suthers G, Borg A, Radice P, Grist SA, Monteiro ANA, Billack B 2009. Analysis of a set of missense, frameshift, and in-frame deletion variants of BRCA1. *Mutation research* 660:1–11. DOI: 10.1016/j.mrfmmm.2008.09.017.
11. Concolino P, Mello E, Minucci A, Santonocito C, Scambia G, Giardina B, Capoluongo E
2014. Advanced tools for BRCA1/2 mutational screening: comparison between two methods
for large genomic rearrangements (LGRs) detection. *Clinical chemistry and laboratory
medicine* 52:1119–1127. DOI: 10.1515/cclm-2013-1114.

12. Coppa A, Buffone A, Capalbo C, Nicolussi A, D’Inzeo S, Belardinilli F, Colicchia V, Petroni
M, Granato T, Midulla C, Zani M, Ferraro S, Serepanti I, Gulino A, Giannini G 2014. Novel
and recurrent BRCA2 mutations in Italian breast/ovarian cancer families widen the ovarian
cancer cluster region boundaries to exons 13 and 14. *Breast Cancer Research and Treatment
148*:629–635. DOI: 10.1007/s10549-014-3196-z.

13. Coppa A, Nicolussi A, D’Inzeo S, Capalbo C, Belardinilli F, Colicchia V, Petroni M, Zani M,
Ferraro S, Rinaldi C, Buffone A, Bartolazzi A, Serepanti I, Ottini L, Giannini G 2018.
Optimizing the identification of risk-relevant mutations by multigene panel testing in selected
hereditary breast/ovarian cancer families. *Cancer medicine* 7:46–55. DOI:
10.1002/cam4.1251.

14. Couch FJ, Antoniou AC, Ramus SJ, Kuchenbaecker KB, Soucy P, Beesley J, Chen X, Wang
X, Kirchhoff T, McGuffog L, Barrowdale D, Lee A, Healey S, Sinilnikova OM, Andrus IL,
for OCGN, Ozcelik H, Mulligan AM, Thomassen M, Gerdes A-M, Skytte A-B, Kruse TA,
Caligo MA, Wachenfeldt von A, Barbany-Bustinza G, Loman N, Soller M, Ehrencrona H,
Karlsson P, for SWE-BRCA, Nathanson KL, Domchek SM, Jakubowska A, Lubinski J,
Jaworska K, Durda K, Zlowocka E, Huzarski T, Byrski T, Gronwald J, Cybulski C, Górski B,
Osorio A, Durán M, Tejada MI, Benítez J, Hamann U, for HEBON, van Os TA, van Leeuwen
FE, Meijers-Heijboer HEJ, Wijnen J, Blok MJ, Kets M, Hooning MJ, Oldenburg RA, Peock
S, Frost D, Ellis SD, Platte R, Fineberg E, Evans DG, Jacobs C, Eeles RA, Adlard J,
Davidson R, Cole T, Cook J, Paterson J, Brewer C, Douglas F, Hodgson SV, Morrison PJ, Walker L, Porteous ME, Kennedy MJ, for EMBRACE, Bove B, Godwin AK, Stoppa-Lyonnet D, for GEMO Study Collaborators, Fassy-Colcombet M, Castera L, Cornelis F, Mazoyer S, Léoné M, Boutry-Kryza N, Bressac-de Paillerets B, Caron O, Pujol P, Coupier I, Delnatte C, Akloul L, Lynch HT, Snyder CL, Buys SS, Daly MB, Terry M, Chung WK, John EM, Miron A, Southey MC, Hopper JL, Goldgar DE, Singer CF, Tea M-KM, Fink-Retter A, Nielsen FC, Arason A, Shah S, Sarrel K, Robson ME, Piedmonte M, Phillips K, Basil J, Rubinstein WS, Boggess J, Wakeley K, Ewart-Toland A, Montagna M, Isaacs C, Lazaro C, Blanco I, Feliubadalo L, Brunet J, Gayther SA, Pharoah PPD, Odunsi KO, Karlan BY, Walsh CS, Olah E, Teo SH, Ganz PA, Beattie MS, Dorfling CM, Diez O, Kwong A, Schmutzler RK, Engel C, Meindl A, Ditsch N, Arnold N, Heidemann S, Niederacher D, Preisler-Adams S, Gadzicki D, Varon-Mateeva R, Deissler H, Gehrig A, Sutter C, Kast K, Fiebig B, Heinritz W, Caldés T, la Hoya de M, Muranen TA, Nevanlinna H, Tischkowitz MD, Spurdle AB, Neuhausen SL, Ding YC, Lindor NM, Frederiksen Z, Pankratz VS, Peterlongo P, Manoukian S, Peissel B, Zaffaroni D, Bernard L, Viel A, Giannini G, Varesco L, Radice P, Greene MH, Easton DF, Chenevix-Trench G, for kConFab investigators, Offit K, Simard J, on behalf of the Consortium of Investigators of Modifiers of BRCA1/2 2012. 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 epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. DOI: 10.1158/1055-9965.EPI-11-0888.
15. EB W 2019. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*:1–1.

16. Economopoulou P, Dimitriadis G, Psyrri A 2015. Beyond BRCA: new hereditary breast cancer susceptibility genes. *Cancer treatment reviews* 41:1–8. DOI: 10.1016/j.ctrv.2014.10.008.

17. Enyedi MZ, Jaksa G, Pintér L, Sükösd F, Gyuris Z, Hajdu A, Határvölgyi E, Priskin K, Haracska L 2016. Simultaneous detection of BRCA mutations and large genomic rearrangements in germline DNA and FFPE tumor samples. *Oncotarget*. DOI: 10.18632/oncotarget.11259.

18. Ewald IP, Ribeiro PLI, Palmero EI, Cossio SL, Giugliani R, Ashton-Prolla P 2009. Genomic rearrangements in BRCA1 and BRCA2: A literature review. *Genetics and molecular biology* 32:437–446. DOI: 10.1590/S1415-47572009005000049.

19. Feliubadalo L, Lopez-Doriga A, Castellsagué E, del Valle J, Menéndez M, Tornero E, Montes E, Cuesta R, Gómez C, Campos O, Pineda M, González S, Moreno V, Brunet J, Blanco I, Serra E, Capellá G, Lazaro C 2013. Next-generation sequencing meets genetic diagnostics: development of a comprehensive workflow for the analysis of BRCA1 and BRCA2 genes. *European journal of human genetics : EJHG* 21:864–870. DOI: 10.1038/ejhg.2012.270.

20. Gad S, Caux-Moncoutier V, Pages-Berhouet S, Gauthier-Villars M, Coujper I, Pujol P, Frenay M, Gilbert B, Maugard C, Bignon Y-J, Chevrier A, Rossi A, Fricker J-P, Nguyen TD, Demange L, Aurias A, Bensimon A, Stoppa-Lyonnet D 2002. Significant contribution of large BRCA1 gene rearrangements in 120 French breast and ovarian cancer families. *Oncogene* 21:6841–6847. DOI: 10.1038/sj.onc.1205685.
21. Giannini G, Capalbo C, Ristori E, Ricevuto E, Sidoni T, Buffone A, Cortesi E, Marchetti P, Scambia G, Tomao S, Rinaldi C, Zani M, Ferraro S, Frati L, Screpanti I, Gulino A 2006. Novel BRCA1 and BRCA2 germline mutations and assessment of mutation spectrum and prevalence in Italian breast and/or ovarian cancer families. Breast Cancer Research and Treatment 100:83–91. DOI: 10.1007/s10549-006-9225-9.

22. Hartmann C, John AL, Klaes R, Hofmann W, Bielen R, Koehler R, Janssen B, Bartram CR, Arnold N, Zschocke J 2004. Large BRCA1 gene deletions are found in 3% of German high-risk breast cancer families. Human Mutation 24:534. DOI: 10.1002/humu.9291.

23. Hasmad HN, Sivanandan K, Lee V, Yip CH, Mohd Taib NA, Teo SH 2015. Identification of a recurrent BRCA1 exon 21-22 genomic rearrangement in Malay breast cancer patients. Clinical genetics 87:392–394. DOI: 10.1111/cge.12451.

24. Hogervorst FBL, Nederlof PM, Gille JJP, McElgunn CJ, Grippeling M, Pruntel R, Regnerus R, van Welsem T, van Spaendonk R, Menko FH, Kluijt I, Dommering C, Verhoef S, Schouten JP, Van't Veer LJ, Pals G 2003. Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. Cancer research 63:1449–1453.

25. Judkins T, Rosenthal E, Arnell C, Burbidge LA, Geary W, Barrus T, Schoenberger J, Trost J, Wenstrup RJ, Roa BB 2012. Clinical significance of large rearrangements in BRCA1 and BRCA2. Cancer 118:5210–5216. DOI: 10.1002/cncr.27556.

26. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J, Jervis S, van Leeuwen FE, Milne RL, Andrieu N, Goldgar DE, Terry MB, Rookus MA, Easton DF, Antoniou AC, BRCA1 and BRCA2 Cohort Consortium, McGuffog L, Evans DG, Barrowdale D, Frost D, Adlard J, Ong K-R, Izatt L, Tischkowitz M, Eeles R, Davidson R, Hodgson S, Ellis S, Nogues C, Lasset C, Stoppa-Lyonnet D, Fricker J-P, Faivre L, Berthet P, Hooning
MJ, van der Kolk LE, Kets CM, Adank MA, John EM, Chung WK, Andrulis IL, Southey M,
Daly MB, Buys SS, Osorio A, Engel C, Kast K, Schmutzler RK, Caldés T, Jakubowska A,
Simard J, Friedlander ML, McLachlan S-A, Machackova E, Foretova L, Tan YY, Singer CF,
Olah E, Gerdes A-M, Arver B, Olsson H 2017. Risks of Breast, Ovarian, and Contralateral
Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA : the journal of the
American Medical Association* 317:2402–2416. DOI: 10.1001/jama.2017.7112.
27. Kuchenbaecker KB, Neuhausen SL, Robson M, Barrowdale D, McGuffog L, Mulligan AM,
Andrulis IL, Spurdle AB, Schmidt MK, Schmutzler RK, Engel C, Wappenschmidt B,
Nevanlinna H, Thomassen M, Southey M, Radice P, Ramus SJ, Domchek SM, Nathanson
KL, Lee A, Healey S, Nussbaum RL, Rebbeck TR, Arun BK, James P, Karlan BY, Lester J,
Cass I, Breast Cancer Family Registry, Terry MB, Daly MB, Goldgar DE, Buys SS,
Janavicius R, Tihomirova L, Tung N, Dorfling CM, van Rensburg EJ, Steele L, O Hansen von
T, Ejlertsen B, Gerdes A-M, Nielsen FC, Dennis J, Cunningham J, Hart S, Slager S, Osorio A,
Benítez J, Durán M, Weitzel JN, Tafur I, Hander M, Peterlongo P, Manoukian S, Peissel B,
Roversi G, Scuvera G, Bonanni B, Mariani P, Volorio S, Dolcetti R, Varesto L, Papi L,
Tibiletti MG, Giannini G, Fostira F, Konstantopoulou I, Garber J, Hamann U, Donaldson A,
Brewer C, Foo C, Evans DG, Frost D, Eccles D, EMBRACE Study, Douglas F, Brady A,
Cook J, Tischkowitz M, Adlard J, Barwell J, Ong K-R, Walker L, Izatt L, Side LE, Kennedy
MJ, Rogers MT, Porteous ME, Morrison PJ, Platte R, Eeles R, Davidson R, Hodgson S, Ellis
S, Godwin AK, Rhiem K, Meindl A, Ditsch N, Arnold N, Plendl H, Niederacher D, Sutter C,
Steinemann D, Bogdanova-Markov N, Kast K, Varon-Mateeva R, Wang-Gohrke S, Gehrig A,
Markiefka B, Buecher B, Lefol C, Stoppa-Lyonnet D, Rouleau E, Prieur F, Damiola F,
GEMO Study Collaborators, Barjhoux L, Faivre L, Longy M, Sevenet N, Sinilnikova OM,
Mazoyer S, Bonadona V, Caux-Moncoutier V, Isaacs C, Van Maerken T, Claes K, Piedmonte M, Andrews L, Hays J, Rodriguez GC, Caldés T, la Hoya de M, Khan S, Hogervorst FBL, Aalfs CM, de Lange JL, Meijers-Heijboer HEJ, van der Hout AH, Wijnen JT, van Roozendaal KEP, Mensenkamp AR, van den Ouweland AMW, van Deurzen CHM, van der Luijt RB, HEBON, Olah E, Diez O, Lazaro C, Blanco I, Teulé A, Menéndez M, Jakubowska A, Lubinski J, Cybulski C, Gronwald J, Jaworska-Bieniek K, Durda K, Arason A, Maugard C, Soucy P, Montagna M, Agata S, Teixeira MR, kConFab Investigators, Olswold C, Lindor N, Pankratz VS, Hallberg E, Wang X, Szabo CI, Vijai J, Jacobs L, Corines M, Lincoln A, Berger A, Fink-Retter A, Singer CF, Rappaport C, Kaulich DG, Pfeiler G, Tea M-K, Phelan CM, Mai PL, Greene MH, Rennert G, Imyanitov EN, Glendon G, Toland AE, Bojesen A, Pedersen IS, Jensen UB, Caligo MA, Friedman E, Berger R, Laitman Y, Rantala J, Arver B, Loman N, Borg A, Ehrencrona H, Olopade OI, Simard J, Easton DF, Chenevix-Trench G, Offit K, Couch FJ, Antoniou AC, CIMBA 2014. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. Breast cancer research : BCR 16:3416. DOI: 10.1186/s13058-014-0492-9.

28. la Hoya de M, Gutiérrez-Enríquez S, Velasco E, Osorio A, Sánchez de Abajo A, Vega A, Salazar R, Esteban E, Llort G, González-Sarmiento R, Carracedo Á, Benitez J, Miner C, Díez O, Díaz-Rubio E, Caldés T 2006. Genomic rearrangements at the BRCA1 locus in Spanish families with breast/ovarian cancer. Clinical chemistry 52:1480–1485. DOI: 10.1373/clinchem.2006.070110.

29. Mazoyer S 2005. Genomic rearrangements in theBRCA1 andBRCA2 genes. Human Mutation 25:415–422. DOI: 10.1002/humu.20169.
30. Moisan A-M, Fortin J, Dumont M, Samson C, Bessette P, Chiquette J, Laframboise R, Lépine J, Lespérance B, Pichette R, Plante M, Provencher L, Voyer P, Goldgar D, Bridge P, Simard J 2006. No Evidence of BRCA1/2 genomic rearrangements in high-risk French-Canadian breast/ovarian cancer families. Genetic testing 10:104–115. DOI: 10.1089/gte.2006.10.104.

31. Montagna M, Dalla Palma M, Menin C, Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E 2003. Genomic rearrangements account for more than one-third of the BRCA1 mutations in northern Italian breast/ovarian cancer families. Human molecular genetics 12:1055–1061.

32. Nicolussi A, Belardinilli F, Mahdavian Y, Colicchia V, D'Inzeo S, Petroni M, Zani M, Ferraro S, Valentini V, Ottini L, Giannini G, Capalbo C, Coppa A 2019. Next-generation sequencing of BRCA1 and BRCA2 genes for rapid detection of germline mutations in hereditary breast/ovarian cancer. PeerJ 7:e6661. DOI: 10.7717/peerj.6661.

33. Ottini L, Silvestri V, Saieva C, Rizzolo P, Zanna I, Falchetti M, Masala G, Navazio AS, Graziano V, Bianchi S, Manoukian S, Barile M, Peterlongo P, D'Amico C, Varesco L, Tommasi S, Russo A, Giannini G, Cortesi L, Viel A, Montagna M, Radice P, Palli D 2013. Association of low-penetrance alleles with male breast cancer risk and clinicopathological characteristics: results from a multicenter study in Italy. Breast Cancer Research and Treatment 138:861–868. DOI: 10.1007/s10549-013-2459-4.

34. Palma M, Ristori E, Ricevuto E, Giannini G, Gulino A 2006. BRCA1 and BRCA2: the genetic testing and the current management options for mutation carriers. Critical reviews in oncology/hematology 57:1–23. DOI: 10.1016/j.critrevonc.2005.05.003.

35. Peterlongo P, Chang-Claude J, Moysich KB, Rudolph A, Schmutzler RK, Simard J, Soucy P, Eeles RA, Easton DF, Hamann U, Wilkening S, Chen B, Rookus MA, Schmidt MK, van der Baan FH, Spurdle AB, Walker LC, Lose F, Maia A-T, Montagna M, Matricardi L, Lubinski J,
Bojesen A, Pedersen IS, Thomassen M, Jensen UB, Laitman Y, Rantala J, Wachenfeldt von A, Ehrencrona H, Askmal M, Borg A, Kuchenbaecker KB, McGuffog L, Barrowdale D, Healey S, Lee A, Pharoah PDP, Chenevix-Trench G, kConFab Investigators, Antoniou AC, Friedman E 2015. Candidate genetic modifiers for breast and ovarian cancer risk in BRCA1 and BRCA2 mutation carriers. Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 24:308–316. DOI: 10.1158/1055-9965.EPI-14-0532.

36. Pietschmann A, Mehdipour P, Atri M, Hofmann W, Hosseini-Asl SS, Scherneck S, Mundlos S, Peters H 2005. Mutation analysis of BRCA1 and BRCA2 genes in Iranian high risk breast cancer families. Journal of cancer research and clinical oncology 131:552–558. DOI: 10.1007/s00432-005-0678-8.

37. Pilato B, Pinto R, De Summa S, Petriella D, Lacalamita R, Danza K, Paradiso A, Tommasi S 2016. BRCA1-2 diagnostic workflow from next-generation sequencing technologies to variant identification and final report. Genes, chromosomes & cancer 55:803–813. DOI: 10.1002/gcc.22383.

38. Rebbeck TR, Friebel TM, Friedman E, Hamann U, Huo D, Kwong A, Olah E, Olopade OI, Solano AR, Teo SH, Thomassen M, Weitzel JN, Chan TL, Couch FJ, Goldgar DE, Kruse TA, Palmero EL, Park SK, Torres D, van Rensburg EJ, McGuffog L, Parsons MT, Leslie G, Aalfs CM, Abugattas J, Adlard J, Agata S, Aittomaki K, Andrews L, Andrus IL, Arason A, Arnold N, Arun BK, Asseryanis E, Auerbach L, Azzollini J, Balmaña J, Barile M, Barkardottir RB, Barrowdale D, Benítez J, Berger A, Berger R, Blanco AM, Blazer KR, Blok MJ, Bonadona V, Bonanni B, Bradbury AR, Brewer C, Buecher B, Buys SS, Caldés T, Caliebe A, Caligo MA, Campbell I, Caputo SM, Chiquette J, Chung WK, Claes KBM, Collée JM, Cook J,
Davidson R, la Hoya de M, De Leeneer K, de Pauw A, Delnatte C, Diez O, Ding YC, Ditsch N, Domchek SM, Dorfling CM, Velazquez C, Dworniczak B, Eason J, Easton DF, Eeles R, Ehrencrona H, Ejlertsen B, EMBRACE, Engel C, Engert S, Evans DG, Faivre L, Feliubadaló L, Ferrer SF, Foretova L, Fowler J, Frost D, Galvão HCR, Ganz PA, Garber J, Gauthier-Villars M, Gehrig A, GEMO Study Collaborators, Gerdes A-M, Gesta P, Giannini G, Giraud S, Glendon G, Godwin AK, Greene MH, Gronwald J, Gutierrez-Barrera A, Hahnen E, Hauke J, HEBON, Henderson A, Hentschel J, Hogervorst FBL, Honisch E, Imyanitov EN, Isaacs C, Izatt L, Izquierdo A, Jakubowska A, James P, Janavicius R, Jensen UB, John EM, Vijai J, Kaczmarek K, Karlan BY, Kast K, Investigators K, Kim S-W, Konstantopoulou I, Korach J, Laitman Y, Lasa A, Lasset C, Lazaro C, Lee A, Lee MH, Lester J, Lesueur F, Liljegren A, Lindor NM, Longy M, Loud JT, Lu KH, Lubinski J, Machackova E, Manoukian S, Mari V, Martínez-Bouzas C, Matrai Z, Mebirouk N, Meijers-Heijboer HE, Meindl A, Mensenkamp AR, Mickys U, Miller A, Montagna M, Moysich KB, Mulligan AM, Musinsky J, Neuhausen SL, Nevanlinna H, Ngeow J, Nguyen HP, Niederacher D, Nielsen HR, Nielsen FC, Nussbaum RL, Offit K, Øverholm A, Ong K-R, Osorio A, Papi L, Papp J, Pasini B, Pedersen IS, Peixoto A, Peruga N, Peterlongo P, Pohl E, Pradhan N, Prazzardone K, Prieur F, Pujol P, Radice P, Ramus SJ, Rantala J, Rashid MU, Rhiem K, Robson M, Rodriguez GC, Rogers MT, Rudaitis V, Schmidt AY, Schmutzler RK, Senter L, Shah PD, Sharma P, Side LE, Simard J, Singer CF, Skytte A-B, Slavin TP, Snape K, Sobol H, Southey M, Steele L, Steinemann D, Sukkinnicki G, Sutter C, Szabo CI, Tan YY, Teixeira MR, Terry MB, Teulé A, Thomas A, Thull DL, Tischkowitz M, Tognazzo S, Toland AE, Topka S, Trainer AH, Tung N, van Asperen CJ, van der Hout AH, van der Kolk LE, van der Luijt RB, Van Heetvelde M, Varesc L, Varon-Mateeva R, Vega A, Villarreal-Garza C, Wachenfeldt von A,
Walker L, Wang-Gohrke S, Wappenschmidt B, Weber BHF, Yannoukakos D, Yoon S-Y, Zanzottera C, Zidan J, Zorn KK, Hutten Selkirk CG, Hulick PJ, Chenevix-Trench G, Spurdle AB, Antoniou AC, Nathanson KL 2018. Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. Human Mutation 39:593–620. DOI: 10.1002/humu.23406.

Rebbeck TR, Mitra N, Wan F, Healey S, McGuffog L, Chenevix-Trench G, Easton DF, Antoniou AC, Nathanson KL, CIMBA Consortium, Laitman Y, Kushnir A, Paluch-Shimon S, Berger R, Zidan J, Friedman E, Ehrencreuna H, Stenmark-Askmalm M, Einbeigi Z, Loman N, Harbst K, Rantala J, Melin B, Huo D, Olopade OI, Seldon J, Ganz PA, Nussbaum RL, Chan SB, Odunsi K, Gayther SA, Domchek SM, Arun BK, Lu KH, Mitchell G, Karlan BY, Walsh C, Lester J, Godwin AK, Pathak H, Ross E, Daly MB, Whittemore AS, John EM, Miron A, Terry MB, Chung WK, Goldgar DE, Buys SS, Janavicius R, Tihomirova L, Tung N, Dorfling CM, van Rensburg EJ, Steele L, Neuhausen SL, Ding YC, Ejlertsen B, Gerdes A-M, Hansen TVO, Ramón y Cajal T, Osorio A, Benítez J, Godino J, Tejada MI, Durán M, Weitzel JN, Bobolis KA, Sand SR, Fontaine A, Savarese A, Pasini B, Peissel B, Bonanni B, Zaffaroni D, Vignolo-Lutati F, Scuvera G, Giannini G, Bernard L, Genuardi M, Radice P, Dolcetti R, Manoukian S, Pensotti V, Gismondi V, Yannoukakos D, Fostira F, Garber J, Torres D, Rashid MU, Hamann U, Peock S, Frost D, Platte R, Evans DG, Eeles R, Davidson R, Eccles D, Cole T, Cook J, Brewer C, Hodgson S, Morrison PJ, Walker L, Porteous ME, Kennedy MJ, Izatt L, Adlard J, Donaldson A, Ellis S, Sharma P, Schmutzler RK, Wappenschmidt B, Becker A, Rhiem K, Hahnem E, Engel C, Meindl A, Engert S, Ditsch N, Arnold N, Plendl HJ, Mundhenke C, Niederacher D, Fleisch M, Sutter C, Bartram CR, Dikow N, Wang-Gohrke S, Gadzicki D, Steinemann D, Kast K, Beer M, Varon-Mateeva R, Gehrig A, Weber BH,
Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *JAMA: the journal of the American Medical Association* 313:1347–1361.

DOI: 10.1001/jama.2014.5985.
Spanish breast/ovarian cancer families: review of the literature, and reevaluation of the genetic mechanisms involved in their origin. *Breast Cancer Research and Treatment*. DOI: 10.1007/s10549-011-1909-0.

41. Schenkel LC, Kerkhof J, Stuart A, Reilly J, Eng B, Woodside C, Levstik A, Howlett CJ, Rupar AC, Knoll JHM, Ainsworth P, Waye JS, Sadikovic B 2016. Clinical Next-Generation Sequencing Pipeline Outperforms a Combined Approach Using Sanger Sequencing and Multiplex Ligation-Dependent Probe Amplification in Targeted Gene Panel Analysis. *The Journal of molecular diagnostics : JMD* 18:657–667. DOI: 10.1016/j.jmoldx.2016.04.002.

42. Schmidt AY, Hansen TVO, Ahlborn LB, Jønson L, Yde CW, Nielsen FC 2017. Next-Generation Sequencing-Based Detection of Germline Copy Number Variations in BRCA1/BRCA2: Validation of a One-Step Diagnostic Workflow. *The Journal of molecular diagnostics : JMD* 19:809–816. DOI: 10.1016/j.jmoldx.2017.07.003.

43. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic acids research* 30:e57.

44. Sluiter MD, van Rensburg EJ 2011. Large genomic rearrangements of the BRCA1 and BRCA2 genes: review of the literature and report of a novel BRCA1 mutation. *Breast Cancer Research and Treatment* 125:325–349. DOI: 10.1007/s10549-010-0817-z.

45. Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M, Taylor M, Hood L, King MC 1996. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1. *Genome research* 6:1029–1049.
46. Tan R, Wang Y, Kleinstein SE, Liu Y, Zhu X, Guo H, Jiang Q, Allen AS, Zhu M 2014. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Human Mutation* 35:899–907. DOI: 10.1002/humu.22537.

47. Tarabeux J, Zeitouni B, Moncoutier V, Tenreiro H, Abidallah K, Lair S, Legoix-Né P, Leroy Q, Rouleau E, Golmard L, Barillot E, Stern M-H, Rio-Frio T, Stoppa-Lyonnet D, Houdayer C 2014. Streamlined ion torrent PGM-based diagnostics: BRCA1 and BRCA2 genes as a model. *European journal of human genetics : EJHG* 22:535–541. DOI: 10.1038/ejhg.2013.181.

48. Ticha I, Kleibl Z, Stribrna J, Kotlas J, Zimovjanova M, Mateju M, Zikan M, Pohlreich P 2010. Screening for genomic rearrangements in BRCA1 and BRCA2 genes in Czech high-risk breast/ovarian cancer patients: high proportion of population specific alterations in BRCA1 gene. *Breast Cancer Research and Treatment* 124:337–347. DOI: 10.1007/s10549-010-0745-y.

49. Trujillano D, Weiss MER, Schneider J, Köster J, Papachristos EB, Saviouk V, Zakharkina T, Nahavandi N, Kovacevic L, Rolfs A 2015. Next-generation sequencing of the BRCA1 and BRCA2 genes for the genetic diagnostics of hereditary breast and/or ovarian cancer. *The Journal of molecular diagnostics : JMD* 17:162–170. DOI: 10.1016/j.jmoldx.2014.11.004.

50. Vasickova P, Machackova E, Lukesova M, Damborsky J, Horky O, Pavlu H, Kuklova J, Kosinova V, Navratilova M, Foretova L 2007. High occurrence of BRCA1 intragenic rearrangements in hereditary breast and ovarian cancer syndrome in the Czech Republic. *BMC medical genetics* 8:32. DOI: 10.1186/1471-2350-8-32.

51. Wallace AJ 2016. New challenges for BRCA testing: a view from the diagnostic laboratory. *European journal of human genetics : EJHG* 24 Suppl 1:S10–8. DOI: 10.1038/ejhg.2016.94.
52. Woodward AM, Davis TA, Silva AGS, Kirk JA, Leary JA, kConFab Investigators 2005. Large genomic rearrangements of both BRCA2 and BRCA1 are a feature of the inherited breast/ovarian cancer phenotype in selected families. *Journal of medical genetics* 42:e31. DOI: 10.1136/jmg.2004.027961.

53. Xia LC, Bell JM, Wood-Bouwens C, Chen JJ, Zhang NR, Ji HP 2018. Identification of large rearrangements in cancer genomes with barcode linked reads. *Nucleic acids research* 46:e19. DOI: 10.1093/nar/gkx1193.
Table 1 (on next page)

LGRs in TS and VS
| sample Id | genomic variant          | exon deletion               | ref                  |
|-----------|--------------------------|------------------------------|----------------------|
| TS        |                          |                              |                      |
| BR59      | c.5407-?_* (1_?)del      | exon 23-24 del               | Buffone et al., 2007 |
| BR328     | c.5075-?_5193+?del       | exon 18-19 del               | Buffone et al., 2007 |
| BR409     | NBR2delEX1_BRCA1delEX1-2 | exon 1 NBR2 del, exon 1-2 BRCA1 del | Coppa et al., 2018   |
| VS        |                          |                              |                      |
| BR963     | NG_005905.2: g.163181_169408del6228 | exon 21-22 del               | /                    |
| BR1154    | NG_005905.2: g.160396_164568del4173 | exon 20 del                  | /                    |
| BR1379    | NG_005905.2: g.145185_151339del6155 | exon 16-17 del               | /                    |
Table 2 (on next page)

CNVs prediction by IR software algorithm in TS and VS

The confidence score is the probability that the number of copies of the region of interest is different from 2, which is the normal value, while the confident precision indicates how much the algorithm is certain of the accuracy of the number of copies estimated by the analysis.
| Sample ID | Locus       | Type | Genes | Location      | Length | Copy Number | CytoBand                                    | CNV Confidence | CNV Precision |
|-----------|-------------|------|-------|---------------|--------|-------------|---------------------------------------------|----------------|---------------|
| TS        | BR59        | CNV  | BRCA1 | exon 23-24    | 2.138kb| 1           | 17q21.31(41197602-41199740)x1                | 5.66           | 5.66          |
|           | BR328       | CNV  | BRCA1 | exon 18-19    | 749kb  | 1           | 17q21.31(41215277-41216026)x1                | 13.05          | 13.05         |
|           | BR409       | CNV  | BRCA1 | exon 2        | 275kb  | 1           | 17q21.31(41275973-41276248)x1                | 1.14           | 1.14          |
| VS        | BR963       | CNV  | BRCA1 | exon 21-22    | 2.18kb | 1           | 17q21.31(41201074-41203254)x1                | 9.14           | 9.14          |
|           | BR1379      | CNV  | BRCA1 | exon 16-18    | 7.44kb | 1           | 17q21.31(41215855-41223295)x1                | 5.11           | 5.11          |
Table 3 (on next page)

Performance of NGS-dependent LGRs analysis

Validation metrics were defined as: Accuracy = (TP + TN)/(TP + FP + TN + FN); Sensitivity = TP/(TP + FN); Specificity = TN/(TN+FP); FDR = FP/(TP + FP), Negative Predictive Value = TN/(TN + FN), where TP = true positives, TN = true negatives, FP = false positives, FN = false negatives.
|   | MLPA |   |   | Results |
|---|------|---|---|---------|
|   | tot  | LGR | no LGR |         |
| DQ | LGR  | 33 | 3 | 30 | 64.7% accuracy [95% CI: 50.6-76.7] |
|   | no LGR | 52 | 0 | 52 | 63.4% specificity [95% CI: 49-75.8] |
| IR | LGR  | 29 | 2 | 27 | 66.7% sensitivity [95% CI: 8.9-98.8] |
|   | no LGR | 56 | 1 | 55 | 67.1% specificity [95% CI: 52.7-78.9] |
Figure 1

DQ analyses for TS samples.

(A) for each sample, every peak represents the ratio of the amplicon read count normalized on BRCA1/BRCA2 total reads and the average of normalized amplicon read counts from all samples on a single chip (DQA). (B) for each sample, every peak represents the ratio of the amplicon read count normalized on BRCA1/BRCA2 total reads and the average of the coverage data of a baseline built from 127 LGRs negative samples (DQB1). The threshold = mean ± 2 SD. Value > mean ± 2 SD is indicative of a duplication; Value < mean ± 2 SD is indicative of a deletion. * indicated the amplicons included in the region involved in the rearrangement as confirmed by MLPA analysis.
Manuscript to be reviewed
Figure 2

BRCA1 MLPA electropherogram showing aberrant profiles in BR963, BR1154, and BR1379 patients.

(A) Wild-type sample (WT). Black arrows indicate the deletion of (B) BRCA1 exons 21-22 (BR963), (C) BRCA1 exon 20 (BR1154), (D) BRCA1 exons 16-17 (BR1379).
Figure 3

DQ analyses for the representative samples for VS

(A) and (B) for each sample, every peak represents the ratio of the amplicon read count normalized on \( BRCA1/BRCA2 \) total reads and the average of normalized amplicon read counts from all samples on a single chip (DQA). (C) for each sample, every peak represents the ratio of the amplicon read count normalized on \( BRCA1/BRCA2 \) total reads and the average of the coverage data of a baseline built from 127 LGRs negative samples considering separately the amplicon pools (DQB2, pool 2). The threshold= mean ± 2 SD. Value > mean ± 2 SD is indicative of a duplication; Value < mean ± 2 SD is indicative of a deletion. * indicated the amplicons included in the region involved in the rearrangement as confirmed by MLPA analysis.
A

B

C

Manuscript to be reviewed
Figure 4

Characterization of BRCA1 LGRs

(A) gel image of PCR products. PCR amplification of the genomic region spanning the *BRCA1* rearrangement resulted in a fragment of approximately 1353 bp present only in the proband BR963. (B) and (C) schematic representation and electropherogram showing the *BRCA1* exons 21 and 22 deletion. The variant arose from an erroneous homologous recombination process between an AluSq2 (Alu family, SINE class; chr17:41206762-41207066) and an AluSz (Alu family, SINE class; chr17:41200521-41200834) motif, and it involved a perfectly repeated stretch of 24 bp. (D) gel image of PCR products. PCR amplification of the genomic region spanning the *BRCA1* rearrangement resulted in a fragment of approximately 872 bp present in the proband BR1154 and in her mother BR1148. (E) and (F) schematic representation and electropherogram showing the *BRCA1* exon 20 deletion. The variant arose from an erroneous homologous recombination process between two AluY motif at chr17:41205398-41205698 and chr17:41205398-41205698, respectively, and it involved a perfectly repeated stretch of 11 bp. (G) gel image of PCR products. PCR amplification of the genomic region spanning the *BRCA1* rearrangement resulted in a fragment of approximately 2027 bp present only in the proband BR1379. (H) and (I) schematic representation and electropherogram showing the *BRCA1* exons 16 and 17 deletion. The variant arose from an erroneous homologous recombination process between an AluSp motif (Alu family, SINE class; chr17:41224585-41224884) and an AluSg (Alu family, SINE class; chr17:41218424-41218724) motif, and it involved a perfectly repeated stretch of 16 bp. MK, marker; NT, no template; CTR healthy individual DNA.
BR963  NG_005905.2:g.163181_169408del6228

A

B

C

BR1154 e BR1148  NG_005905.2:g.160396_164568del4173

D

E

F

BR1379  NG_005905.2:g.145185_151339del6155

G

H

I
Figure 5

Pedigree of the HBC or HBOC family carriers of BRCA1 novel LGRs

(A) exons 21-22 deletion (BR963). (B) exon 20 deletion (BR1154). (C) exons 16-17 deletion (BR1379). Probands are indicated with an arrow. Cancer type and age at diagnosis are reported and described as: BC, breast cancer; Pan, pancreas; Leu, leukemia; Lung; bil BC, bilateral breast cancer; OC, ovarian cancer; Hep, hepatoma; CC, colon cancer.
Figure 6

Operative algorithm to select samples for MLPA analysis.
Coverage Data analysis (average base coverage depth ≥ 500)

DQA

DQB

double negative

clinical report

DQ positive (A and B)*

DQ and IR positive*

MLPA analysis

negative

positive

clinical report

confirmation assays/breakpoint characterization

positive

negative

clinical report

IR