

International distribution and age estimation of the Portuguese *BRCA2*

c.156_157insAlu founder mutation

Ana Peixoto, Catarina Santos, Manuela Pinheiro, Pedro Pinto, Maria José Soares, Patrícia Rocha, Leonor Gusmão, António Amorim, Annemarie van der Hout, Anne-Marie Gerdes, Mads Thomassen, Torben Kruse, Dorthe Cruger, Lone Sunde, Yves-Jean Bignon, Nancy Uhrhammer, Lucie Cornil, Etienne Rouleau, Rosette Lidereau, Koulis Yannoukacos, Drakoulis Yannoukacos, Maroulio Pertesi, Steven Narod, Robert Royer, Maurício M. Costa, Conxi Lazaro, Lidia Feliubadaló, Begoña Graña, Ignacio Blanco, Miguel de la Hoya, Trinidad Caldés, Philippe Maillet, Gaelle Benais-Pont, Bruno Pardo, Eitan Friedman, Eladio Velasco, Mercedes Durán, Maria-Dolores Miramar, Ana R. Valle, Ana Vega, Ana Blanco, Orland Diez, Teresa Ramon, Carmen Alonso, Montserrat Baiget, Judith Balmaña, William Foulkes, Marc Tischkowitz, Rachel Kyle, Nelly Sabbaghian, Patricia Ashton-Prolla, Ingrid P. Ewald, Thangarajan Rajkumar, Luisa M. Vieira, Giuseppe Giannini, Alberto Gulino, Maria I. Achatz, Dirce M. Carraro, Brigitte Bressac, Audrey Remenieras, Cindy Benson, Maureen Chung, Erik Teugels, Manuel R. Teixeira

Insert affiliations here:

Running title: Portuguese *BRCA2* founder mutation

Key words: c.156_157insAlu *BRCA2* mutation, age estimation, hereditary breast/ovarian cancer.

Correspondence: Manuel R. Teixeira, MD, PhD, Department of Genetics, Portuguese Oncology Institute, Rua Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal

Phone: +351 225084000; Fax: +351 225084016

E-mail: manuel.teixeira@ipoporto.min-saude.pt

Abstract

The c.156_157insAlu *BRCA2* mutation has so far only been reported in hereditary breast/ovarian cancer (HBOC) families of Portuguese origin. Since this mutation is not detectable using the commonly used screening methodologies and must be specifically sought, we screened for this rearrangement in a total of 5,440 suspected HBOC families from 22 labs from 13 countries from several continents. Whereas the c.156_157insAlu *BRCA2* mutation was detected in 11 of 149 suspected HBOC families from Portugal, representing 37.9% of all deleterious mutations, in other countries it was detected only in one proband living in France and in four individuals requesting predictive testing living in France and in the USA, all having in common the fact that they are relatively recent immigrants of Portuguese origin in those countries. After performing an extensive haplotype study in carrier families, we estimate that this founder mutation has occurred 558 ± 215 years ago. We further demonstrate significant quantitative differences regarding the production of the *BRCA2* full length RNA and the transcript with exon 3 skipping in c.156_157insAlu *BRCA2* mutation carriers and in controls, indicating that disruption of alternative transcript ratios is the mechanism causing hereditary breast/ovarian cancer associated with this *BRCA2* rearrangement. We further show that the cumulative incidence of breast cancer in c.156_157insAlu *BRCA2* mutation carriers does not differ from that of other *BRCA2* and *BRCA1* pathogenic mutations, further strengthening its role as the major contributor to hereditary predisposition to breast cancer in Portugal. We recommend that all suspected HBOC families from Portugal or with Portuguese ancestry are specifically tested for this rearrangement, ideally prior to screening of the entire coding regions of *BRCA1* and *BRCA2*.

Introduction

The pattern of *BRCA1* and *BRCA2* mutations in hereditary breast/ovarian cancer (HBOC) families varies widely among different populations. Many present a wide spectrum of different mutations throughout these genes, while some ethnic groups show a high frequency of particular mutations due to founder effects (Fackenthal and Olopade, 2007; Ferla *et al.*, 2007). Identification of founder mutations in the various populations makes possible the use of more specific approaches to molecular testing (Filippini *et al.*, 2007), allowing the analysis of more patients with less stringent selection criteria. Furthermore, a frequent founder mutation in a population allows a more accurate estimation of mutation-specific cumulative cancer incidence, facilitating also identification of genetic and environmental risk modifiers.

The c.156_157insAlu *BRCA2* mutation was first described by Teugels *et al.* (2005) in a Portuguese patient residing in Belgium. These authors demonstrated that this exon 3 Alu insertion originates at the mRNA level an in-frame deletion of that exon, which encodes a transcriptional activation domain (Teugels *et al.*, 2005). Machado *et al.* (2007) later described a regional founder effect for this rearrangement in HBOC families mostly originated from central/southern Portugal. We recently evaluated the contribution of the c.156_157insAlu *BRCA2* mutation to inherited predisposition to breast/ovarian cancer in families originated mostly from northern/central Portugal (Peixoto *et al.*, 2009a) and found that this rearrangement is responsible for more than half of all deleterious *BRCA2* mutations and about one-fourth of all deleterious mutations in HBOC families. Additionally, in light of some doubts raised about the pathogenic effect of *BRCA2* exon 3 skipping (Diez *et al.*, 2007), we demonstrated its pathogenic effect by showing that the *BRCA2* full length transcript is produced almost exclusively from the wildtype allele in patients carrying the c.156_157insAlu *BRCA2*

rearrangement and that this mutation co-segregates with the disease in HBOC families and is absent in healthy blood donors, although residual exon 3 skipping in *BRCA2* mRNA can be found in negative controls (Peixoto *et al.*, 2009a; Peixoto *et al.*, 2009b).

Although all reported c.156_157insAlu *BRCA2* mutations have so far been identified in Portuguese HBOC families (Teugels *et al.*, 2005; Machado *et al.*, 2007; Peixoto *et al.*, 2009a), this mutation is not detected using the commonly used screening methodologies and must be specifically sought (Teugels *et al.*, 2005; Peixoto *et al.*, 2009b), so one can not currently rule out its presence in other populations. To gain insight into the ancestral origin and population spread of the c.156_157insAlu *BRCA2* mutation, we screened for this rearrangement in 5,440 suspected HBOC patients from several countries and performed an extensive haplotype study using closely linked microsatellite markers and single nucleotide polymorphisms (SNPs) in carrier families. In addition to estimate the age of the c.156_157insAlu *BRCA2* mutation, we used real-time RT-PCR to quantify the preferential production of the transcript lacking exon 3 by the mutated allele and its (presumably residual) production by both alleles in non-carriers.

Materials and methods

Families

This study comprised a total of 5,440 suspected HBOC families from 22 labs from 13 countries from several continents. From Portugal, 149 new suspected HBOC families were selected for *BRCA1* and *BRCA2* mutation screening using previously described criteria (Peixoto *et al.*, 2006; Peixoto *et al.*, 2009a) after written informed consent. Molecular testing at the Department of Genetics of the Portuguese Oncology

Institute, Porto, Portugal (IPO-Porto) started by looking for the c.156_157insAlu *BRCA2* mutation, followed by full *BRCA1* and *BRCA2* mutation screening with the previously reported methodology (Peixoto *et al.*, 2006; Peixoto *et al.*, 2009a; Peixoto *et al.*, 2009b). The c.156_157insAlu *BRCA2* mutation was additionally screened in 5,291 suspected HBOC families living in countries other than Portugal in whom no deleterious *BRCA1/BRCA2* mutations had previously been found, with the following distribution: 1,209 from Spain (356 from Girona, 341 from Madrid, 151 from Valladolid, 132 from Zaragoza, 123 from Santiago de Compostela, and 106 from Barcelona), 1,087 from France (650 from Clermont-Ferrand, 428 from Saint-Cloud, and 9 from Villejuif, all the latter with Portuguese ancestry), 820 from Holland (Groningen), 758 from Denmark (Odense), 400 from Greece (Athens), 219 from Switzerland (Geneva), 200 from Belgium (Brussels), 185 from Israel (Tel Aviv), 144 from Brazil (98 from Porto Alegre and 46 from S. Paulo), 103 from Canada (Montreal), 91 from India (Chennai), and 75 from Italy (Rome). Besides the suspected HBOC families, two consecutive series of breast cancer patients from Rio de Janeiro, Brazil (390), and Ponta Delgada, Azores (86), were also screened for the c.156_157insAlu *BRCA2* mutation. Additionally, predictive testing was performed in four individuals from two additional families (two relatives from each family living in Rhode Island, USA, and in Villejuif, France, respectively) with the c.156_157insAlu *BRCA2* mutation identified elsewhere.

For the purpose of haplotype studies and age estimation of the c.156_157insAlu *BRCA2* mutation, the 14 HBOC families we previously reported (Peixoto *et al.*, 2009a) and the family (four c.156_157insAlu carriers) initially identified by Teugels *et al.* (2005) were also included. The geographic origin of the c.156_157insAlu *BRCA2* positive families was inferred from the birthplace of the oldest carrier or of the oldest family member most likely to be a carrier.

Screening for the c.156_157insAlu *BRCA2* mutation

The detection of the c.156_157insAlu *BRCA2* mutation was performed at the Department of Genetics of IPO-Porto for the suspected HBOC families from Portugal, for those originating from the Athens, Barcelona, Madrid and Zaragoza labs, as well as for predictive testing of four individuals from two additional families leaving in Rhode Island and Villejuif, respectively. The remaining cases were analyzed at the respective labs (except the cases from Rio de Janeiro, which were analyzed in Toronto) using the same protocol and a positive control provided by the Portuguese lab.

Screening for the c.156_157insAlu *BRCA2* mutation was performed using two independent PCRs (Peixoto *et al.*, 2009a; Peixoto *et al.*, 2009b), one for exon 3 amplification and another specific for the Alu rearrangement. Using this strategy, we expect two bands in positive cases in the first PCR (one band if negative) and one band in the second PCR (none if negative). The second PCR helps to control the first PCR for eventual problems with preferential amplification of the shorter fragment (wild type), whereas the first PCR controls for eventual absence of amplification in the second PCR. This strategy of two independent PCRs, followed by sequencing of the genomic fragments in positive cases, allows the unambiguous detection of the c.156_157insAlu *BRCA2* mutation (Peixoto *et al.*, 2009a; Peixoto *et al.*, 2009b). Positive and negative controls were used in all runs and all positive cases were confirmed in a second independent sample.

Real-time RT-PCR analysis

Primers and probes for the transcripts *BRCA2* wild type (*BRCA2*-wt) and *BRCA2* lacking exon 3 (*BRCA2*- Δ ex3) were designed with Primer Express 2.0 (Applied

Biosystems, Foster City, USA). The forward primer for the *BRCA2*-wt transcript was TGAAATTTTAAAGACACGCTGCAA and the reverse primer was GGTTTCGTAATTGTTGTTTTTATGTTTCAG. The forward primer for the *BRCA2*- Δ ex3 transcript was TCCAAAGAGAGGCCAACATTTT and the reverse primer was TTCACTGTGCGAAGACTTTTTATGTC. The probes were 5'-FAM-AAGCAGATTTAGGACCAATA-MGB-3' for the normal transcript (exon 2-3) and 5'-FAM-AACAAAGCAGGAAGGAAT-MGB-3' for the *BRCA2*- Δ ex3 transcript (exon 2-4). Primers and probes for the *BRCA2* and *GUSB* (endogenous control) genes were purchased from Applied Biosystems. PCR reactions were performed in a 20 μ l volume reaction containing 9 μ l of synthesized cDNA, 10 μ l of TaqMan universal PCR master mix, 1 μ M of each primer and 0.25 μ M of probe. PCR was performed in separate wells for each primer/probe set and each sample was run in triplicate. PCR parameters were as follows: 50°C for 2 min, 95°C for 10 min, followed by 50 cycles at 95°C for 15 s and 60°C for 1 min. Each plate included multiple non-template controls and serial dilutions of a control to construct the standard curve of each transcript and endogenous control. To determine the relative expression levels of the target transcripts in each sample, the comparative C_T method was performed as described by Schmittgen and Livak (2008). The relative expression of the transcripts in two different groups (that included ten carriers and eight controls) was calculated using the $2^{-\Delta CT}$ method. The ratio $2^{-\Delta CT_{BRCA2-\Delta ex3}}/2^{-\Delta CT_{BRCA2-wt}}$ was calculated for each sample. The Mann-Whitney U Test was used to compare the relative expression of those transcripts between the two groups. Statistical analysis was performed with SPSS version 11 and statistical significance was considered whenever $P < 0.05$.

Mutation-specific cumulative incidence

The cumulative incidence of breast cancer in women with the c.156_157insAlu *BRCA2* mutation was derived using the method of Kaplan and Meier, with unaffected individuals censored at the age of last follow-up or death without breast cancer. Only individuals shown to be carriers or obligate carriers were used for this calculation.

Microsatellite and SNP typing

Haplotype analysis was carried out in families in which the c.156_157insAlu *BRCA2* mutation was detected in at least one family member in addition to the proband. A total of 15 probands and 62 family members, including the three informative families previously reported (Peixoto *et al.*, 2009a) and the one described by Teugels *et al.* (2005), were genotyped for polymorphic microsatellite markers flanking *BRCA2* as described (Peixoto *et al.*, 2009a). The D13S1699 marker was not included in this study because we could not rule out preferential allelic amplification. The physical distances of the genetic markers were derived from the National Center for Biotechnology Information (NCBI) Map Viewer (genome build 36.3) (<http://www.ncbi.nlm.nih.gov/projects/mapview/>). All nine markers were assayed by PCR using fluorescently 5'-labeled primers. PCR products were run on an ABI PRISM 310 Genetic Analyser (Applied Biosystems) together with the fluorescence labeled DNA fragment size standard TAMRA.

Single-nucleotide polymorphism (SNP) markers were used to obtain a haplotype spanning ~1.1Mb encompassing the region between the D13S260 and D13S1695 microsatellite markers, where the first recombinant and/or mutational events were observed. In order to capture most of the genetic variation in this region and to avoid redundant SNP markers (i.e., markers in strong linkage disequilibrium), we performed Tag-SNP, namely Tagger Multimarker, using International HapMap Project CEPH

(Utah residents with ancestry from northern and western Europe) population data (www.hapmap.org). We developed SNaPshot assays for 19 SNP markers by multiplexed nucleotide primer extension reaction using dye label terminators (Applied Biosystems). The primers for multiplex amplification and single base extension (Supplementary Table 1) were designed using the online Primer-BLAST tool (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). AutoDimer (www.cstl.nist.gov/strbase/NIJ/AutoDimer.htm) was used to test for potential hairpin structures and primer dimers. The 19 SNPs were PCR amplified in four multiplex reactions with amplicon length between 100 bp and 450 bp. Amplification was carried out in a 9700 Thermocycler (Applied Biosystems). After a 95°C pre-incubation step for 5 min, PCR was performed in a total of 35 cycles using the following conditions: 95°C denaturation for 30 s, annealing at 55°C for 30 s and extension at 72°C during 30 s, followed by 10 min of final extension at 72°C. The multiplex SNaPshot reaction and capillary electrophoresis was done following the manufacturer's protocol (Applied Biosystems).

Haplotype construction and estimation of mutation age

Haplotype construction was performed manually based on the genotypes obtained of index cases and family members. We estimated the age of the c.156_157insAlu *BRCA2* mutation from the variation accumulated in their ancestral haplotypes, as described by Martins *et al.* (2007). This method takes into account both recombination (c) and mutation (μ) rates in the generation of variation. The probability of change per generation (ϵ) is given by $\epsilon=1-[(1-c)(1-\mu)]$, and the average of mutation and recombination events (λ) equals ϵt , where t is the number of generations. The recombination rate (c) was estimated from the physical distance between the two most

distant markers (D13S1700 and D13S267) using a conversion factor calculated in Rutgers Map Interpolator (<http://compgen.rutgers.edu/old/map-interpolator/>). The estimate of average mutation rate used was 7.8×10^{-4} (Gyapay *et al.*, 1994) for dinucleotides and two times lower for tetranucleotides.

Results

Detection of the c.156_157insAlu *BRCA2* mutation

Of the 149 Portuguese probands studied for germline mutations in the *BRCA1* and *BRCA2* genes at IPO-Porto, 11 patients presented the c.156_157insAlu *BRCA2* mutation (Figure 1) and 18 patients presented other deleterious mutations in either *BRCA1* (10 patients) or *BRCA2* (8 patients) genes (data not shown). Together with the 14 probands we previously reported with this mutation (Peixoto *et al.*, 2009a), a total number of 25 HBOC families have been uncovered at IPO-Porto at the time of writing. Altogether, 68 individuals from these 25 HBOC families have so far been tested for the c.156_157insAlu *BRCA2* mutation and 39 of them were shown to be carriers. The known geographic origins of all the c.156_157insAlu *BRCA2* positive families are shown in Supplementary Figure 1. Although most of the families are originated from northern/central Portugal, most likely reflecting our target population for genetic testing, we also detected the c.156_157insAlu mutation in families from southern Portugal and Madeira Island.

Of the 5,291 suspected HBOC families with no known deleterious mutation originated from other countries, only one proband tested in Clermont-Ferrand was shown to carry the c.156_157insAlu *BRCA2* mutation. Interestingly, this patient belongs to a family of Portuguese origin living in France. Additionally, the two relatives living

in Rhode Island (family with origin in Mangualde, central Portugal) and the two relatives living in Villejuif (family with origin in Porto, Portugal) for whom we performed predictive testing, were carriers of the c.156_157insAlu *BRCA2* mutation that had previously been identified elsewhere in Portuguese family members. Finally, the patient originally reported by Teugels *et al.* (2005) belongs to a Portuguese family originally from the region of Guarda (central Portugal).

Quantitative transcript analysis

Real-time RT-PCR showed quantitative differences between the full length and the *BRCA2*- Δ ex3 transcripts in c.156_157insAlu *BRCA2* mutation carriers and controls. The relative expression of the *BRCA2*- Δ ex3 transcript was six fold higher in carriers compared with controls, whereas a three fold decrease was observed for the *BRCA2*-wt transcript in patients compared with controls (Figure 2). The difference observed between patients and controls was statistically significant ($P = 0.00032$).

Mutation-specific cumulative incidence

Using the method of Kaplan and Meier, the cumulative incidence of breast cancer in women carrying the c.156_157insAlu *BRCA2* mutation was 90% until the age of 60 years (Figure 3).

Ancestral STR-based haplotypes and age estimate

Nine different haplotypes were phased for 11 out of the 15 families, three of them reported earlier (Peixoto *et al.*, 2009a). The results of the haplotype analyses for the 11 informative families are shown in Table 1 and the most parsimonious relationships among flanking haplotypes are presented as a phylogenetic network in

Figure 4. The probability of mutation *versus* recombination was evaluated, considering the minimum number of stepwise mutations. In the 11 informative families, SNP haplotypes were constructed in order to establish if a specific microsatellite was different from the consensus because of a recombination event rather than a mutation (Supplementary Figure 2).

Based on the mutation and recombination events observed in microsatellite haplotypes and assuming a generation time of 25 years, the age estimate for the c.156_157insAlu *BRCA2* mutation is 558 ± 215 years (Table 1).

Discussion

The c.156_157insAlu *BRCA2* mutation has so far only been reported in HBOC families of Portuguese origin. We here show that this rearrangement accounts for 57.8% of the *BRCA2* mutations and 37.9% of all deleterious mutations in HBOC families originating mostly from northern/central Portugal. This study confirms our and other earlier findings indicating that this is by far the most common BRCA mutation in Portuguese families with hereditary predisposition to breast/ovarian cancer, being detected in about 8% of all probands tested and presenting a nation-wide distribution (Machado *et al.*, 2007; Peixoto *et al.*, 2009a). This high frequency makes it cost-effective to test specifically for this rearrangement prior to screening the entire coding regions of *BRCA1* and *BRCA2* in suspected HBOC families from Portugal or with Portuguese ancestry. Furthermore, complementing earlier data showing that the c.156_157insAlu *BRCA2* mutation originates skipping of exon 3 (Teugels *et al.*, 2005) and that residual exon 3 skipping in *BRCA2* mRNA can be found in negative controls (Peixoto *et al.*, 2009a; Peixoto *et al.*, 2009b), we here demonstrate by real-time RT-PCR that carriers present significantly more *BRCA2*- Δ ex3 transcripts and much less full

length transcripts than controls, indicating that disruption of alternative transcript ratios is the mechanism causing hereditary breast/ovarian cancer associated with this *BRCA2* rearrangement. We further show that the cumulative incidence of breast cancer in c.156_157insAlu *BRCA2* mutation carriers does not differ from that of other *BRCA2* and *BRCA1* pathogenic mutations in our population (data not shown) or elsewhere (Ford *et al.*, 1998), further strengthening its role as the major contributor to hereditary predisposition to breast cancer in Portugal.

Since the c.156_157insAlu *BRCA2* mutation had only been reported in HBOC families of Portuguese origin (Teugels *et al.*, 2005; Machado *et al.*, 2007; Peixoto *et al.*, 2009a) and is not detectable with commonly used screening methodologies, one can not exclude that it is present in other populations until it is specifically sought. To further evaluate whether or not it constitutes a population-specific founder mutation, we screened for the c.156_157insAlu *BRCA2* rearrangement outside Portugal in more 5,291 suspected HBOC families with no known deleterious *BRCA1/BRCA2* mutations coming from several countries mainly from Europe, but also from Asia and North and South America. In addition to the family identified in Belgium by Teugels *et al.* (2005), we now detected this mutation in one proband living in France and in four individuals requesting predictive testing living in France and in the USA, all having in common the fact that they are relatively recent immigrants of Portuguese origin in those countries. Interestingly, c.156_157insAlu *BRCA2* mutation was not detected in 1.209 suspected HBOC families from Spain, including those from Galicia, the Spanish region with which Portugal shares more linguistic and cultural links, as also demonstrated by our recent finding of a common ancestry for the Portuguese HBOC families presenting the R71G *BRCA1* founder mutation of Galician origin (Santos *et al.*, 2009).

The findings we here present indicate that, within the relatively large sample population studied, the c.156_157insAlu *BRCA2* mutation is unique to HBOC of Portuguese ancestry, a fact that is hardly compatible with the age of about 2500 years previously estimated by Machado *et al.* (2007). Although geographic distribution of mutations is only an indirect measure of mutation age, more widespread mutations tend to be older than mutations showing a regional distribution. In order to get a more accurate mutation age estimate of the c.156_157insAlu *BRCA2* rearrangement, we performed an extensive haplotype analysis having in mind that the size of an ancestral haplotype around a mutation is inversely correlated with the number of generations separating the common ancestor from the families carrying that rearrangement. After performing the haplotype reconstruction in the 11 informative families and assuming a generation time of 25 years, we estimate the age of the c.156_157insAlu *BRCA2* mutation to be 558 ± 215 years, that is, most likely well after Portugal became politically independent (in 1143). Our estimate is consistent with the widespread distribution of the mutation in Portugal (Machado *et al.*, 2007; Peixoto *et al.*, 2009a; present report), the country demographic history (the North has been and still is consistently the source of migrants to the South), its occasional finding in countries with strong Portuguese immigration, and with its absence in the other populations studied (*e.g.* absence of the mutation in Spain, namely in Galicia). Nevertheless, statistical methods for estimating mutation ages are relatively crude (Rannala and Bertorelle, 2001), are dependent on sample representativeness, and estimate only the age of the common ancestor to the informative families that have been identified. The older age estimate advanced by Machado *et al.* (2007) was based upon a different sample of Portuguese patients (mostly from Center and South) and using a different age estimate method. However these authors recognize that the age of the mutation may be «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». On the other hand, although the mutation has so far only been detected in Portugal and in a few families with Portuguese ancestry living in Belgium, France or the USA, we can not conclusively exclude its presence in other countries that have strong historical links with Portugal, such as those having Portuguese as official language (Brazil, Angola, Mozambique, Cape Verde, Guinea-Bissau, São Tomé and Príncipe, East Timor, and Macau) or other countries with a large community of Portuguese immigrants. In fact, one of our probands with the c.156_157insAlu *BRCA2* mutation illustrates well this possibility: although she is now living in Portugal, her ancestors originating from North Portugal had moved several generations ago to Brazil and later to Angola, where reportedly various affected relatives lived.

In conclusion, we showed that the c.156_157insAlu *BRCA2* rearrangement is a Portuguese founder mutation originated about 558 ± 215 years ago, accounting for the majority of the *BRCA2* mutations and for about one third of all deleterious germline mutations in Portuguese HBOC families. We therefore recommend that all suspected HBOC families from Portugal or with Portuguese ancestry are specifically tested for this rearrangement, ideally prior to screening the entire coding regions of *BRCA1* and *BRCA2*. Furthermore, we showed that disruption of alternative transcript ratios is the mechanism causing hereditary breast/ovarian cancer associated with this *BRCA2* rearrangement and that the cumulative incidence of breast cancer in c.156_157insAlu *BRCA2* mutation carriers does not differ from that of other *BRCA2* and *BRCA1* pathogenic mutations.

Acknowledgements

This study was supported by Ministério da Saúde (Project N° 15/2007).
IPATIMUP was funded by Fundação para a Ciência e Tecnologia, through POCI
(Programa Operacional Ciência e Inovação 2010).

References

Diez, O., Gutierrez-Enriquez, S., Ramon y Cajal, T., Alonso, C., Balmana, J., and Llorc, G. (2007). Caution should be used when interpreting alterations affecting the exon 3 of the BRCA2 gene in breast/ovarian cancer families. *J Clin Oncol* 25, 5035-5036.

Fackenthal, J.D., and Olopade, O.I. (2007). Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer* 7, 937-948.

Ferla, R., Calo, V., Cascio, S., Rinaldi, G., Badalamenti, G., Carreca, I., Surmacz, E., Colucci, G., Bazan, V., and Russo, A. (2007). Founder mutations in BRCA1 and BRCA2 genes. *Ann Oncol* 18 Suppl 6, vi93-98.

Filippini, S., Blanco, A., Fernandez-Marmiesse, A., Alvarez-Iglesias, V., Ruiz-Ponte, C., Carracedo, A., and Vega, A. (2007). Multiplex SNaPshot for detection of BRCA1/2 common mutations in Spanish and Spanish related breast/ovarian cancer families. *BMC Med Genet* 8, 40.

Ford, D., Easton, D.F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D.T., Weber, B., Lenoir, G., Chang-Claude, J., Sobol, H., Teare, M.D., Struwing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T.R., Tonin, P., Neuhausen, S., Barkardottir, R., Eyfjord, J., Lynch, H., Ponder, B.A., Gayther, S.A., Zelada-Hedman, M., et al. (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 62, 676-689.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993-94 Genethon human genetic linkage map. *Nat Genet*, pp. 246-339.

Machado, P.M., Brandao, R.D., Cavaco, B.M., Eugenio, J., Bento, S., Nave, M., Rodrigues, P., Fernandes, A., and Vaz, F. (2007). 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* 25, 2027-2034.

Martins, S., Calafell, F., Gaspar, C., Wong, V.C., Silveira, I., Nicholson, G.A., Brunt, E.R., Tranebjaerg, L., Stevanin, G., Hsieh, M., Soong, B.W., Loureiro, L., Durr, A., Tsuji, S., Watanabe, M., Jardim, L.B., Giunti, P., Riess, O., Ranum, L.P., Brice, A., Rouleau, G.A., Coutinho, P., Amorim, A., and Sequeiros, J. (2007). Asian origin for the worldwide-spread mutational event in Machado-Joseph disease. *Arch Neurol* 64, 1502-1508.

Peixoto, A., Salgueiro, N., Santos, C., Varzim, G., Rocha, P., Soares, M.J., Pereira, D., Rodrigues, H., Bento, M.J., Fraguas, A., Moura, G., Regateiro, F., Castedo, S., and Teixeira, M.R. (2006). BRCA1 and BRCA2 germline mutational spectrum and evidence for genetic anticipation in Portuguese breast/ovarian cancer families. *Fam Cancer* 5, 379-387.

Peixoto, A., Santos, C., Rocha, P., Pinheiro, M., Principe, S., Pereira, D., Rodrigues, H., Castro, F., Abreu, J., Gusmao, L., Amorim, A., and Teixeira, M.R. (2009a). The

c.156_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. *Breast Cancer Res Treat* 114, 31-38.

Peixoto, A., Santos, C., Rocha, P., Pinto, P., Bizarro, S., and Teixeira, M.R. (2009b). Molecular diagnosis of the Portuguese founder mutation BRCA2 c.156_157insAlu. *Breast Cancer Res Treat* 117, 215-217.

Rannala, B., and Bertorelle, G. (2001). Using linked markers to infer the age of a mutation. *Hum Mutat* 18, 87-100.

Santos, C., Peixoto, A., Rocha, P., Vega, A., Soares, M.J., Cerveira, N., Bizarro, S., Pinheiro, M., Pereira, D., Rodrigues, H., Castro, F., Henrique, R., and Teixeira, M.R. (2009). Haplotype and quantitative transcript analyses of Portuguese breast/ovarian cancer families with the BRCA1 R71G founder mutation of Galician origin. *Fam Cancer* 8, 203-208.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3, 1101-1108.

Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., and De Greve, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum Mutat* 26, 284.

Figure legends

Figure 1 - Molecular diagnosis of the *BRCA2* c.156_157insAlu mutation using two independent PCR analyses, showing positive cases in lanes 1 and 2 and a negative case in lane 3. (A) PCR specific for *BRCA2* exon 3, showing an additional band resulting from the insertion of a DNA fragment of about 350bp long within exon 3 of *BRCA2* (c.156_157insAlu mutation; lane 1 and 2). (B) PCR specific for the c.156_157insAlu *BRCA2* mutation with positive cases being then confirmed by sequencing the amplified genomic fragment. Lane 4 corresponds to a positive control and NTC is a non template control. MW refers to 100 bp DNA standard.

Figure 2 - Real-time RT-PCR quantification of the altered transcript ratios in *BRCA2* c.156_157insAlu carriers as compared with controls. The relative expression of the *BRCA2*- Δ ex3 transcript was six fold higher in carriers compared with controls, whereas a three fold decrease was observed for the *BRCA2*-wt transcript in patients compared with controls.

Figure 3 - Cumulative incidence of breast cancer among c.156_157insAlu *BRCA2* germline mutation carriers, reaching 90% at about 60 years of age.

Figure 4 - Phylogenetic network showing the most parsimonious relationships among flanking short tandem repeat-based haplotypes in families carrying the c.156_157insAlu *BRCA2* mutation. Circle and line sizes are proportional to the number of families and stepwise mutations, respectively. Dashed diamonds indicate recombination events.

Supplementary Figure 1 - Map of Portugal showing the known geographical origin of the families with the c.156_157insAlu *BRCA2* germline mutation. The black circles indicate the origin of the 25 families detected in Portugal (present report and those reported in Peixoto *et al.*, 2009a), the open circle the origin of the family previously identified by Teugels *et al.* (2005) in Belgium, and the triangles the origin of the families of the four individuals (two from each) subjected to predictive testing living in Rhode Island, USA, and in Villejuif, France, respectively.

Supplementary Figure 2 - SNP marker haplotype spanning ~1.1Mb, encompassing the region between the D13S260 and D13S1695 microsatellite markers, in the 11 informative families.

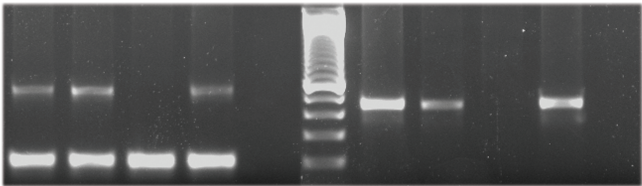
Table 1- Age estimation of the ancestral c.156_157insAlu *BRCA2* mutation.

Family ID	Haplotype^a	Families, No.	Mutation steps/ recombination events, No.	Age$\pm\delta$, y^b
4001; 6216	H1: 317-160-156-299-230-242-228-144-158	2	0	558 \pm 215
5931	H2: 317-160-156-295-230-242-228-144-158	1	1	
6045	H3: 313- 160-156-299-230-242-228-144-158	1	1	
14230	H4: 269- 160-156-299-230-242-228-144-158	1	1	
7772; 14331	H5: 317-160-156-299-230-234-234-148-144	2	1	
14459	H6: 321- 160-156-299-230-234-234-148-144	1	2	
3854	H7: 317-160-156-307-226-252-234-148-144	1	1	
12703	H8: 317-160-156-299-230-252-234-144-156	1	1	
3286	H9: 309-162-160- 299-230-242-228-144-158	1	1	
Total		11	9	

a The nine microsatellite markers used were: D13S1700, D13S260, D13S1698, D13S1701, D13S171, D13S1695, D13S1694, D13S310 and D13S267 (from left to right). Alleles conserved in each population appear in bold.

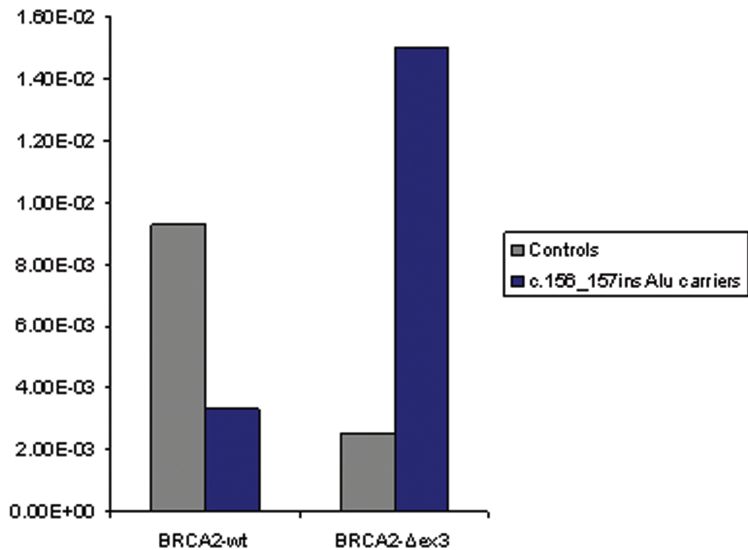
b The recombination rate (*c*) was based on the physical distance between the two most distant markers (1930.8 kilobases; *c*=0.030597 cM) using a conversion factor calculated in Rutgers Map Interpolator. The estimated probability of mutation per generation and per haplotype was 0.00624 (as seven dinucleotide and two tetranucleotide short tandem repeats were studied).

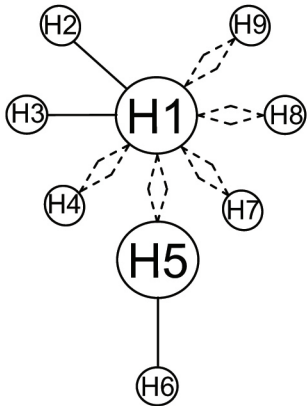
1 2 3 4 NTC MW 1 2 3 4 NTC

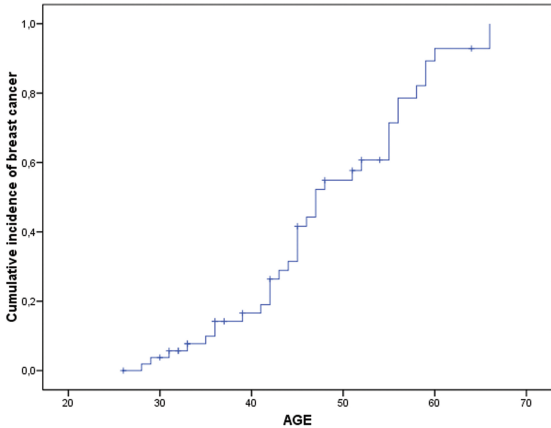


A

B







Supplementary table 1 – SNaPshot assay primers used for SNP genotyping around *BRCA2*.

	SNP ID	PCR amplification primers	Single base extension primers	Base change
1	RS379693-F RS379693-R	ACCAGGCACGCATGTAATTC TGCCCAGCCAGAAACTTAAT	ACAGTTGAAAAAGTACTTTCA	A>T
	RS9590624-F RS9590624-R	TTGGGAGGGAGAGCACTAGA TGGGCTCCTACATCCAACCTC	(GACT) ₂ GAAAGGTTATTGGAGACCTGC	A>T
	RS916732-F RS916732-R	TCACAGAGCTGTGCAAACC GCGGAATACCATCAACCATC	(GACT) ₃ TAGAAAGTAAATGTTTGAAAATT	C>T
	RS2806638-F RS2806638-R	TTCATCCAACCCTTCCAGTC AGCCACAGCAGTGAGAAAAT	(GACT) ₅ ATTTTCTTCCTTTCTTTACAC	A>G
	RS206102-F RS206102-R	AAATGAAGCCTCTGTGACAAAA CCAGGGTGTGAGAGGAAAAA	(GACT) ₇ GTTTGCATCTTTTTTGAAGG	C>G
2	RS9567445-F RS9567445-R	GCTATGAGGCATAGTCAGCG GCATTTGATGTGGTTAGCATT	GGGAAACTCAGAAATTCAGT	A>T
	RS798972-F RS798972-R	ACGGCCACCATCTAACAGAC GCAGACGCCTCCTCTCTCTA	GACTCAATTCACCAAAATTTTACCA*	A>T
	RS703219-F RS703219-R	TTAGTGGCCCTGTTCTTGCT AAAAACACCTGGCACAAAGC	(GACT) ₃ GATAGTATTCCATTCCATGTCA*	C>T
	RS582274-F RS582274-R	TCTCACCTGGCTGAAACTCC CTGCTTAGCAACAGGAAGGG	(GACT) ₄ GATTGAAAATAGTTTGATCTCTTT	C>T
	RS176059-F RS176059-R	AAGAAAAAGGGAGAAGTTTGCC TAAACAAATCTGCCCCATC	(GACT) ₆ TTATATTTGAGTCTAATTTGG	A>G

Supplementary table 1 (continued)

Multiplex	SNP ID	PCR Amplification primers	Single base extension primers	Base change
3	RS1207952-F RS1207952-R	TGCATGAAGCACACTGTGAA CTCCTTGCTAGCCTCAGGG	GTTCAGCAGTTCTCTTTCAG	A>C
	RS2761367-F RS2761367-R	AGCTGAGGATGGCTGAAAAA GACAAGCAGCCTAAGGGAAG	GACTTCATTTTACTTTTTGTTTTGTTT*	G>T
	RS207632-F RS207632-R	GCAACGTTACAGAGCAGTTTG ATGCTCCATCTCCACGTTTC	(GACT)₄GAGCATCATGTCAATCAATG	C>G
	RS6561643-F RS6561643-F	CCTCAGCAATCCTGTGTGTG GGGACTGGAGGATTTTCTCA	(GACT)₅GATGTTAGCTTGTTTTCCAAGT*	A>T
	RS762900-F RS762900-R	GGCAAAAGGCAGTGCTAGTT TCTATACCAATGATGAGCAATCTT	(GACT)₆GACTGAATAGGCTTAAAATCTGAA*	A>G
4	RS9595946-F RS9595946-R	TGCCTTTCCCACTCGTTAGT GGCTGAATTGAGTGATGGGT	TCTGAAAAGATTTGTGTTAATA*	A>G
	RS472873-F RS472873-R	TTTCTGCATTGTGACTCTGCTT CACGGTAAAGGTTTCTGGGA	(GACT)₂AACCAACTTCCTTCGTATTT	T>C
	RS2146994-F RS2146994-R	TCTCACCTGTGGGACCAAAT TAGTGCTGTTCCCTGGGGAC	(GACT)₅GATGGAGCTTTGGCCTCTTG	A>C
	RS7983065-F RS7983065-R	TGATGAGACCTGGTAGTCTGTAATG CCCTAGCAATTCTTCTGTATTTG	(GACT)₆GAAAAAAATGATTTCCCTGAAA	C>T

D13S1695

rs656143

rs7983065

rs762900

rs9595946

rs207632

rs2146994

rs472873

rs2761367

rs1207952

BRCA2

rs206102

rs2806638

rs9567445

rs703219

rs916732

rs798972

rs176059

rs379693

rs9590624

rs582274

D13S260

A

C

G

G

G

A

C

T

C

G

G

T

T

C/T

A

A

A

T

C

A

C

G

G

G

A

C

T

C

G

G

T

T

T

A

A

A

T

C

A

C/T

A/G

A/G

G

A

C

T

C

G

G

T

T

T

A

A

A

T

C

A

C

G

G

G

C/A

C

G/T

C

G

G

T

T

T

A

A

T

A

C

A

C

G

G

G

A

C

T

C

G

G

T

T

T

A

A

A

T

C

A

C

G

G

G/C

A

C

T

C

G

G

T

T

T

A

A

A

T

C/T

T/A

C

A

G

C

A

C

T

C

G

G

T

T

C/T

A

A/T

A/T

T

C

A

C

G

G

G

A

C

T

C

G

G

T

C

T

A/T

A

A

A

C/T