

Manuscript title The highly prevalent *BRCA2* mutation c.2808_2811del (3036delACAA) is located in a mutational hotspot and has multiple origins.

Short Title: *BRCA2* 3036delACAA is located in a mutation hotspot

Mar Infante,¹ Mercedes Durán,¹ Alberto Acedo,² Eva María Sánchez-Tapia,³ Beatriz Díez-Gómez,² Alicia Barroso,⁴ María García-González,⁵ Lúdia Feliubadaló,⁶ Adriana Lasa,⁷ Miguel de la Hoya,⁸ Eva Esteban-Cardenosa,⁹ Orland Díez,¹⁰ Cristina Martínez-Bouzas,¹¹ Javier Godino,¹² Alexandre Teulé,¹³ Ana Osorio,⁴ Enrique Lastra,⁵ Rogelio González-Sarmiento,³ Cristina Miner¹ and Eladio A. Velasco^{2,*} .

¹ Cancer Genetics, Instituto de Biología y Genética Molecular (UVa-CSIC), Valladolid, Spain.

² Splicing and Cancer, Instituto de Biología y Genética Molecular (UVa-CSIC), Valladolid, Spain.

³ Laboratorio de Diagnóstico Genético en Cáncer Hereditario. Instituto de Biología Molecular y Celular del Cáncer. Universidad de Salamanca-CSIC, Salamanca, Spain.

⁴ Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, Madrid, Spain and Spanish Network on Rare Diseases (CIBERER).

⁵ Servicio de Oncología. Complejo Hospitalario de Burgos, Burgos, Spain.

⁶ Unitat de Diagnòstic Molecular. Programa de Càncer Hereditari. Institut Català d'Oncologia-IDIBELL. Hospital Duran i Reynals, L'Hospitalet de Llobregat, Spain.

⁷ Servicio de Genética. Hospital de la Santa Creu i Sant Pau. CIBERER (U-705). Barcelona, Spain.

⁸ Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos, Madrid, Spain.

⁹ Laboratorio de Biología Molecular, Servicio de Biopatología Clínica, Hospital Universitario La Fe, Valencia, Spain. Current address: Cancer Genetics, Instituto de Biología y Genética Molecular (UVa-CSIC), Valladolid, Spain.

¹⁰ Oncogenetics Laboratory. University Hospital Vall d'Hebron. Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain

¹¹ Servicio de Bioquímica, Hospital de Cruces, Barakaldo, Bizkaia, Spain.

¹² Instituto Aragonés de Ciencias de La Salud; Servicio de Oncología Médica; Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain.

¹³ Unitat de Consell Genètic. Programa de Càncer Hereditari. Institut Català d'Oncologia-IDIBELL. Hospital Duran i Reynals, L'Hospitalet de Llobregat, Spain.

* To whom correspondence should be addressed. Tel+34 983184829; Fax: +34 983184800 Email: eavelsam@ibgm.uva.es

Abstract

BRCA2-c.2808_2811del (3036delACAA) is one of the most reported germline mutations in non-Ashkenazi breast cancer patients. We investigated its genetic origin in 51 Spanish carrier families that were genotyped with eleven 13q polymorphic markers. Three independent associated haplotypes were clearly distinguished accounting for 23 (WCL), 20 (ECL) and 6 families (SoS). Mutation age was estimated with the DMLE software in a range of 45-68 and 45-71 generations for WCL and ECL haplotypes, respectively. The most prevalent variants, c.2808_2811del and c.2803G>A, were located in a double-hairpin loop structure (c.2794-c.2825) predicted by Quikfold that was proposed as a mutational hotspot. To check this hypothesis, random mutagenesis was performed over a 923-bp fragment of *BRCA2* and 86 DNA variants were characterized. Interestingly, three mutations reported in the mutation databases (c.2680G>A, c.2944del and c.2957dup) were replicated and 20 affected the same position with different nucleotide changes. Moreover, five variants were placed in the same hairpin-loop of c.2808_2811del, and one affected the same position (c.2808A>G). In conclusion, our results support that at least three different mutational events occurred to generate c.2808_2811del. Other highly prevalent DNA variants, such as *BRCA1*-c.68_69delAG and *BRCA2*-c.5946delT and c.8537delAG, are concentrated in hairpin-loops suggesting that these structures may represent mutational hotspots.

Introduction

Hereditary predisposition accounts for 5–10% of all breast cancer cases [1;2], 15%-20% of which are caused by mutations in the two major susceptibility genes *BRCA1* (MIM# 113705) and *BRCA2* (MIM# 600185)[3]. More than 3,500 distinct DNA changes in *BRCA1* and *BRCA2* sequences have been characterized in the reference databases (The Breast Cancer Core Database, BIC database and the Universal Mutation Database, UMD) [4;5], which are spread throughout their entire coding and flanking intronic regions. Only a small number of DNA variations are predominant in some specific ethnic groups as a consequence of a founder effect originated in a common ancestor [6]. Spanish population has shown a specific distribution of founder mutations reflecting the existence of populations substructures in the Iberian Peninsula [7-10].

According to the BIC database, c.2808_2811del (p.Ala938Profs*21) is the second most reported frameshift *BRCA2* mutation, and the most prevalent mutation in Castilla y León (Spain) [9-11]. This mutation has been reported in almost all Western European countries [12] and in African women with premenopausal breast cancer [13]. Furthermore, the same mutation was found *de novo* in a breast cancer case that suggested that these *BRCA2* nucleotides were a mutation hotspot [14].

Knowledge of the molecular origin of a mutation is critical to the understanding of human disease and evolution. Mutations in humans may be caused by exogenous mutagens or endogenous mechanisms [15]. Mutation rates can vary substantially across genomic regions, so that they exceed the average rate at some extremely mutable sites (mutation hotspots) [16]. Moreover, a previous study of the secondary structure of the DNA showed that the most hypermutable bases of the *TP53* tumor suppressor gene were located in stable DNA stem-loop arrangements [17]. In this study we set out to verify the presence of a possible unique founder effect of mutation c.2808_2811del or several independent mutational events occurred. This feature could be relevant in populations where the pre-screening of specific mutation panels is a cost-effective strategy prior the *BRCA1/BRCA2* mutation scanning. To address this issue, we performed haplotype analysis in the region surrounding the *BRCA2* gene in 51 carrier families and then, we estimated the mutation age according to each haplotype and its geographical distribution.

Moreover, a DNA secondary structure prediction and a random mutagenesis experiment were performed to identify mutational hotspots.

Materials and methods

Patients

Fifty-one unrelated families harboring *BRCA2*-c.2808_2811del from ten Spanish familial cancer centers were included in this study (Table 1). Patient data regarding age of diagnosis, current age or age at death, familial cancer history and geographical origin were collected. After written informed consent DNA samples were obtained from at least one affected member of each family. To date 23 c.2808_2811del families were identified in the IBGM laboratory after mutation scanning by CSGE (Conformational Sensitive Gel Electrophoresis) or HA-CAE (Heteroduplex Analysis by Capillary Array Electrophoresis) and confirmed by DNA sequencing as described elsewhere [18;19], but only sixteen affected families were included in this study. The screening of the rest of the patients was performed at the respective participant laboratories with different methods [10]. All mutation carriers from other laboratories were further confirmed at the IBGM. The control population (75 anonymous and unrelated samples) used was obtained from the National DNA Bank, a collection of representative DNA samples of the Spanish population [20]. This study was approved by the Ethics committee of the Faculty of Medicine, University of Valladolid (Spain).

Genotyping

A total of 221 DNA samples representing 51 index cases, 95 family members and 75 controls were typed with seven short tandem repeats (STRs) markers and four *BRCA2* SNPs covering 1.34 Mb of chromosome 13 surrounding the c.2808_2811del *BRCA2* mutation. Primer sequences, order of markers and physical distances among D13S260, D13S1699, D13S1698, D13S1697, D13S171, D13S1695 and D13S1694 were obtained from the Ensembl Genome Browser [21]. Polymerase chain reaction (PCR) products of STRs were generated using fluorescently labelled primers and subsequently electrophoresed on an ABI3130 sequencer (Applied Biosystems, Carlsbad, CA). Analysis was performed with the GeneMapper software v.3.7 (Applied Biosystems). Additionally, four *BRCA2* SNPs: rs766173 (c.865A>C), rs144848 (c.1114C>A), rs1801406 (c.3396A>G) and rs543304 (c.3807T>C) were typed by TaqMan assays in a 7500 Real Time PCR System (Applied Biosystems).

Estimation of the mutation age

We used the DMLE method [22] to estimate the mutation age. For this purpose, genotypes from 75 control individuals were submitted to the Phase v2.1 software to predict their haplotypes and estimate their frequencies (Supplemental Table S1). Then the phased data together with the inferred consensus carrier haplotype of 46 families and the two more prevalent haplotypes were independently introduced into the DMLE program, which calculated the mutation age. Population growth rate (r) was estimated by the formula $t_1 = t_0 * e^{gr}$, where g is the number of generations between two time points, t_1 is the population size in 2010 (47 million) and t_0 the ancestral population size. We used demographic data (INE: Population Census) [23] of two different years for t_0 : 1594 (~8.2 million) or 1857 (~15 million), and g was estimated in 16.6 and 6.12 generations, respectively (considering 25 years per generation). Thus, r was estimated to be 0.124 and 0.190 for 1594 and 1857 data, respectively.

Random mutagenesis

DNA of a control individual was amplified with Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA) and primers containing a 5'-tail with a restriction site (underlined) either for *EcoRI* (Forward, 5' CACACAGAATTCAGAGGCAAAGAATCATACAAA 3') or for *XhoI* (Reverse, 5' CACACACTCGAGTTGGAAAATAACATCTGAGGG 3') (923 bp, c.2359-c.3281). After restriction enzyme digestion, insert and plasmid pBluescript II KS (Stratagene, La Jolla, CA) were ligated with the Fast-Link DNA Ligation Kit (Epicentre, Madison, WI) to transform *Escherichia coli* DH5 α cells. Random mutagenesis was carried out with the Genemorph II EZClone Domain PCR mutagenesis kit (Agilent, Santa Clara, CA, USA). A PCR fragment (megaprimer) of 419 bp (c.2610-c.3028) was generated with nested *BRCA2* primers 11E-Fw, 5'-TTCAAAAATAACTGTCAATCC-3', and 11F-Rev, 5'-TGAAGCTACCTCCAAAACACTGT -3', and the error prone Mutazyme-II DNA polymerase. Mutant PCR products serve as megaprimers for the EZClone reaction during which they are denatured and annealed to the original plasmid and extended with a high fidelity DNA polymerase. Then, the reaction was treated with *DpnI*, specific of methylated sequences, to eliminate the wild type plasmid that had been used as template. XL10-Gold ultracompetent cells (Agilent) were transformed with 1.5 μ L of the EZClone reaction following the manufacturer's protocol

and were plated in LB-agar (Sigma-Aldrich, St. Louis, MO) plates with ampicillin 100 µg/ml, X-Gal 40 µg/ml and IPTG 0.1 mM. Seventy-five different colonies were picked individually and grown in liquid LB medium to purify plasmids with the Pure Yield™ Miniprep system (Promega, Madison, WI). Each clone was sequenced with *BRCA2* primers 11E-Fw and 11F-Rev to characterize the generated mutations.

Sequence variation nomenclature

Nomenclature of DNA variants followed the current guidelines of the Human Genome Variation Society (HGVS) [24] where nucleotide +1 is the A of the ATG-translation initiation codon of the *BRCA2* cDNA sequence NM_000059.1.

Results

We have analyzed seven STR and four SNP markers in a total of 51 families, one of them of British origin, harbouring the c.2808_2811del mutation. Of these, at least two carrier relatives were tested in 37 families whilst only the index case was available in other 14 families. Three different haplotypes were assigned to c.2808_2811del positive families; except for families CN-428 and CN-565 that could not be successfully classified (Figure 1). The geographical origin of the families and the corresponding haplotypes are shown in Figure 2.

The most common haplotype was 157-144-163-225-A-A-A-T-195-253-234 (WCL haplotype) that accounted for 23 families, mainly from the West area of Castilla y León (8 out of 9 SA-families; Figure 2). A second haplotype 159-148-175-225-A-A-G-T-195-253-236 (ECL haplotype) was present in 20 pedigrees which included the largest sample from the East area of Castilla y León (12/16 carrier BU-families, Figure 2). Both haplotypes were very similar in their central core of 6 markers (Supplemental Table S1) but SNP c.3396A>G (rs1801406) was key to distinguish between them ("A"- WCL vs. "G"-ECL; Figure 1). Six other families, four of which were native from Southern Spain, were grouped in a third haplotype 157-144-163-221-A-C-A-T-[195 or 205]-251-228 (SoS haplotype).

The most prevalent disease in carriers was breast cancer with 62 cases (Table 1), nine of which were bilateral and two were male BC cases. The medium age for BC cases was 43 years (range 26-82) whereas for OC was 58 years (49-69). Remarkably, 70 carriers (45 females and 25 males) were asymptomatic, although 60% of them (41/70) were under 40 years old (Supplemental Table S2). It was noteworthy the different number of bilateral breast cancer cases observed (2/20 in WCL haplotype vs. 7/23 in ECL haplotype).

Dating the origin of the mutation

The results were population growth rate depending, and the oldest demographic data in Spain were in year 1594 but they were not reliable until 1857. Accordingly we used both records to calculate the Spanish population growth rate (r). The two main haplotypes and the complete set of 46 families with recognizable haplotype were submitted independently to DMLE analysis using the two r data (0.124 or 0.190; Supplemental Figure S1). Another key parameter in this software is the proportion of population sampled. The

frequency of *BRCA2* mutations in BOC patients is 1-3% [25], thus we obtained 0.000217 assuming that 2% of 47 million people in Spain is a mutation carrier. In the whole set of 46 Spanish families DMLE estimated a mutation age range between 72 and 110 generations ($r=0.124$) or 49-73 ($r=0.190$; Supplemental Figure S1). The 23 WCL families showed a range of 73-112 generations ($r=0.124$) or 45-68 generations ($r=0.190$) whereas the 20 ECL families resulted in a range of 67-105 generations ($r=0.124$) or 45-71 generations ($r=0.190$). Therefore, the mutation associated with ECL and WCL haplotypes could be dated on average around 55 generations ago (about 1375 years, in the VI-VII century during the Visigothic Kingdom; Supplemental Figure S1) when we used the most reliable estimation (0.190). Finally, mutation age of SoS haplotype could not be calculated as there had been collected only 6 families. However, all of them shared a common haplotype in 10 out of 11 markers suggesting a more recent origin than the other two haplotypes.

Random mutagenesis

The absence of a unique ancestral haplotype as well as the identification of one *de novo* c.2808_2811del [14] suggested that this sequence was a mutation hotspot. It has been reported that mutations occur non-randomly since they are preferentially located in vulnerable stem-loops structures [17]. The DNA secondary structure of a 419-nucleotide region spanning this mutation was predicted with the Quikfold software [26;27]. We found a total of 14 stem-loops structures (Supplemental Figure S2) where 82 mutations of the databases were mapped (Supplemental Table S3). Forty-two out of them (51.2%) were positioned in single stranded loops or bulges that represent 41.6% of the analysed sequence (157/377 nucleotides, excluding primer sequences), and 7 variants were placed in the same loop of mutation c.2808_2811del within a complex double hairpin-loop structure, sustaining that this structure may be a mutation hotspot.

With a view to investigating this hypothesis we designed a random mutagenesis experiment over a fragment surrounding c.2808_2811del. Seventy-five different clones were sequenced, 48 of which contained one (21 clones) or two or more (27 clones) DNA variants. We identified 86 different DNA changes, three of which were recurrent and were found twice in independent clones: c.2683G>A, c.2771A>G and c.2957dup (Supplemental Table S3). Seven mutations were single nucleotide deletions, one was a single nucleotide insertion (c.2957dup) and 78 were nucleotide substitutions. In addition, three mutations

reported in the mutation databases were replicated in the random mutagenesis experiment: c.2680G>A, c.2944del and c.2957dup (twice, see above). Other 20 variants affected identical positions of databases variants although they consisted of different nucleotide changes. Of all changes produced, 53/86 (61.6%) were located into the stem-loops predicted by Quikfold (Figure 3). Remarkably, random mutagenesis generated five mutations in the double hairpin-loop (#7) of the c.2808_2811del mutation, including a single nucleotide substitution in position 2808 (c.2808A>G). The highest incidence of mutations was found in hairpin 8 with 1 mutation every 1.1 nucleotide (Table 2). Hairpins 4, 5, and 7 and inter-sequence 4 were also especially vulnerable to mutations with 1 mutation every 1.7-1.8 nucleotides.

Discussion

Multiple origins of c.2808_2811del

We found three different haplotypes linked to the mutation, which were correlated with the geographic distribution of families. Patients from Castilla y León were distributed into two haplotypes, which basically corresponded to the West and East areas of this region. Both areas had differences in the mutational spectrum of the *BRCA* genes [11;28], supporting two different genetic backgrounds and, therefore, two independent mutational events. The alternative hypothesis is the existence of a unique genetic origin that might have been originated either by a double recombination event between c.2808_2811del-rs1801406-rs543304 or a *de novo* mutation at nucleotide c.3396 (A to G, rs1801406; Figure 1). Both options seem unlikely, since, first, rs1801406 was only 588 nucleotides downstream the deletion and 411 bp upstream rs543304 (totalling 999 bp), too close to undergo two crossovers, and, second, the mutation rate of single-nucleotide variants in the human genome has been recently estimated in 1.2×10^{-8} per base pair per generation [29], disregarding the hypothesis of a recent *de novo* event. A third haplotype (SoS) was distinguished in six families, four of which had origins in Southern provinces of Spain, suggesting a common founder effect. The deduced core haplotype (157-144-163-221-A-C-A-T-[195/205]-251-228) shared alleles at ten out of eleven markers in the six families suggesting a more recent origin. Previous reports of a “*de novo*” mutation in a Dutch [14] and an African breast cancer patients without familial antecedents [13] give further support to multiple historical occurrences of c.2808_2811del. Indeed, one of our patients (a 32 year-old BC patient) was suspected to carry a *de novo* c.2808_2811del mutation, because none of the parents was a mutation carrier. Unfortunately, the haplotype construction cannot be performed because DNA patient sample was depleted. All the data presented here support that at least three different mutational events took place to originate c.2808_2811del, so that this sequence could be a mutation hotspot.

Interestingly, c.2808_2811del was not previously reported in Sephardic Jewish breast cancer patients [30] despite almost 20% of Y-STR haplotypes of the Iberian Peninsula and Balearic Island are of Sephardic origin [31]. This might be due to the small sample size of the first report (only 16 patients) and the low frequency of the Sephardic ancestry in Castilla y León (East, 2.5%; West, 12.9%) where c.2808_2811del probably arose

(major haplotypes ECL and WCL). Surprisingly, a putative founder *BRCA1* mutation in Sephardic breast cancer patients, c.5123C>A (p.A1708E) [30], is precisely the most prevalent *BRCA1* mutation in Castilla y León (12.2% of *BRCA1*+ families, data not shown) and the rest of Spain (11.3% on average) [9], suggesting that its genetic origin requires further investigation.

In earlier reports we estimated mutation age with a method based on the linkage disequilibrium due to recombination between the nearest recombinant marker and the mutation [28;32;33]. This calculation was impossible to conduct in c.2808_2811del families because the ancestral allele was more frequent in non-carrier chromosomes than in mutation-carriers in all cases. In this report, we estimated by DMLE that the mutation arose in Spain around 58 generations ago (1450 years), dating the mutation back to the VI century during the Visigothic Kingdom, with a range between 49 and 73 generations, and began to spread to the rest of Spain. Mutation age was previously estimated in ~80 generations with the maximum likelihood method in eleven families from different countries, including Spain [12].

Mutation hotspot

It was reported that the most hypermutable nucleotides of the *TP53* gene are located in secondary structures called hairpin-loops and bulges [17]. These structures contain regions of single-stranded DNA with unpaired or mispaired bases that are more vulnerable to mutation because they are exposed to endogenous (e.g. metabolites or DNA polymerase errors) or exogenous mutagens [15]. Furthermore, Quikfold data illustrate a complex structure that the polymerase might not copy properly (Figure 3) leading to different DNA changes around this position [26;34]. Moreover, the most reported sequence changes in the BIC and UMD mutation databases between nucleotides c.2610-c.3028 were precisely this mutation (c.2808_2811del, 199 records in the BIC and UMD databases) and c.2803G>A (151 records), both in the single stranded portion of this loop. Random mutagenesis created 5 different mutations in the predicted arrangement, with even a single nucleotide substitution in position c.2808. We previously showed that two founder mutations of Castilla y León, *BRCA1* c.5153-1G>A and *BRCA2* c.5146_5149delTATG, were also located in such secondary structures [28]. Altogether these data, we can conclude that the double-hairpin structure of c.2808_2811del is a

mutation hotspot. It is also worth mentioning that hairpin-loops #4, #5 and #8 have a high incidence of reported and “artificial” mutations (Table 2), suggesting that they could also be hotspots. However, the mutagenesis experiment did not generate c.2808_2811del as Mutazyme-II polymerase only introduced insertions and deletions of one nucleotide despite we almost doubled the indel rate described with this polymerase (10.1% vs. 5.5%). In any case, we should be cautious when interpreting these results since this assay does not reproduce the physiological conditions: first, Mutazyme-II is a modified prokaryotic polymerase; second, the supercoiled structure of the plasmid where the template is cloned may influence this process; and, third, the genomic and nucleosomal contexts may also play a role in the natural phenomenon of mutagenesis. Twenty different variants (three of them in two independent clones) of the mutagenesis experiment affected the same positions of reported variants but they consisted of different nucleotide changes, and three mutations produced the same nucleotide-change: one single-nucleotide substitution (c.2680G>A), one insertion (c.2957dup in two independent clones) and one deletion (c.2944del). Curiously, none of these three variants was placed in hairpin-loops suggesting that their occurrence is not probably related to the secondary structure of DNA.

On the other hand, the most prevalent mutations of *BRCA1* (c.68_69del) and *BRCA2* (c.5946delT), both of Ashkenazi origin, are located in a hairpin-loop structure and in a complex hairpin arrangement with multiple loops, respectively (Figure 4). Interestingly, c.68_69del has been proven to have at least three founder effects [35;36] suggesting that this sequence is also a mutation hotspot. The observation of a recurrent *BRCA* mutation with multiple origins has also been documented in *BRCA2* c.8537delAG that was found in distinct populations with three differentiated haplotypes [37]. The authors suggested that the presence of repetitive sequences could allow for mismatching during replication. As expected, *BRCA2* c.8537delAG also lies in a large loop (Figure 4).

In conclusion, the presence of several haplotypes linked to c.2808_2811del indicated that three independent events occurred. This might be due to the presence of a mutation hotspot in this region that would justify its high frequency worldwide. This sequence lies within a double-hairpin-loop structure that we postulate as a mutation hotspot. The identification of recurrent mutations in a specific population allows the

redefinition of the genetic screening strategy as well as a more accurate estimation of the risk associated with this mutation.

Supplementary material

Supplementary Tables 1, 2 and 3 and Supplementary Figures 1 and 2 can be found at <http://carcin.oxfordjournals.org/>

Funding

Instituto de Salud Carlos III (PI10/2910 to E.A.V., ISCIIIRETIC: RD06/0020/1051 and PI10/01422 to L.F.); Junta de Castilla y León, Consejería de Educación (CSI004A10-2 to E.A.V.), Consejería de Sanidad (BIO39/VA27/10 to E.A.V.) and the Cancer Prevention Program; The European Social Fund and Consejería de Educación de la Junta de Castilla y León under the P.O. Castilla y León 2007-2013 to A.A.; Catalan Health Institute, Government of Catalonia (2009SGR290 to L.F.); the Spanish Association Against Cancer (to L.F.).

Acknowledgements

We are grateful to the breast/ovarian cancer patients who participated in this study. We are also grateful to Lara Hernández Sanz and Noemi Martínez Martín for their excellent technical support. This work was supported by grants PI10/2910 of the Instituto de Salud Carlos III, grants CSI004A10-2 (Consejería de Educación), BIO39/VA27/10 (Consejería de Sanidad) and the Cancer Prevention Program of the Regional Government of Castilla y León. Alberto Acedo is supported by the European Social Fund and Consejería de Educación de la Junta de Castilla y León under the P.O. Castilla y León 2007-2013. Lúcia Feliubadaló is supported by grants of the Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute, Government of Catalonia and the Spanish Association Against Cancer; contract grant numbers: ISCIIIRETIC: RD06/0020/1051, PI10/01422 and 2009SGR290.

Conflict of Interest Statement: None declared.

References

1. Claus,E.B.et al. (1996) The genetic attributable risk of breast and ovarian cancer. *Cancer*, **77**, 2318-2324.
2. Narod,S.A.et al. (2004) BRCA1 and BRCA2: 1994 and beyond. *Nature Reviews Cancer*, **4**, 665-676.
3. Stratton,M.R.et al. (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet*, **40**, 17-22.
4. The Breast Cancer Information Core Database.
<http://research.nhgri.nih.gov/projects/bic/Member/index.shtml> (June 2011, date last accessed).
Ref Type: Generic
5. The Universal Mutation Database-BRCA2. <http://www.umd.be/BRCA2/> (June 2012, date last accessed).
Ref Type: Generic
6. Ferla,R.et al. (2007) Founder mutations in BRCA1 and BRCA2 genes. *Ann.Oncol.*, **18 Suppl 6**, vi93-vi98.
7. Vega,A.et al. (2002) Analysis of BRCA1 and BRCA2 in breast and breast/ovarian cancer families shows population substructure in the Iberian peninsula. *Ann Hum Genet*, **66**, 29-36.
8. Beristain,E.et al. (2007) Differences in the frequency and distribution of BRCA1 and BRCA2 mutations in breast/ovarian cancer cases from the Basque country with respect to the Spanish population: implications for genetic counselling. *Breast Cancer Res Treat.*, **106**, 255-262.
9. Diez,O.et al. (2010) Heterogeneous prevalence of recurrent BRCA1 and BRCA2 mutations in Spain according to the geographical area: implications for genetic testing. *Fam.Cancer*, **9**, 187-191.
10. Diez,O.et al. (2003) Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: a high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat*, **22**, 301-312.
11. Salazar,R.et al. (2006) BRCA1-2 mutations in breast cancer: identification of nine new variants of BRCA1-2 genes in a population from central Western Spain. *Cancer Lett.*, **233**, 172-177.
12. Neuhausen,S.L.et al. (1998) Haplotype and phenotype analysis of nine recurrent BRCA2 mutations in 111 families: Results of an international study. *Am J Hum Genet*, **62**, 1381-1388.
13. Gao,Q.et al. (2000) Protein truncating BRCA1 and BRCA2 mutations in African women with pre-menopausal breast cancer. *Hum.Genet.*, **107**, 192-194.
14. van der Luijt,R.B.et al. (2001) De novo recurrent germline mutation of the BRCA2 gene in a patient with early onset breast cancer. *J.Med.Genet.*, **38**, 102-105.
15. Muniappan,B.P.et al. (2002) The DNA polymerase beta replication error spectrum in the adenomatous polyposis coli gene contains human colon tumor mutational hotspots. *Cancer Res.*, **62**, 3271-3275.
16. Kondrashov,A.S.et al. (2004) Context of deletions and insertions in human coding sequences. *Hum.Mutat.*, **23**, 177-185.
17. Wright,B.E.et al. (2002) Hypermutable bases in the p53 cancer gene are at vulnerable positions in DNA secondary structures. *Cancer Res.*, **62**, 5641-5644.
18. Velasco,E.et al. (2007) Heteroduplex analysis by capillary array electrophoresis for rapid mutation detection in large multiexon genes. *Nat Protoc.*, **2**, 237-246.
19. Esteban-Cardena, E.et al. (2004) High-throughput mutation detection method to scan BRCA1 and BRCA2 based on heteroduplex analysis by capillary array electrophoresis. *Clinical Chemistry*, **50**, 313-320.

20. The National DNA Bank. <http://www.bancoadn.org/en/home.htm>.
Ref Type: Generic
 21. Ensembl Genome Browser. <http://www.ensembl.org> (January 2010, date last accessed). (7 January 2004, date last accessed).
Ref Type: Generic
 22. Rannala,B.et al. (1998) Likelihood analysis of disequilibrium mapping, and related problems. *Am.J.Hum.Genet.*, **62**, 459-473.
 23. Instituto Nacional de Estadística. <http://www.ine.es/inebaseweb/71807.do?language=0> (June 2011, date last accessed).
Ref Type: Generic
 24. Human Genome Variation Society. GUIDELINES FOR MUTATION NOMENCLATURE.<http://www.hgvs.org/rec.html> (September 2011, date last accessed).
Ref Type: Generic
 25. Balmaña,J.et al. (2010) BRCA in breast cancer: ESMO Clinical Practice Guidelines. *Ann.Oncol.*, **21 Suppl 5**, v20-v22.
 26. Markham,N.R.et al. (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research*, **33**, W577-W581.
 27. Quikfold. The DINAMelt Web Server.<http://mfold.rna.albany.edu/?q=DINAMelt/Quickfold> (September 2012, date last accessed).
Ref Type: Generic
 28. Infante,M.et al. (2010) BRCA1 5272-1G>A and BRCA2 5374delTATG are founder mutations of high relevance for genetic counselling in breast/ovarian cancer families of Spanish origin. *Clin.Genet.*, **77**, 60-69.
 29. Campbell,C.D.et al. (2012) Estimating the human mutation rate using autozygosity in a founder population. *Nat.Genet.*, **44**, 1277-1281.
 30. Sagi,M.et al. (2011) Two BRCA1/2 founder mutations in Jews of Sephardic origin. *Fam.Cancer*, **10**, 59-63.
 31. Adams,S.M.et al. (2008) The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am.J Hum.Genet*, **83**, 725-736.
 32. Risch,N.et al. (1995) Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. *Nat Genet*, **9**, 152-159.
 33. Infante,M.et al. (2010) Two founder BRCA2 mutations predispose to breast cancer in young women. *Breast Cancer Res.Treat.*, **122**, 567-571.
 34. Pilato,B.et al. (2010) Mutations and polymorphic BRCA variants transmission in breast cancer familial members. *Breast Cancer Res.Treat.*
 35. Berman,D.B.et al. (1996) Two distinct origins of a common BRCA1 mutation in breast-ovarian cancer families: a genetic study of 15 185delAG-mutation kindreds. *Am J Hum Genet*, **58**, 1166-76.
 36. Laitman,Y.et al. (2012) Haplotype analysis of the 185delAG BRCA1 mutation in ethnically diverse populations. *Eur.J.Hum.Genet.*
 37. Palomba,G.et al. (2007) Origin and distribution of the BRCA2-8765delAG mutation in breast cancer. *BMC.Cancer*, **7**, 132.
-

TABLE AND FIGURES LEGENDS

Table 1. Clinical characteristics of the BRCA2- c.2808_2811delACAA families.

Table 2. Summary of DNA variants reported in mutation databases and detected in the random mutagenesis experiment between hairpins 4 and 10.

Fig. 1. Haplotype tree of eleven 13q polymorphic markers. The three main core haplotypes are in bold font and surrounded by squares. The markers used are ordered according to their position in the Ensembl Genome Browser. Asterisk indicates families where only the index case was available.

Fig. 2. Geographic distribution of the 51 BRCA2-c.2808_2811del families analyzed and their associated haplotypes. The three different haplotypes are represented by squares (WCL), circles (ECL) and triangles (SoS). CN-428 and CN-565 are families with only the index case whose core haplotype could not be ascertained.

Fig. 3. Quikfold prediction of the DNA secondary structure of the DNA fragment between nucleotides c.2714 and c.2902. Mutation c.2808_2811del is shown in hairpin #7. DNA variants detected in the mutagenesis experiment are indicated. Designation of DNA variants is based on the BRCA2 cDNA sequence NM_000059.1. Green squares highlight mutations that affect the same nucleotide of reported mutations but differ in the nucleotide change. Mutations that have been identified twice in two independent clones are shown in italics. Variants that have been detected in clones that contain more than one change are shown in bold type.

Fig. 4. DNA secondary structure predictions with Quikfold of sequences flanking highly prevalent mutations: A) BRCA1-c.68_69del; B) BRCA2-c.5946delT; C) BRCA2-c.8537_8538delAG. Affected nucleotides of each variant are circled.

Table 1. Clinical characteristics of the BRCA2- c.2808_2811delACAA families.

Family ^a	ADx ^{c,d,e}	Number of cancer cases in not tested relatives					
		BC<50 ^f	BC>50 ^f	OC	BOC	MBC	Other cancers (number) ^{c,d}
BU-48	53	3	-	-	-	-	-
BU-79	41, 41, 42, 44, 60^{OC}	-	-	-	-	-	<u>NHL</u>
BU-89	29	3	-	-	-	-	-
BU-260	32	-	-	-	-	-	Pancreas
BU-265	0 (43)	2(1)	-	1	-	-	-
BU-385	29, 35	2	-	-	-	-	Gastric, leukemia, lung
BU-389	52	1(1)	-	1	-	-	CRC, gastric (2), leukemia
BU-398	28	-	-	-	-	-	Gastric
BU-446	48, 62^{OC}	1	-	-	-	-	Larynx, renal
BU-456	44	1	1(1)	-	-	-	CRC (2)
BU-466	35-39^{bil}, 43	-	-	-	-	-	Esophagus, sarcoma
BU-477	47	-	1	-	-	-	Head and neck
BU-732	54^{bil}, 55	1	1	-	-	-	Head and neck, leukemia, lung
BU-852	58^{OC}, 39-52^{bil}	-	2	-	-	-	Esophagus, lung, gastric, CRC (3)
BU-855	36^{bil}	1	1	-	-	-	NHL, Endometrial, Prostate (2)
BU-904	45-64^{bil}	1	2 (1)	-	-	-	CNS, lung
SA-1	37	3	1	-	-	-	Pancreas (2), prostate (2)
SA-2	40	1	1(1)	-	-	-	Pancreas, prostate
SA-3	43	5	-	-	-	-	-
SA-4	30	-	-	-	1	-	CNS, gastric
SA-5	55	3	2	-	-	-	CRC, lung, gastric, pancreas (2)
SA-6	0 (32)	3	-	1	-	-	CNS, pancreas
SA-7	46, 70	5(2)	7	1	-	-	-
SA-8	48-52^{bil}-49^{OC}	1	4(1)	-	-	-	Pancreas
SA-9	34	1	-	1	-	-	Pancreas, prostate (2)
BI-167	76^{MBC}	-	1	-	-	-	Lung
M-110	46	2	2	-	-	-	-
M-171	65^{bil}	1	-	-	-	-	Gastric, Prostate
CN-59 ^e	32, 35, 40	2	-	-	-	-	<u>Cervix, Prostate</u>
CN-87	48-53^{OC}	-	-	-	-	-	Intestinal, <u>Pancreas</u> , Prostate
CN-93	41, 46	3	-	-	-	-	CRC, Sarcoma
CN-428	49^{OC}	3	-	1	-	-	Leukemia
CN-565	58^{MBC}	-	-	-	-	-	-
CN-827 M-178	31	1	-	-	-	-	Prostate
CN-877	69^{OC}	-	-	-	-	-	-
CN-1357	46, 49	-	-	-	-	1	Mediastinum, Prostate
ICO-61	26-30^{bil}	-	1	-	-	-	-
ICO-337	53^{OC}	-	-	-	-	-	CRC, Endometrial (2)
ICO-358	65	1	1	-	-	1	Leukemia, LNH, Renal
ICO-847	67^{OC}	-	2	-	-	-	Gastric (2), Head and Neck, Hepatic (2), Prostate
ICO-1188	34-50^{bil}	-	1	-	-	-	Bladder, Hepatic, Lung, Prostate
VH-60	42, 58	1	2	-	-	-	Prostate
VH-209	36	2	1	-	-	-	Peritoneal, Liver
VH-400	29, 36^{bil}, 53	2	-	-	-	-	<u>Stomach</u>
SP-121	38, 50-52^{OC}, 68^{OC}	-	-	-	-	-	-
SP-122	82	3	-	-	-	-	<u>Lung</u>
SP-378	45	1	2	-	-	-	Lung, Gastric
V-44 / SP-395	30	-	-	3	-	-	Leukemia, Lung
V-156	42	-	-	1	-	-	-
V-309	29, 38, 39	-	1	-	-	-	Prostate
Z-195	28, 43, 37, 72	1	-	-	-	-	<u>Prostate</u> (2), Bladder

^a Samples were recruited in the following collaborating centers: Hospital Universitario de Burgos (16 BU families), Centro de Investigación del Cáncer in Salamanca (nine SA families), Centro Nacional de Investigaciones Oncológicas in Madrid (nine CN families), Institut Català d'Oncologia in L'Hospitalet (five ICO families), Hospital de la Santa Creu i Sant Pau in Barcelona (three SP families), Hospital Clínico San Carlos in Madrid (three M families), Hospital Universitario La Fe in Valencia (three V families), University Hospital Vall d'Hebron in Barcelona (three VH families), Hospital de Cruces in Barakaldo (one BI family) and Hospital Clínico Universitario Lozano Blesa in Zaragoza (one Z family). ^b ADx: age of diagnosis. Actual

age is indicated between parentheses in asymptomatic carriers ^c Carriers or obligate carriers are underlined. ^d Type of cancer: BC, Breast Cancer; OC, Ovarian Cancer; BOC, Breast and Ovarian Cancer; MBC, Male Breast Cancer; bil, Bilateral Breast Cancer; NHL, Non-Hodking Lymphoma, CRC; Colorectal Cancer. ^e Patient CN-59 had cervix carcinoma at 38 years. ^f Number of bilateral BC cases are indicated between parentheses.

Table 2. Summary of DNA variants reported in mutation databases and detected in the random mutagenesis experiment between hairpins 4 and 10.

Hairpin-Loop	cDNA ^a	Databases	Mutagenesis	Nt / mutation ^b
4	c.2719-c.2727	1	4	1.8
Inter-sequence 4	c.2728-c.2738	3	3	1.8
5	c.2739-c.2769	7	11	1.7
Inter-sequence 5	c.2770-c.2776	1	2	2.3
6	c.2777-c.2792	3	3	2.7
7	c.2794-c.2825	13	5	1.8
8	c.2830-c.2840	7	3	1.1
9	c.2842-c.2871	5	6	2.7
Inter-sequence 9	c.2872-c.2883	3	3	2.2
10	c.2884-c.2895	0	3	4.0
TOTALS		43	42	2.1

^a Nucleotide numbering is based on *BRCA2* cDNA sequence NM_000059.1 where nucleotide 1 is the A of the ATG-translation initiation codon.

^b Nucleotides per mutation. Rates below the average (2.1) are shown in bold type.

Figure 1

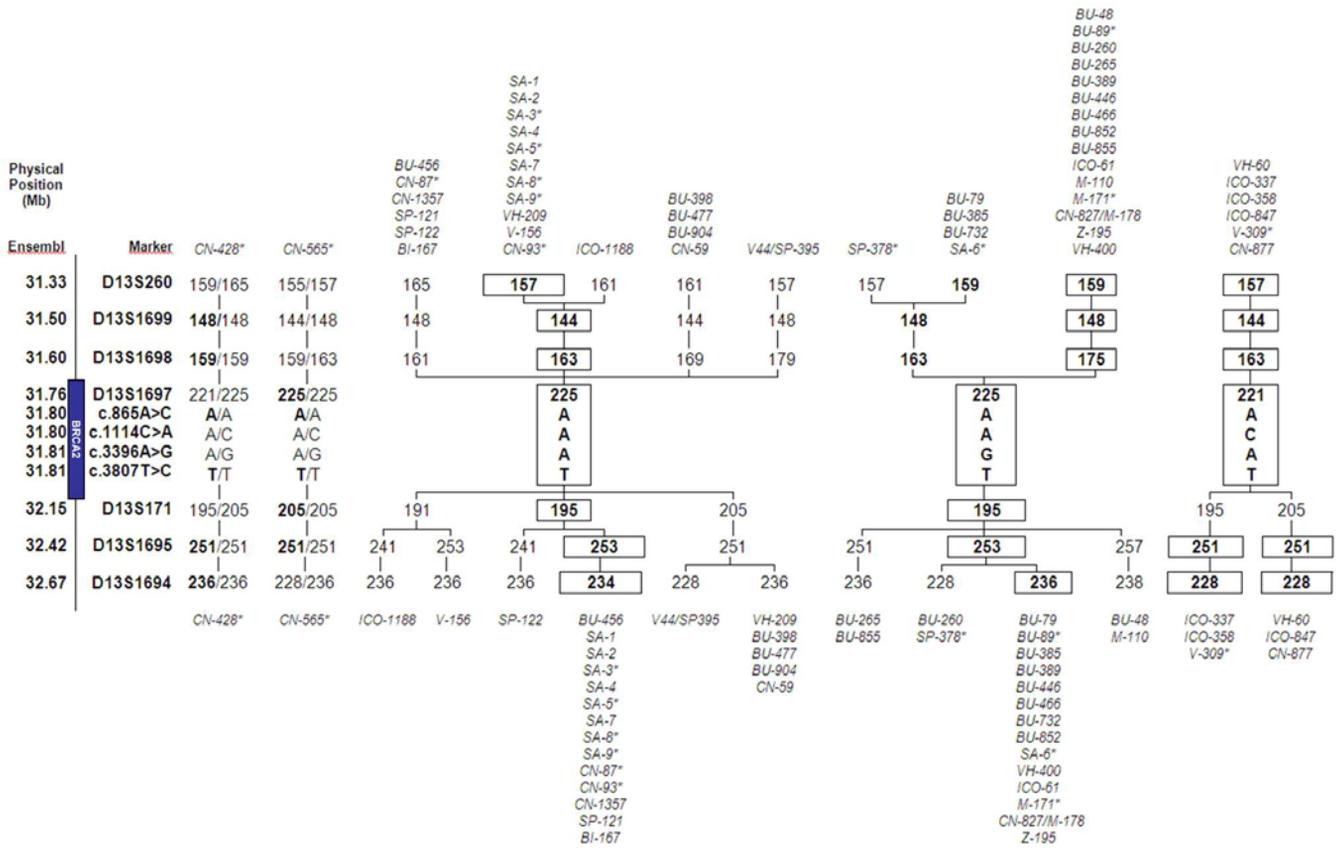
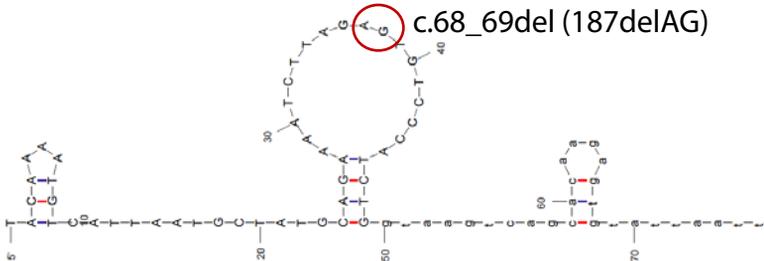
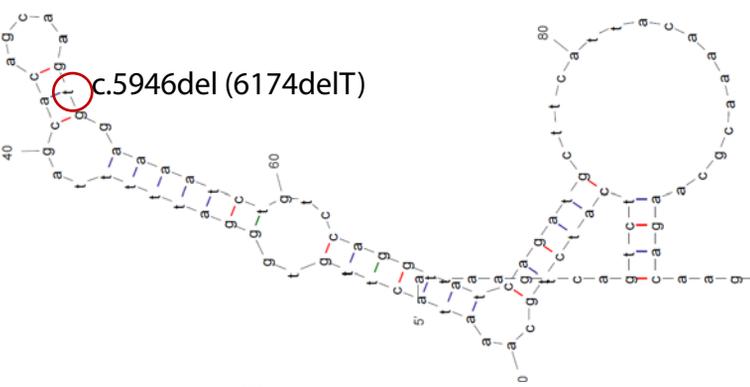


Figure 4

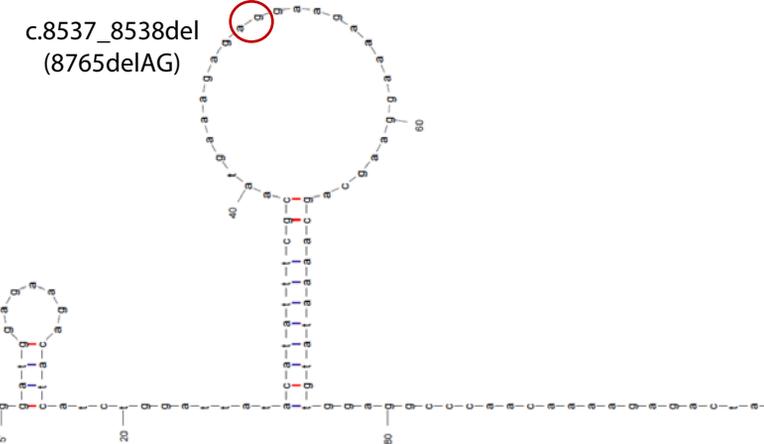
A)



B)



C)



Supplemental Table S1. Haplotype reconstruction of five and six markers with the PHASE program from the genotype data of 75 control individuals. The haplotypes linked to c.2808_2811del (WCL, ECL and MC) are highlighted.

Five markers¹ D13S1697- c.865A>C- c.1114C>A-c.3396A>G-c.3807T>C			Six markers¹ D13S1697-c.865A>C-c.1114C>A-c.3396A>G-c.3807T>C-D13S171		
Haplotype	Frequency	S.E	Haplotype	Frequency	S.E
2ACAT	0.2238	0.0115	2ACAT6	0.0975	0.0200
2AAAT-WCL	0.1929	0.0080	2ACAT7	0.0912	0.0222
2AAGT-ECL	0.1787	0.0141	2AAAC6	0.0863	0.0262
2AAAC	0.1560	0.0112	2AAAT-WCL	0.0799	0.0287
3AAGT	0.0933	0.0146	2AAGT6-ECL	0.0668	0.0102
3ACAT-SoS	0.0522	0.0111	2AAAT7	0.0584	0.0261
2CAAT	0.0186	0.0027	2AAAC7	0.0568	0.0187
1AAAT	0.0171	0.0050	3AAGT7	0.0419	0.0156
1AAGT	0.0109	0.0032	2AAGT7	0.0403	0.0137
3AAAT	0.0098	0.0079	3AAGT6	0.0384	0.0154
2ACAC	0.0020	0.0031	2AAAT2	0.0338	0.0054
1ACAT	0.0013	0.0026	2AAGT2	0.0290	0.0073
2AAGC	0.0013	0.0026	3AAGT2	0.0257	0.0049
3ACAC	0.0007	0.0020	2ACAT8	0.0243	0.0101
3AAAC	0.0349	0.0080	2AAGT8	0.0208	0.0072
1AAAC	0.0040	0.0051	3ACAT8-SoS	0.0157	0.0088
3CAGT	0.0013	0.0027	3ACAT6-SoS	0.0138	0.0067
3AAGC	0.0011	0.0025	2ACAT2	0.0136	0.0096
TOTAL: 18 HAPLOTYPES			1AAAT7	0.0124	0.0059
			2AAAC2	0.0122	0.0060
			2AAAT8	0.0120	0.0050
			3ACAT7	0.0097	0.0074
			2AAAC3	0.0097	0.0049
			2CAAT6	0.0090	0.0083
			3AAAC2	0.0083	0.0055
			3AAAT7	0.0078	0.0066
			2AAAC8	0.0077	0.0052
			2CAAT7	0.0075	0.0079
			1AAAT6	0.0069	0.0051
			1AAGT6	0.0067	0.0000
			3AAGT8	0.0066	0.0042
			2ACAT5	0.0053	0.0040
			3AAAC7	0.0051	0.0076
			3AAAC8	0.0050	0.0055
			3ACAT2	0.0040	0.0049
			3AAAC6	0.0040	0.0044
			1AAGT2	0.0040	0.0033

2AAAT5	0.0033	0.0033
1AAAC2	0.0027	0.0033
3ACAT5	0.0027	0.0033
2AAGT5	0.0020	0.0031
3AAGT3	0.0020	0.0031
2AAAT3	0.0017	0.0029
2CAAC7	0.0014	0.0028
3AAAT6	0.0014	0.0028
2CAGT6	0.0013	0.0029
3AAAT8	0.0007	0.0020
1ACAT6	0.0007	0.0020
2ACGT7	0.0007	0.0020
3CAGT6	0.0007	0.0020
2ACAC8	0.0006	0.0020
TOTAL: 51 HAPLOTYPES		

¹ D13S1697 Alleles: 1, 229 bp; 2, 225 bp; 3, 221 bp. D13S171 alleles: 2, 205 bp; 3, 203; 5, 197; 6, 195 bp; 7, 193 bp; 8, 191 bp.

Supplemental Table S2. Current ages of asymptomatic *BRCA2*-c.2808_2811delACAA carriers.

FAMILY	Number of carriers		3036delACAA+ (current age)									
			<40		41-50		51-60		61-70		>71	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
BU-48		1		38								
BU-79	2	2	18		44			52				90
BU-260										66		
BU-265	2		40		43							
BU-385	2		26				53					
BU-398	4	2	26, 27	25		43	55, 59					
BU-456	2		21, 22									
BU-466		2		25				60	63			
BU-477	3	1	35	25	46							
BU-904	1		37									
SA-1	1	1	21	19								
SA-2	3		24,30				54					
SA-4	1		40									
SA-6	1		32									
SA-7	1		33									
BI-167	1						56					
M-110	2 (BPM)*	2	33	30	44*	50						
CN-59	2	3	19	40		48, 50			65			
CN-87	2		33, 35									
CN-877	1	1			46	44						
CN-1357	1	1	27	30								
ICO-61		2		38						70		
ICO-337	1		33									
ICO-358	2			40		46						
ICO-847		1				45						
ICO-1188	1		26									
VH-60	3		34,37				58					
SP-121	2	1	35	39					62			
V-44/ SP-395	1				43							
V-156	2	1	26, 30	24								
Z-195	1	3	35	33				51 / 59				

* BPM, Bilateral prophylactic Mastectomy.

Supplemental Table S3. List of sequence variants detected in the random mutagenesis experiment and reported in patients in the mutation databases.

List of BIC and UMD variants ¹	HAIRPIN-LOOP	RANDOM MUTAGENESIS	
		DNA change ^{1,2}	Clone # ³
c.2635T>A		c.2633A>T	60
c.2636_2637delCT		c.2634C>A	35
c.2643A>G	2	c.2659G>A	32
c.2653_2656delGACA		c.2660A>T	6
c.2657A>T		c.2662A>G	7
c.2657delA		c.2668T>C	72
c.2660A>G		c.2670T>A	3
		c.2676C>A	13
c.2677C>T		c.2677delC	22
c.2678A>G	c.2680G>A	15	
c.2680G>A	c.2681T>A	55	
c.2684delC	c.2683G>A	31	
	c.2683G>A	35	
	c.2685T>C	25	
	c.2687A>G	44	
c.2690A>G	3	c.2691A>T	28
c.2691A>G		c.2699A>G	37
c.2698A>G		c.2699A>T	14
c.2701delC		c.2701C>A	36
		c.2702T>A	52
		c.2703T>A	38
		c.2703T>C	9
	c.2706T>C	55	
	c.2707T>C	15	
c.2716A>G			
c.2719A>G	4	c.2720A>T	72
		c.2720delA	22
		c.2722delG	43
		c.2725C>A	67
c.2731delG	5	c.2728C>T	23
c.2735C>T		c.2736A>C	24
c.2735C>A		c.2737G>T	27
c.2739C>T		c.2739C>G	27
c.2743_2747delACTTG		c.2744C>T	20
c.2745_2746delTT		c.2745T>A	54
c.2751A>G		c.2746T>A	21
c.2755G>A	c.2748T>C	60	
c.2758C>A	c.2756A>G	25	
c.2760delC	c.2762T>C	43	
	c.2763T>C	52	
	c.2768A>C	5	
	c.2768A>T	9	
	c.2769G>A	12	
c.2771A>T		c.2771A>G	1
		c.2771A>G	64
c.2780T>C	6	c.2777C>T	26
c.2786T>C		c.2783T>A	62
c.2786dup		c.2786delT	74
c.2798delC	7	c.2804A>G	5
c.2798_2799delCA		c.2808A>G	3
c.2798C>G		c.2812G>T	12
c.2803G>A		c.2816C>T	67
c.2803G>C		c.2819A>G	54
c.2805_2808delTAAA			
c.2808_2811delACAA			
c.2810_2811delAA			
c.2812_2815delGCAA			
c.2813C>A			

c.2817C>T c.2818C>T c.2821G>A			
c.2830A>T	8	c.2830A>G	12
c.2833_2834ins TT		c.2835delA	7
c.2834_2835delAA		c.2835A>G	52
c.2836_2837delGA c.2836delG c.2836G>C c.2837A>G			
c.2842dup c.2849T>A c.2854G>T	9	c.2848G>A	2
c.2860_2862delGAG		c.2851C>T	38
c.2862G>A		c.2857G>T	37
		c.2860G>T	22
		c.2867A>G	40
		c.2868A>T	50
c.2872A>G c.2881C>T c.2883G>A		c.2873G>T	44
		c.2875G>A	25
		c.2882A>C	71
	10	c.2886T>A	6
		c.2893delA	50
		c.2893A>C	56
c.2899_2900delCT c.2899C>G		c.2896A>G	42
		c.2898T>C	28
		c.2899C>A	34
		c.2900T>A	37
		c.2902G>A	63
c.2905C>T c.2918C>G c.2919G>A	11	c.2907A>G	52
		c.2914A>C	54
		c.2915A>C	67
c.2926_2927delinsAT c.2927_2929delCCT c.2930_2940del11 c.2944A>C		c.2924T>C	56
c.2944delA		c.2927C>T	10
c.2946A>G		c.2930T>A	41
c.2956A>G		c.2932A>G	46
c.2957_2958insG		c.2934T>A	23
c.2957dup		c.2935A>T	49
c.2957A>T		c.2938G>T	7
c.2957A>G		c.2944delA	72
		c.2946A>T	60
		c.2950G>T	63
		c.2957dup	67
		c.2957dup	28
c.2960A>T c.2965T>G	12		
c.2971A>G c.2977T>C c.2978G>A c.2979G>A		c.2970G>T	74
c.2987T>G	13		
		c.2999T>A	28

Nucleotide sequences of Forward (c.2610-c.2630) and Reverse primers (c.3008-c.3028) have been excluded from this table.

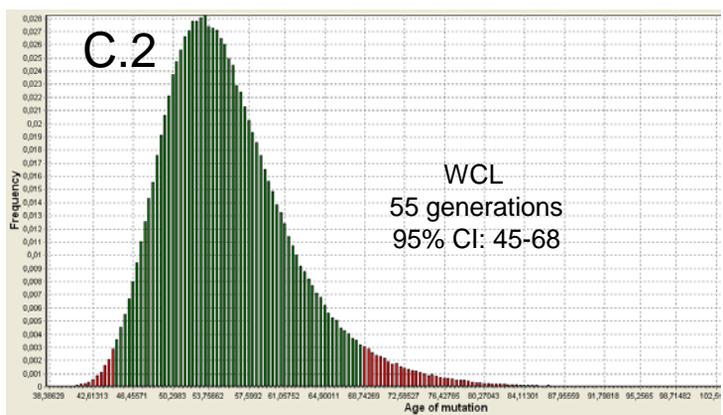
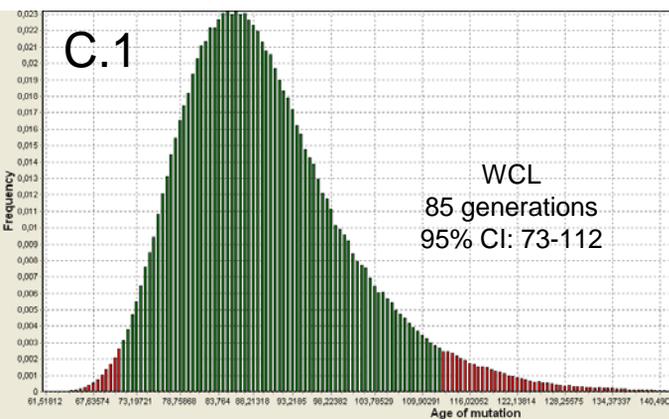
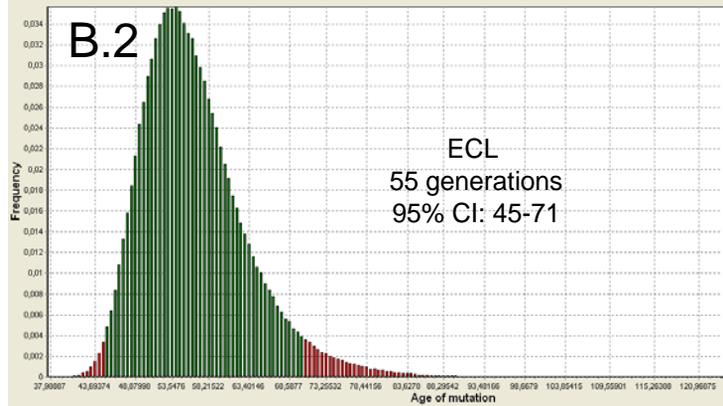
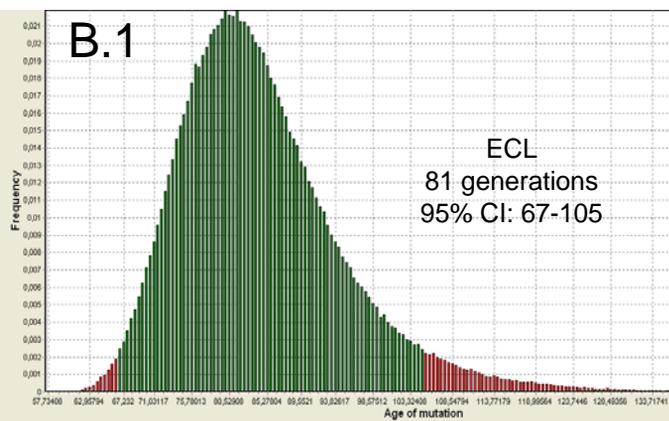
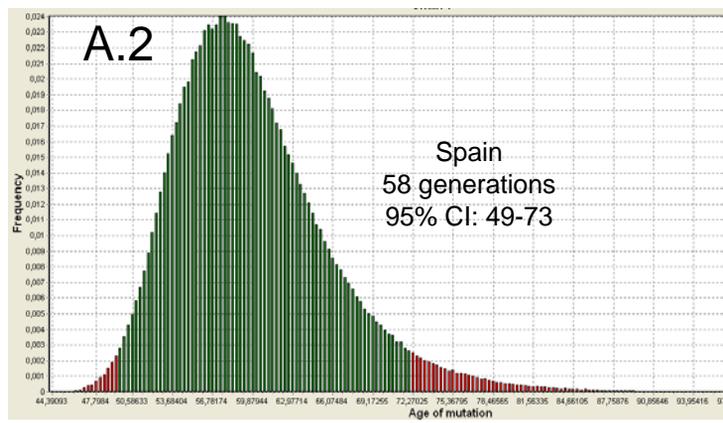
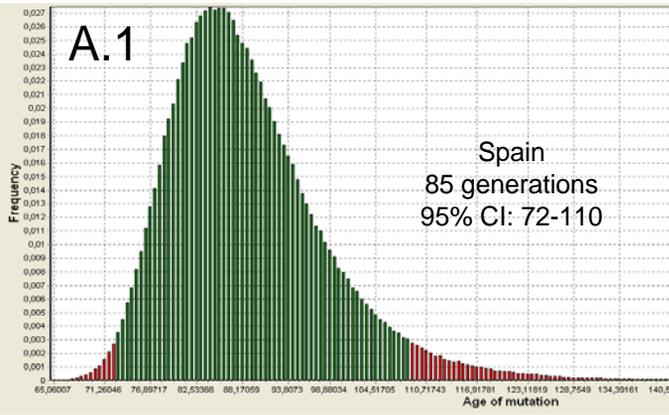
¹ Nucleotide numbering is based on *BRCA2* cDNA sequence NM_000059.1 where nucleotide 1 is the A of the ATG-translation initiation codon. Yellow squares denote mutations reported in the databases that have been replicated in the random mutagenesis experiment. Green squares indicate mutations that affect the same nucleotide in databases and mutagenesis experiment but with a different DNA change.

² Artificial mutations that have been identified in two independent clones are shown in bold and italics.

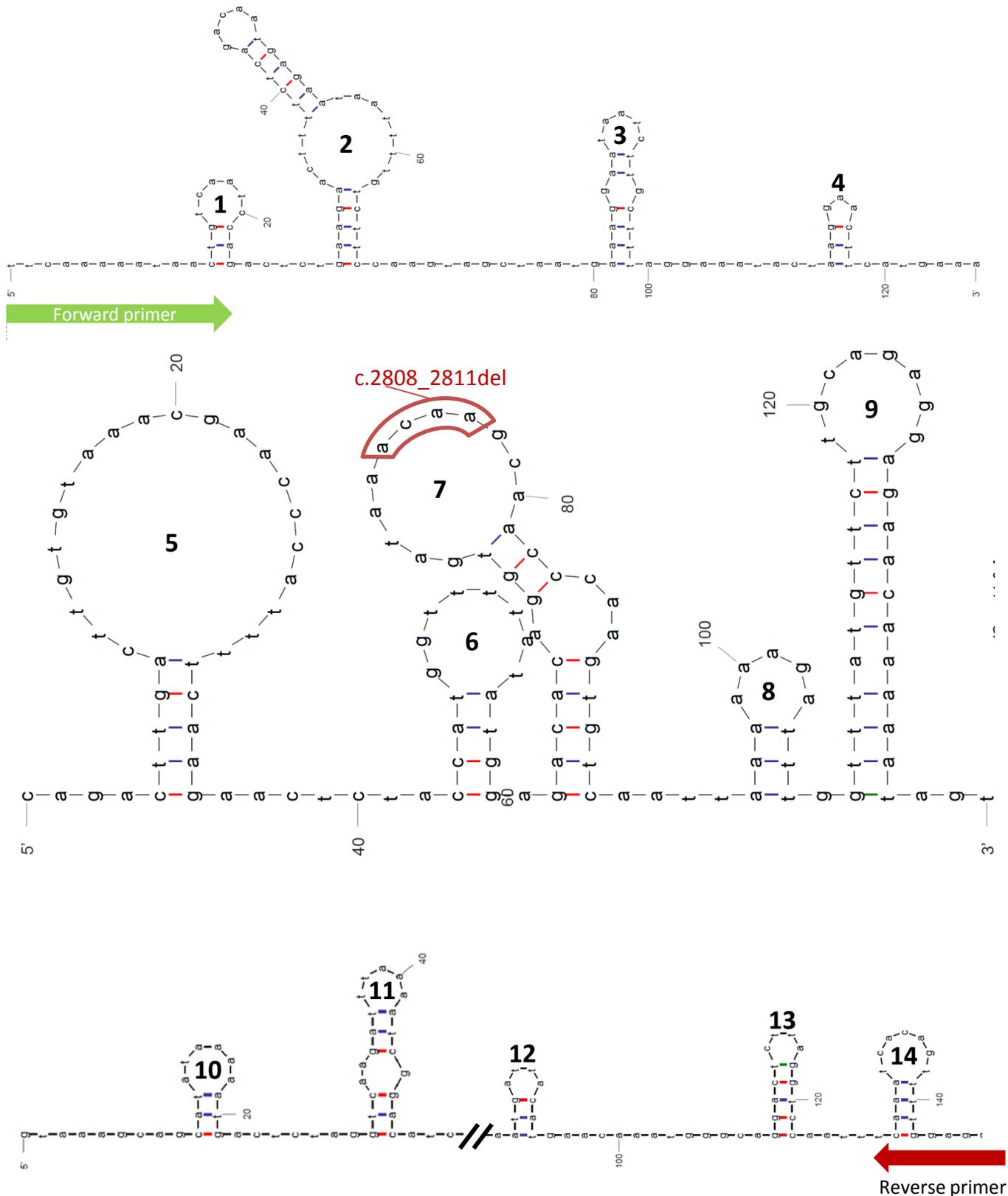
³ Clones that contain more than one mutation are shown in bold type

r=0.124

r=0.190



Supplemental figure S1. Estimation of 3036delACAA age in generations. DMLE+ 2.3 results in the Spanish context (A.1 and A.2), in the East CYL haplotype (B.1 and B.2) and in the West CYL haplotype (C.1 and C.2). A growth rate of 0.124 (A.1, B.1, C.1 figures) or 0.19 (A.2, B.2, C.2 figures) were considered for each analysis. The 95% confidence interval (CI) for each analysis is shown.



Supplemental Figure S2. Quikfold prediction of the DNA secondary structure of the DNA fragment between nucleotides c.2610 and c.3028 of the *BRCA2* cDNA. Primers 11E-Fw and 11F-Rev that were used for the random mutagenesis assay are indicated.