Gene Ther Mol Biol Vol 12, 175-180, 2008

A characterization of genetic haplotypes in *BRCA1* identifies linkage disequilibrium with a novel polymorphism in intron 7

Research Article

Josefa Salgado[1,*], Carmen Gil[1], Maitane Robles[1], Cristina Gutierrez[1], Carmen Reyna[2], Jesús García-Foncillas[1,2]

[1]Clinical Genetics Unit

[2]Department of Oncology, University Clinic of Navarra (CUN). Avda. Pio XII, 36. 31008, Pamplona. Navarra. Spain

---

***Correspondence**: Dr. Josefa Salgado, Clinical Genetic Unit, University Clinic of Navarra, Avda. Pio XII, 36, 31008 Pamplona, Navarra, Spain; Tel: + 34-948-255400 Ext: 1104; Fax: + 34-948-296795; e-mail: jsalgadog@unav.es

Key words: *BRCA1*-novel polymorphism, *BRCA1*-haplotypes, breast cancer, linkage disequilibrium

Abbreviations: Breast Cancer Information Core (BIC)

Received: 27 June 2008; Revised: 22 July 2008
Accepted: 23 July 2008; electronically published: September 2008

Summary

Inherited mutations in the *BRCA1* gene are known to confer a predisposition to breast and ovarian cancer. Mutations, and low-frequency variants, have been invariably detected on any of the two major haplotypes H1 and H2, albeit more frequently on the dominant haplotype H1. Deleterious mutations, detected in our patients, were studied in relation with different haplotypes. A group of 77 patients with hereditary and familiar breast and/or ovarian cancer have been studied. *BRCA1* gene analysis was done by direct sequencing. A seven polymorphic site cassette is used to define the *BRCA1*-haplotypes in our population. The frequency of a novel polymorphism found was studied in a control population of 100 unrelated healthy volunteers. Two new haplotypes (H2+H4 and H2+H5) not defined before have been found. We have first characterized a novel polymorphism (IVS7+16(TTC)nTTTTC) at intron 7 of *BRCA1* gene. The IVS7+16(TTC)nTTTTC polymorphism shows a significant linkage disequilibrium with a cassette of seven polymorphisms that define H2 haplotype in this population. Seven out of eleven patients with deleterious mutations show the polymorphic site cassette. The (TTC)7/7 and (TTC)6/7 genotypes in the intron 7 could be markers for H1 and H2 haplotypes respectively. Larger patient population studies would be needed to study the association between *BRCA1* mutations and the presence of the cassette.

I. Introduction

Mutations in the tumor suppressor breast/ovarian cancer susceptibility gene *BRCA1* have been identified in patients with high incidence of familial breast and ovarian cancer (Peto et al, 1999; Frank et al, 2002). According to BIC (Breast Cancer Information Core) database, more than 1500 distinct *BRCA1* mutations and sequence variants have been reported. Approximately 10% of these variants reside within introns, or close to exon junctions, and may impact RNA stability and/or splicing. On the other hand, 10 different *BRCA1*-haplotypes were described (Shattuck-Eidens et al, 1997). Since then, other haplotypes have been described according to the population studied and the SNPs used (Freedman et al, 2005). Mutations, and low-frequency variants, have been detected on any of the two major haplotypes H1 and H2, albeit more frequently on the dominant haplotype H1 (Frost et al, 2005). During *BRCA1*-haplotypes characterization we have found two new haplotypes (H2+H4 and H2+H5), and a novel polymorphism in *BRCA1*-intron 7 that shows significant linkage disequilibrium with a seven polymorphic site cassette. Deleterious mutations, detected in our patients, have been studied in relation to the haplotypes and the novel polymorphism.

II. Materials and Methods

A. Patients

77 cases that fulfilled familiar hereditary breast cancer criteria (Eccles et al, 2000) were selected. In addition, 100 healthy blood donors were also studied. All participants gave their informed consent prior to blood sample extraction.

B. DNA isolation and sequencing

Genomic DNA was automatically extracted (MagNaPure, Roche) from peripheral blood. Direct sequencing of the *BRCA1* gene was done on an automated sequencer ABI PRISM® 3130XL (Applied Biosystems). Genetic variants were detected by comparison with the consensus wild-type sequence (Genbank Nº U14680), and were confirmed by repeated analysis, including reverse-primer sequencing. *BRCA1* gene variants were used to define the *BRCA1* haplotype, as described by Frost and colleagues in 2005.

C. Statistical Analysis

The χ^2 or Fischer's two-tailed exact test was used to compare the observed genotype distributions with those expected by the Hardy-Weinberg equilibrium. Linkage disequilibrium between *BRCA1*-intron 7 polymorphism and *BRCA1*-haplotypes was also evaluated.

III. Results

BRCA1 haplotypes detected fell into two major (H1 and H2), two minor (H4 and H6) and three low-frequency (H5, H2+H4 and H2+H5) groups (**Table 1**). The most common haplotype (H1), which corresponds to the consensus sequence (Miki et al, 1994; Smith et al, 1996) was found in the 36% of the alleles. The second most common haplotype H2 (28%) was characterized by a seven polymorphic site cassette (S694S, L771L, P871L, E1038G, K1183R, S1436S, S1613G). The H4 (10%) and H5 (3%) haplotypes contained respectively the sequence variants Q356R and S1040N, when compared to H1 haplotype. The H6 haplotype (17%) was identical to H2 except for the one additional variant D693N. Finally, two first described haplotypes were present: H2+H4 (4%) and H2+H5 (2%).

A novel triplet deletion (TTC) in the intron 7 of the *BRCA1* was identified in our patients, in a $(TTC)_7TTTTC$ region. Resulting genotypes were (TTC)7/7, (TTC)6/7 and (TTC)6/6 (**Figure 1**). The frequencies were 47% (TTC)7/7, 41% (TTC)6/7 and 12% (TTC)6/6 These variant alleles were not described before and, therefore, there were no information about its distribution in normal population. We studied the segregation of these alleles in the 77 patients and in 100 healthy volunteers, in order to compare if there were a particular (TTC)n genotype associated to disease. We could not find any association. Results showed frequencies of 44% (TTC)7/7, 46% (TTC)6/7 and 10% (TTC)6/6 in the 100 healthy donors. The statistical analysis indicated that the distribution was in agreement with that predicted by the Hardy-Weinberg equilibrium, with no statistical differences between patients and healthy donors (p=0.8). We studied the distribution of the IVS7+16(TTC)nTTTTC genotype among the different *BRCA1* haplotypes (**Table 2**) in the patients. Given the fact that some haplotypes were present in a very low frequency we defined the "H1-like" haplotype (H1, H4 and H5, all of them lacking the cassette) and the "H2-like" haplotype (H2, H6, H2+H4 and H2+H5, all of them carrying the cassette). Results indicated that (TTC)7/7 and (TTC)6/7 polymorphisms were preferentially, and respectively, associated to "H1-like" and "H2-like haplotypes; while (TTC)6/6 polymorphism was present in the two groups. We could establish that the (TTC)7/7 genotype segregated with the absence of the cassette, and the (TTC)6/7 genotype was associated with the presence of the cassette, with a statistically significance of p<0.001. Moreover, 11 out of 77 patients studied had deleterious variants (**Table 3**), and 7 out of these 11 patients carried the seven polymorphic site cassette (**Table 2**).

IV. Discussion

The seven polymorphic site cassette identified in this study, S694S, L771L, P871L, E1038G, K1183R, S1436S and S1613G, have been reported previously to have no significant differences in allele frequencies between familial breast/ovarian cancer patients and the general population (Peto et al, 1999). Although these polymorphisms are actually part of the same haplotype, they are usually reported individually. In so doing, the reported degree of variation in the *BRCA1* sequence may appear to be more substantial than it is. Shattuck-Eidens and coworkers described in 1997 a total of 10 haplotypes among 1590 alleles from a US population. Their reported haplotypes included those identified in our study (H1, H2, H4, H5 and H6), and up to ten low-frequency haplotypes absent among our 77 patients. It has been suggested that haplotypes H1 and H2 represent common chromosomes on which mutations and other variants occur.

Table 1. BRCA1-haplotypes found.

		HAPLOTYPES						
Sequence Variant	**Exon**	**H1**	**H2**	**H4**	**H5**	**H6**	**H2+H4**	**H2+H5**
c.1186A>G(p.Q356R)	11	-	-	+	-	-	+	-
c.2196G>A(p.D693N)	11	-	-	-	-	+	-	-
c.2201C>T(p.S694S)	11	-	+	-	-	+	+	+
c.2430T>C(p.L771L)	11	-	+	-	-	+	+	+
c.2731C>T(p.P871L)	11	-	+	-	-	+	+	+
c.3232A>G(p.E1038G)	11	-	+	-	-	+	+	+
c.3238G>A(p.S1040N)	11	-	-	-	+	-	-	+
c.3667A>G(p.K1183R)	11	-	+	-	-	+	+	+
c.4427T>C(p.S1436S)	13	-	+	-	-	+	+	+
c.4956A>G(p.S1613G)	16	-	+	-	-	+	+	+
Haplotypes frequency		36%	28%	10%	3%	17%	4%	2%

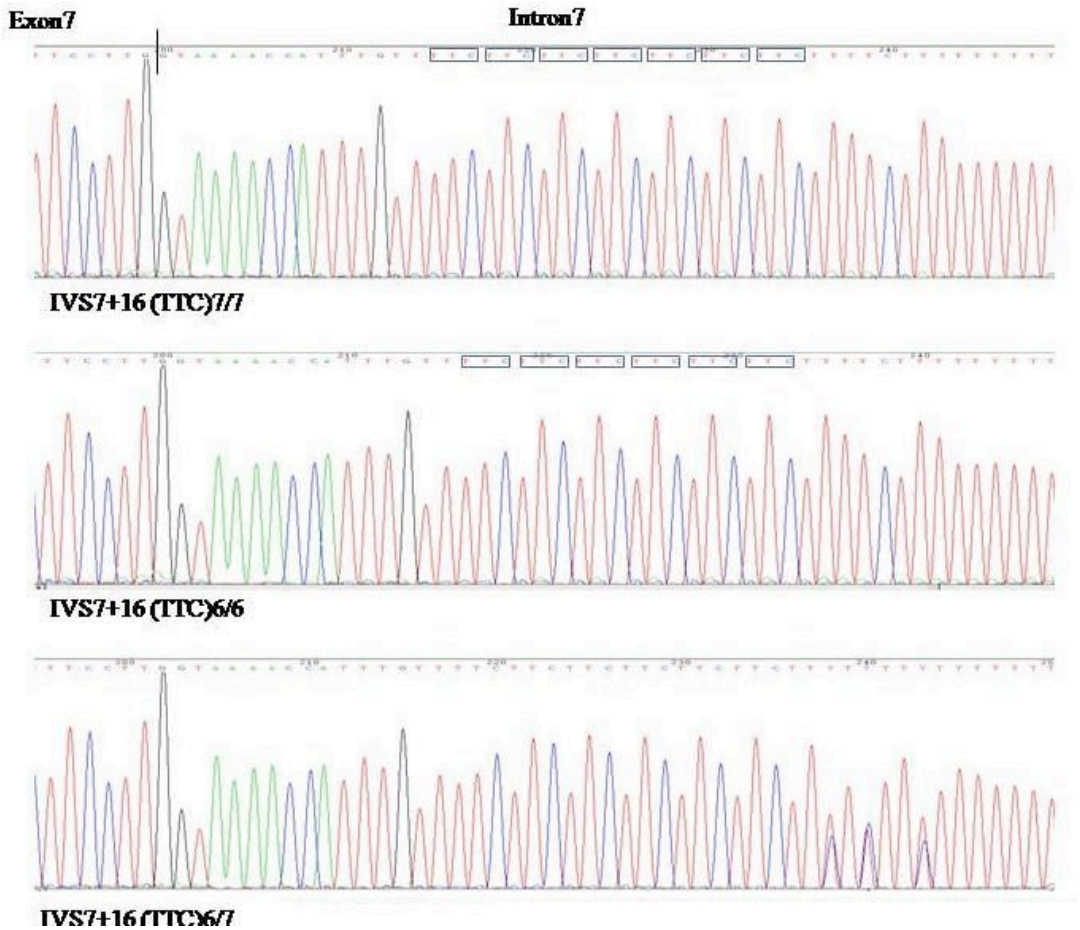

Figure 1. Direct sequencing showing the IVS 7 + 16(TTC)nTTTTC region where a TTC deletion has been found. The resulting variant alleles (TTC)7/7, (TTC)6/7 and (TTC)6/6 are shown in the figure.

Table 2. Distribution of the IVS7+16(TTC)nTTTTC genotype among the different BRCA1 haplotypes found.

	HAPLOTYPES							
Intron 7 genotypes	**Intron**	**H1**	**H2**	**H4**	**H5**	**H6**	**H2+H4**	**H2+H5**
IVS7+16 (TTC) 7/7	7	26*		8	2			
IVS7+16 (TTC) 6/7	7		18*			10*	3	1
IVS7+16 (TTC) 6/6	7	2*	4			3		

* Deleterious mutations have been found in patients of these groups.

Table 3. Deleterious mutations found in 11 patients.

Patient	Haplotype	IVS7 variant	Mutation type
1	H1	IVS7+16 (TTC) 6/6	Arg71Gly
2	H1	IVS7+16 (TTC) 6/6	Arg71Gly
3	H1	IVS7+16 (TTC) 7/7	Ala1708Glu
4	H1	IVS7+16 (TTC) 7/7	Ala1708Glu
5	H2	IVS7+16 (TTC) 6/7	4314delAC
6	H2	IVS7+16 (TTC) 6/7	4156-7delAA
7	H2	IVS7+16 (TTC) 6/7	Tyr261Stop codon
8	H6	IVS7+16 (TTC) 6/7	Arg71Gly
9	H6	IVS7+16 (TTC) 6/7	Arg71Gly
10	H6	IVS7+16 (TTC) 6/7	Arg71Gly
11	H6	IVS7+16 (TTC) 6/7	1191delC

Haplotype H3 would be related to haplotypes H1 and H2 by a recombination event between exons 11 and 13. Haplotypes H4 through H10 occur at low frequency and would require more recombination events to be related to the two common haplotypes (Shattuck-Eidens et al, 1997). Similarly, in our study, haplotypes H1 and H2 appear to be made up of ancient alleles that have independently acquired additional variants. In this way, H4 and H5 haplotypes on one hand, and H6 haplotype on the other, are identical to H1 and H2 respectively except for the one additional variant acquired on each case. Frequencies of our haplotypes vary compared to bibliographic data (Shattuck-Eidens et al, 1997; Frost et al, 2005), markedly for the minor haplotypes H4 and H6, suggesting a variety of ancestral populations. We have found two new low-frequency haplotypes (H2+H4 and H2+H5) that resemble H2 haplotype but acquiring the variants that define H4 and H5 respectively. All these less frequent haplotypes could be expected to originate from different ancestral populations or they could have been transferred from one haplotype to the other by gene conversion. We have detected deleterious mutations in 11 of our patients. Considering the seven polymorphic site cassette, H2 and H6 haplotypes carry this cassette while it is absent in H1. Seven out of these 11 patients with deleterious mutations (64%) belong to the H2 or H6 haplotypes. Although larger patient population studies would be needed, there seems to exist a weak association between *BRCA1* mutations and the presence of the cassette.

To our knowledge, no triplet deletion in the intron 7 of the *BRCA1* has yet been published. Analysis of intron 7 among breast cancer patients revealed size variations in the IVS7+16 $(TTC)_7$TTTTC region due to a triplet deletion, being the variant alleles (TTC)7/7, (TTC)6/6 and (TTC)6/7. We have study the segregation of these alleles in the 77 patients and in 100 healthy volunteers, in order to compare if there is a particular (TTC)n genotype associated to disease. We could not find any association. Moreover, the frequency of this intron 7 variant among the healthy donors suggests that it would be a polymorphic site. TTC repeats are one of the most ubiquitous short tandem repeats, among 10 possible trinucleotide sequences, in the human genome (Subramanian et al, 2003a,b) with several important biological functions *in vivo*. In *Mycoplasma* GAA-TCC repeats are responsible for the regulation of gene expression (Glew et al, 1998; Liu et al, 2000). Large expansions of the GAA-TTC sequence have been found in human genes like *FXN*, in association with the autosomal recessive Friedreich´s ataxia (Pandolfo, 2000). A TTC deletion in intron 5 of the gene *CYP19* encoding the P450 aromatase protein has been described, although without clear evidence of its relationship with breast cancer (Probst-Hensch et al, 1999). Biochemical studies suggest that GAA-TCC repeats *per se* are preferred sites for possible intramolecular recombinations. Although the frequency of recombination between any two homologous sequences, in the vast majority of cases, increases upon lengthening the recombining DNA fragment, GAA-TCC sequences demonstrated an exactly reverse relationship (Napierala et al, 2004). The role of the newly identified TTC deletion would deserves further investigation in this context. We have studied the relevance of these size variations within the different haplotypes detected in our patients. From table 2 we can say that (TTC)6/6 genotype is detected on any of the two major haplotypes, H1 or H2, while (TTC)7/7 and (TTC)6/7 are in linkage disequilibrium invariably associated to the H1 haplotype (or its variants H4 and H5) and H2 haplotype (or its variants H6, H2+H4 and H2+H5), respectively. From the point of view of the seven polymorphic site cassette, (TTC)6/7 segregates with this cassette, opposite to (TTC)7/7 genotype that never carries this cassette. The (TTC)7/7 and (TTC)6/7 genotypes in the intron 7 could be markers for H1 and H2 haplotypes respectively.

Acknowledgments

We are grateful to the patients are their families for their participation in our research project. We thank other members of the Clinical Genetics Unit for their support and encouragement.

References

Eccles D, Evans D, Mackay J (**2000**) Guidelines for a genetic risk based approach to advising women with a family history of breast cancer. UK Cancer Family Study Group (UKCFSG), **J Med Genet** 37, 203-209.

Frank T.S, Deffenbaugh A.M, Reid J.E, Hulick M, Ward B.E, Lingengelter B, Gumpper K.L, Scholl T, Tavtigian S.V, Pruss D.R,Critchfield G.C (**2002**) Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: Analysis of 10,000 individuals. **J Clin Oncol** 20, 1490-2002.

Freedman M, Penney K, Stram D, Riley S, McKean-Cowdin R, Le Marchand L, Altshuler D, Haiman C (**2005**) A haplotype-based case-control study of *BRCA1* and sporadic breast cancer risk, **Cancer Res** 65, 7516-7522.

Frost P, Jugessur A, Apold J, Heimdal K, Aloysius T, Eliassen AK, Fauske L, Matre G, Eiken HG (**2005**) Complete mutation screening and haplotype characterization of the *BRCA1* gene in 61 familial breast cancer patients from Norway, **Dis Markers** 21, 29-36.

Glew MD, Baseggio N, Markham PF, Browning GF, Walker ID (**1998**) Expression of the pMGA genes of Mycoplasma gallisepticum is controlled by variation in the GAA trinucleotide repeat lengths within the 5' noncoding regions, **Infect Immun** 66, 5833-5841.

Liu L, Dybvig K, Panangala VS, van Santen VL, French CT (**2000**) GAA trinucleotide repeat region regulates M9/pMGA gene expression in Mycoplasma gallisepticum, **Infect Immun** 68, 871-876.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al (**1994**) A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*, **Science** 266, 66-71.

Napierala M, Dere R, Vetcher A, Wells RD (**2004**) Structure-dependent recombination hot spot activity of GAA.TTC sequences from intron 1 of the Friedreich's ataxia gene, **J Biol Chem** 279, 6444-6454.

Pandolfo M (**2000**) The molecular basis of Friedreich ataxia, **Neurologia** 15, 325-9.

Peto J, Collins N, Barfoot S, Seal S, Warren W, Rahman N, Easton D, Evans C, Deacon J, Stratton M (**1999**) Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. **J Natl Cancer Inst** 91, 943-949.

Probst-Hensch NM, Ingles SA, Diep AT, Haile RW, Stanczyk FZ, Kolonel LN, Henderson BE (**1999**) Aromatase and breast cancer susceptibility, **Endocr Relat Cancer** 6, 165-173.

Shattuck-Eidens D, Oliphant A, McClure M, McBride C, Gupte J, Rubano T, Pruss D, Tavtigian SV, Teng DH, Adey N, Staebell M, Gumpper K, Lundstrom R, Hulick M, Kelly M, Holmen J, Lingenfelter B, Manley S, Fujimura F, Luce M, Ward B, Cannon-Albright L, Steele L, Offit K, Thomas A, et al (**1997**) *BRCA1* sequence analysis in women at high risk for susceptibility mutations. Risk factor analysis and implications for genetic testing. **JAM**. 278, 1242-1250.

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 Res** 6, 1029-1049.

Subramanian S, Madgula VM, George R, Mishra RK, Pandit MW, Kumar CS, Singh L (**2003a**) Triplet repeats in human genome: distribution and their association with genes and other genomic regions, **Bioinformatics** 19, 549-552.

Subramanian S, Mishra RK, Singh L (**2003b**) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions, **Genome Biol** 4, R13.