J Hum Genet (2008) 53:490–498
DOI 10.1007/s10038-008-0276-0

ORIGINAL ARTICLE

Analysis of *GADD45A* sequence variations in French Canadian families with high risk of breast cancer

Sylvie Desjardins · Geneviève Ouellette · Yvan Labrie · Jacques Simard · INHERIT BRCAs · Francine Durocher

Received: 21 December 2007 / Accepted: 21 February 2008 / Published online: 19 March 2008

Abstract *GADD45A* is an evolutionary conserved gene whose expression is regulated by two major tumor suppressor proteins involved in breast cancer etiology, namely, p53 and BRCA1, and which acts primarily in the control of the G2/M cell-cycle transition, apoptosis, and DNA repair. Following genotoxic stress, the p53 protein activates *GADD45A* transcription, whereas in absence of DNA damage, BRCA1 represses *GADD45A* expression through interaction with the zinc finger protein ZNF350. Moreover, BRCA1 can activate *GADD45A* gene expression through interactions with transcription factors binding to the gene promoter. On the basis of the intricate network of interactions between GADD45A, p53, and BRCA1, and the fact that both *BRCA1* or *TP53* mutations are involved in breast cancer tumorigenesis, we undertook the characterization of the entire coding sequence, intron/exon boundaries, and p53- and ZNF350-binding sequences of this potential breast cancer susceptibility candidate gene in a sample set of 96 women affected with breast cancer from non-*BRCA1* and *BRCA2* French Canadian families with a high risk of breast cancer and 95 healthy controls from the same population. Although none of the 12 identified sequence variations show a significant difference in frequency between both sample sets, haplotype phasing and frequency estimations identified a common haplotype displaying a higher frequency among the control group. As the variants present on this particular haplotype are non-coding variants in either intron 2 or 3, this finding will have to be further investigated in larger cohorts and other populations. In this regard, our study also identified tagging single nucleotide polymorphisms (tSNPs), providing useful data for other large-scale association studies.

Keywords Breast cancer · *GADD45A* gene · Sequence analysis · Haplotype · Tagging SNP

Introduction

Breast cancer is a complex disease involving genetic and lifestyle components. Our current understanding of breast cancer susceptibility is based upon a polygenic model in which several alleles conferring variable risk act in concert. Under this model, it is hypothesized that 88% of all breast cancer cases may be found in 50% of the population more at risk (Pharoah et al. 2002). Thus, finding the underlying genetic components conferring such risk may have a great impact on breast cancer prevention and treatment. However, our knowledge of breast cancer susceptibility remains incomplete. Mutations and/or allelic loss of tumor suppressor genes, such as *BRCA1*, *BRCA2*, *TP53*, as well as other genes, are associated with an increased predisposition to breast cancer tumorigenesis (Antoniou and Easton 2006). However, mutations in the two major predisposition genes *BRCA1* and *BRCA2* account for less than 25% of families with hereditary breast cancer (Antoniou and Easton 2006), indicating that many of the underlying genetic

Other members of INHERIT BRCAs involved in clinical aspects of the study are listed in the Appendix.

Electronic supplementary material The online version of this article (doi:10.1007/s10038-008-0276-0) contains supplementary material, which is available to authorized users.

S. Desjardins · G. Ouellette · Y. Labrie · J. Simard · F. Durocher (✉)
Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University, 2705 Laurier Boulevard, T2-53, G1V 4G2 Québec City, QC, Canada
e-mail: francine.durocher@crchul.ulaval.ca

factors involved in breast cancer susceptibility remain to be uncovered.

GADD45A (*G*rowth *A*rrest and *D*NA *D*amage-induced 45, *A*lpha) is a p53- and BRCA1-regulated gene acting in the control of the G2/M cell-cycle transition, apoptosis, and DNA repair (Wang et al. 1999; Harkin et al. 1999; Smith et al. 1996). Following genotoxic stress, the p53 protein binds directly to a consensus sequence located in *GADD45A* third intron and indirectly, through its interaction with transcription factors, to the promoter region, therefore activating *GADD45A* transcription (Zhan et al. 1998; Jin et al. 2001). In absence of DNA damage, BRCA1 represses *GADD45A* expression through interaction with the zinc finger protein ZNF350 (ZBRK1), this complex recognizing a DNA binding site also located in the *GADD45A* third intron (Zheng et al. 2000). Moreover, BRCA1 can activate *GADD45A* gene expression through interactions with transcription factors binding to the gene promoter (Fan et al. 2002).

Given that *GADD45A* is a target of two major tumor suppressor proteins involved in breast cancer, namely, BRCA1 and p53, and that specific haplotypes of the gene encoding BRCA1 corepressor ZNF350 have also been recently associated with a modulation of breast cancer risk in our cohort of non-*BRCA1/2* high-risk breast cancer families (Desjardins et al. 2008), *GADD45A*, therefore, represents an attractive candidate gene with regard to breast cancer susceptibility. Although no alterations of *GADD45A* have been found in previous studies performed on a limited set of breast tumors, human cancer cell lines of various origins, and familial breast carcinomas (Blaszyk et al. 1996; Campomenosi and Hall 2000; Sensi et al. 2004), *Gadd45a*−/− cells and mice display genomic instability (Hollander et al. 1999, 2005). The absence of *Gadd45a* has been associated with an acceleration of Ras-driven mammary tumor formation in mice (Tront et al. 2006).

Therefore, we investigated the plausible implication of *GADD45A* in breast cancer susceptibility by analyzing 96 individuals affected with breast cancer from our cohort of high-risk non-*BRCA1/2* breast and ovarian cancer families, along with 95 healthy controls from the same origin. The entire *GADD45A* coding region, intron–exon boundaries, and the p53- and BRCA1/ZNF350-binding consensus sequences were examined.

Materials and methods

Ascertainment of families and DNA extraction

All 96 non-*BRCA1/2* individuals from French Canadian families with a high risk of breast and ovarian cancer, along with 95 healthy controls included in this study, provided written informed consent. The research project has also been reviewed by the ethics committee of each participating institution. Details regarding selection criteria for breast cancer cases, experimental and clinical procedures, and the INHERIT BRCAs research program have been described previously (Simard et al. 2007; Durocher et al. 2006; Desjardins et al. 2008). Control blood DNA was either used directly or subjected to whole-genome amplification using Illustra GenomiPhi V2 DNA amplification kit (GE Healthcare, formerly Amersham Biosciences, Piscataway, NJ, USA) according to the manufacturer's instructions.

Polymerase chain reaction amplification, mutation analysis, and variant characterization

The *GADD45A* gene (NM_001924.2) consists of four exons covering more than 3 kb of genomic DNA. Direct sequencing of genomic sequence was performed using sequencing primers listed in Supplemental Table 1. Direct sequencing was performed using an ABI3730XL automated sequencer (Applied Biosystems, Foster City, CA, USA), according to manufacturer's instructions. Staden preGap4 and Gap4 programs were used for sequence data analysis. Deviation from Hardy–Weinberg equilibrium (HWE) and allelic differences between both series were measured using a X^2 test with 1 degree of freedom. The effect of a given variant on splice-site consensus strength was evaluated with the Splice Site Prediction Program using Neural Networks (SSPNN) (Reese et al. 1997) with default parameters. The putative impact of the exonic variant on exonic splicing enhancers was also assessed using ESE Finder (Cartegni et al. 2003; Smith et al. 2006).

Linkage disequilibrium analysis, haplotype estimation, and tagging single nucleotide polymorphism selection

The Linkage Disequilibrium Analysis (LDA) program (Ding et al. 2003) was used to calculate pairwise LD defined by Lewontin's $|D'|$ and r^2 measures on cases and controls combined (Lewontin 1964; Devlin and Risch 1995). The PHASE 2.1.1 software (Stephens et al. 2001) was used with default parameters to estimate the total set of haplotypes present among the case and control data sets and to evaluate the global test of significance associated with a case-control comparison. Evaluation of a positive association of a specific haplotype displaying a frequency >5% with breast cancer was further analyzed using the WHAP program (Purcell et al. 2007), implementing a regression-based association test. The determination of haplotype blocks and tagging single nucleotide polymorphisms (tSNPs) for each LD block was performed with

Table 1 *GADD45A* sequence variations and genotype distribution in familial breast cancer cases and controls

SNP	SNP ID[a]	dbSNP ID	Location	Series	No. Individuals	No. observed (expected)[b]			MAF	HWE p value[c]	χ^2 p value[d]	MAF in NCBI[e]
						Common homozygote	Heterozygote	Rare homozygote				
1	c.45-23C/T	rs3783466	Intron 1	Cases	96	71 (71.8)	24 (22.5)	1 (1.7)	0.135	0.507	0.889	0.148
				Controls	92	71 (69.5)	18 (20.9)	3 (1.6)	0.130	0.187		
2	c.147-103G/A	NA	Intron 2	Cases	84	83 (83.0)	1 (1.0)	0 (0.0)	0.006	0.956	0.967	–
				Controls	89	88 (88.0)	1 (1.0)	0 (0.0)	0.006	0.957		
3	c.147-53G/T	rs2759219	Intron 2	Cases	95	76 (74.3)	16 (19.4)	3 (1.3)	0.116	0.083	0.112	0.206
				Controls	89	59 (60.7)	29 (25.6)	1 (2.7)	0.174	0.210		
4	c.384 + 93A/G	rs3783468	Intron 3	Cases	96	32 (33.2)	49 (46.5)	15 (16.3)	0.411	0.597	0.465	0.025–0.472
				Controls	90	31 (27.2)	27 (44.6)	22 (18.2)	0.450	0.108		
5	c.384 + 116T/C	rs3783469	Intron 3	Cases	96	83 (81.6)	11 (13.8)	2 (0.6)	0.078	0.045	0.556	0.061–0.083
				Controls	89	78 (78.4)	11 (10.3)	0 (0.3)	0.062	0.534		
6	c.384 + 118C/T	rs681673	Intron 3	Cases	96	56 (55.5)	34 (35.0)	6 (5.5)	0.240	0.784	0.070	0.214–0.333
				Controls	89	41 (40.4)	38 (39.1)	10 (9.5)	0.326	0.790		
7	c.384 + 168T/C	rs532446	Intron 3	Cases	96	56 (55.5)	34 (35.0)	6 (5.5)	0.240	0.784	0.078	0.300–0.400
				Controls	88	41 (40.2)	37 (38.6)	10 (9.2)	0.324	0.708		
8	c.385-205A/G	rs3783474	Intron 3	Cases	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959	1.00	0.006
				Controls	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959		
9	c.385-174C/T	NA	Intron 3	Cases	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959	1.00	–
				Controls	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959		
10	c.385-137T/C	rs3171012	Intron 3	Cases	95	71 (71.6)	23 (21.7)	1 (1.6)	0.132	0.563	0.765	0.067–0.147
				Controls	95	72 (69.9	19 (23.2)	4 (1.9)	0.142	0.080		
11	c.492A/G	rs3783478	Exon 4	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.959	0.994	0.006
				Controls	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959		
12	c.498 + 27C/T	rs3783479	3′-UTR	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.959	0.994	0.006
				Controls	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.959		

SNP single nucleotide polymorphism, *ID* identification number, *dbSNP* Single Nucleotide Polymorphism Database, *MAF* minor allele frequency, *NCBI* National Center for Biotechnology Information, *UTR* untranslated region

[a] According to the nomenclature of the Human Genome Variation Society

[b] As expected under Hardy–Weinberg equilibrium (HWE)

[c] p value for deviation from HWE

[d] p value based on single-marker analysis for allelic frequency between both series

[e] Reported MAFs in the NCBI database for global samples and populations of European descent

genotyping data from both sample sets combined using the Haploview software (Barrett et al. 2005).

Electronic links

UCSC Genome Bioinformatics: http://genome.ucsc.edu/; NCBI dbSNP: http://www.ncbi.nlm.nih.gov/SNP/; SSPNN: http://www.fruitfly.org/seq_tools/splice.html; ESE Finder: http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi; PHASE: http://www.stat.washington.edu/stephens/software.html; WHAP: http://pngu.mgh.harvard.edu/~purcell/whap/; Haploview: http://www.broad.mit.edu/mpg/haploview; HapMap: http://www.hapmap.org;

Results

Analysis of *GADD45A* sequence variations

Although no truncating mutation was found in the *GADD45A* coding region of our French Canadian breast cancer cases, we identified 12 variants in *GADD45A* exonic and flanking intronic sequences, including two novel sequence variations (c.147-103G/C and c.385-174C/T) not reported in the NCBI Single Nucleotide Polymorphism Database (dbSNP Build 128) (Table 1). Among these sequence variations, one is a coding silent variant (c.492A/G, p.E164E), whereas the nine remaining sequence changes are intronic nucleotide substitutions. No deviation from HWE was observed for any of the nucleotide changes identified, with the exception of one intronic variation (c.384 + 116T/C), displaying borderline significance ($p = 0.045$) due to an excess of rare homozygotes among cases (Table 1). When considering all nucleotide variations, seven are common variants with minor allele frequency (MAF) higher than 5%, whereas five are considered as rare sequence variations, as they display an MAF below 5%. Among the rare variants, two (c.492A/G and c.498 + 27C/T) are observed simultaneously within the same breast cancer and control individuals, suggesting that both individuals carry a specific allele. The allele frequency of variants observed in the breast cancer series was also genotyped in healthy French Canadian controls, as denoted in Table 1. No significant difference in MAF was observed. In addition, all variants identified also displayed similar frequencies to those reported in the dbSNP, including the CEPH-Utah residents with ancestry from Northern and Western Europe (CEU) population.

As consensus-binding sites located in the *GADD45A* third intron are important regulatory sequences, the relevant genomic regions were also analyzed by direct sequencing. No sequence alteration was found directly within the consensus p53-binding region located in the 5′ end of the intron, the closest variation being situated 20 nucleotides downstream (c.384 + 168T/C, Table 1). Likewise, amplification of the BRCA1/ZNF350-binding region near the beginning of exon 4 in the 3′ part of the intron did not lead to identification of any sequence alteration directly within the consensus sequence. However, a rare variant (c.385-205A/G) located two nucleotides apart was identified in only one control and one breast cancer case, but no genomic DNA from other affected family members was available for testing. Another rare nucleotide change, present within the same amplicon (c.385-174C/T), was also identified in only one breast cancer case and one control individual, whereas variant c.385-137T/C (rs3171012) situated 137 nucleotides upstream of the fourth exon of *GADD45A* was present at an MAF that was similar for both series and within the range of frequencies reported (Table 1).

In silico analysis of the effect of variants on splicing

The possible effect of all intronic variants on splicing consensus sequences was also assessed using in silico analysis with the SSPNN program. None of the genetic variations observed displayed a significant change in the splicing score (data not shown), including the intronic variant most likely to have a potential effect (c.45-23C/T) given its proximal location to a known exon boundary. The strongest effects on the splicing score change were observed for the c.384 + 93A/G (rs3783468) and the c.385-174C/T variants. The c.384 + 93A/G variant could affect a pseudo donor site located 86 nucleotides downstream of the exon 3 3′ boundary by decreasing the splicing score from 0.60 to 0.41, whereas c.385-174C/T could create a weak pseudo acceptor site located 189 nucleotides upstream of exon 4 (splicing score of 0.42). In addition, this variant (c.385-174C/T) could also strengthen another pseudo acceptor site situated 164 nucleotides before exon 4 by increasing the splicing score from 0.68 to 0.76, which would be as strong as the splice site currently used, although no evidence indicates that this pseudo donor site is utilized by the splicing machinery. The putative impact of the c.492A/G exonic variant was also assessed using the ESE Finder program, which examines sequences for exonic splicing enhancer motifs. Although the c.492A/G variant slightly modified the putative ESE motifs identified, all changes were very close to the default thresholds and thus would be expected to act on relatively weak motifs, which would most likely not be used by the splicing machinery (data not shown).

Evaluation of *GADD45A* intragenic linkage disequilibrium

LD calculations for each SNP pair were performed in both series combined using $|D'|$ and r^2 measures and are

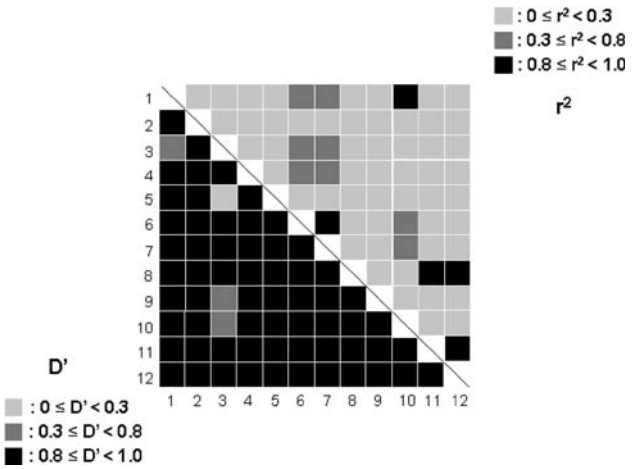

Fig. 1 Pairwise linkage disequilibrium (LD) measures of |D′| and r^2 for the 12 single nucleotide polymorphisms (SNPs) identified in the breast cancer and control series. SNP numbers are denoted according to Table 1

represented in Fig. 1. Perfect LD (|D′| = 1) was observed between the two most distant intragenic variants (SNP1: c.45-23C/T and SNP12: C.498 + 27C/T), which indicates that LD at the *GADD45A* locus does not decline significantly with distance. This is not surprising given that the *GADD45A* gene covers less than 4 kb of genomic sequence. Indeed, the lowest pairwise LD value involves c.147-53G/T (SNP3) with c.384 + 116T/C (SNP5) (|D′| = 0.158). As expected, r^2 coefficients calculated for *GADD45A* genomic region displayed lower values, as this measure is sensitive to variation in allelic frequency, which is well represented by the large spectrum of r^2 values ranging from 0.0001 to 1.0. A high r^2 coefficient was observed between c.384 + 118C/T (SNP6) and c.384 + 168T/C (SNP7), which displays high and similar MAF, whereas the majority of lowest r^2 values were observed for the three SNPs displaying an MAF below 1% (c.147-103G/C, c.492A/G and c.498 + 27C/T). As described above, both c.492A/G and c.498 + 27C/T are rare and are always observed within the same individuals. Hence, this pair displays perfect LD (r^2 coefficient of 1.0).

Haplotype analysis

Analysis of *GADD45A* haplotypes with the PHASE program, using the 12 sequence variations genotyped in both series, led to 18 estimated haplotypes (Table 2). Among these, five haplotypes (PH1, 4, 7, 9, and 15) displayed a frequency >5%, which represent 95.5% of all haplotypes estimated in both sample sets combined, whereas the remaining haplotypes displayed frequencies <1%. Estimated *p* value for case–control permutation did not indicate any significant difference in the global haplotype composition and frequency distribution between both groups

Table 2 Haplotypes as estimated by PHASE and WHAP programs and their estimated frequencies in cases and controls using all single nucleotide polymorphisms (SNPs) genotyped in both case and control series

Haplotype	SNP-1, -2, -3, -4, -5, -6, -7, -8, -9, -10, -11, -12	Haplotype frequencies (estimated)		
		Cases	Controls	Combined
PHASE				
PH1	CGGATCTACTAC	0.581	0.545	0.563
PH2	CGGATCTACCAC	0.003	0.000	0.002
PH3	CGGATTTACTAC	0.000	0.003	0.001
PH4	CGGGTCTACTAC	0.083	0.061	0.072
PH5	CGGGTTCACCAC	0.005	0.000	0.003
PH6	CGGGTTCGCTGT	0.005	0.005	0.005
PH7	CGGGCCTACTAC	0.070	0.053	0.061
PH8	CGTGTCTACTAC	0.005	0.000	0.003
PH9	CGTGTTCACTAC	0.103	0.174	0.139
PH10	CGTGCCTACTAC	0.000	0.005	0.003
PH11	CAGGCCTACTAC	0.003	0.003	0.003
PH12	TGGGTCCACCAC	0.005	0.005	0.005
PH13	TGGGTTTACCAC	0.005	0.000	0.003
PH14	TGGGTTCACTAC	0.010	0.000	0.005
PH15	TGGGTTCACCAC	0.104	0.133	0.118
PH16	TGGGTTCATCAC	0.000	0.003	0.001
PH17	TGGGCCTACTAC	0.003	0.000	0.001
PH18	TGTGTTCATCAC	0.004	0.000	0.002
WHAP				
WH1	CGGATCTACTAC	0.618	0.565	0.597
WH2	CGTGTTCACTAC	0.100	0.176	0.146
WH3	TGGGTTCACCAC	0.118	0.140	0.123
WH4	CGGGTCTACTAC	0.094	0.065	0.076
WH5	CGGGCCTACTAC	0.071	0.053	0.058

($p = 0.94$). Among the common haplotypes, only the haplotype PH9 showed a notable difference in frequency, with this haplotype being more represented among the controls. To further ascertain a possible association of this specific haplotype with breast cancer, a second haplotype estimation program, WHAP, was used, allowing for regression-based haplotype association testing. Estimation of haplotypes within the same data sets yielded the same five common haplotypes, displaying remarkably similar estimated frequencies as those obtained with the PHASE program (Table 2). Although the global test for association was not significant ($p = 0.16$), this haplotype-specific testing indicated a weak but significant association of haplotype WH2 with breast cancer ($p = 0.032$), with an overrepresentation of this specific haplotype among controls. Using a sliding window, we were able to circumscribe the specificity of this haplotype to the first six genetic variants without loosing the significant association

($p = 0.0158$, data not shown), with the strongest effect observed when considering the first nine variants ($p = 0.0155$).

Tagging SNPs determination

In order to reduce genotyping costs and efforts in future association studies without loss of power, the selection of a small fraction of SNPs, identified as tSNP—which likely represent with a good reliability the French Canadian population and the underlying variation at the *GADD45A* locus—represents an avenue of choice. Indeed, the selection of tSNPs that would be both frequent (MAF $\geq$0.05) and in strong LD with the other variants will allow capture of the common alleles of *GADD45A*. Thus, a causal variant may not be genotyped but will still be represented by the set of tSNPs. Therefore, genotyping data of the 12 variants from both series have been used for final haplotype block as well as tSNP identification. Based on an algorithm of solid block of LD, only one region of strong LD among the French Canadians was identified using the Haploview software (expectation maximization algorithm) (Fig. 2). This LD block encompasses the whole exonic region, given that SNPs 1 and 12 are comprised within the same LD block. Thereafter, considering exclusively haplotypes having a frequency $\geq$5%, four tSNPs were identified within this LD block, namely, variants c.384 + 93A/G, c.384 + 116T/C, c.384 + 118C/T, and c.385-137T/C (rs3783468, rs3783469, rs681673, and rs3171012). Given that only three genomic variants were available from HapMap data (rs3783468, rs681673, and rs532446) and that they are all located in intron 3, it was not relevant to perform similar analyses using Centre d'Etude du Polymorphisme Humain (CEPH)/CEU sample data. To further test the usefulness of our selected tSNPs in capturing haplotype diversity, a second haplotype analysis was performed using only the tSNPs selected here. These four tSNPs still allowed us to distinguish the five common haplotypes present in our data set, demonstrating the efficacy of the tSNPs identified here to accurately pinpoint all common haplotypes found in our population. However, when using only the three tSNPs present in the HapMap database, just three haplotypes could be determined, delineating the importance of the added tSNPs from the study.

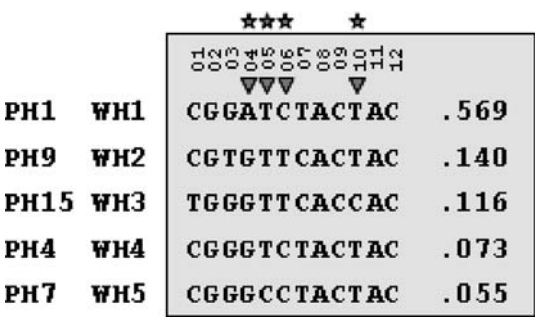

Fig. 2 Haploview predictions of haplotype blocks using single nucleotide polymorphisms (SNPs) displaying a major allele frequency (MAF) higher than 5%. The identification of SNPs is performed on a block-by-block basis and are denoted with an *asterisk* (*) above the SNP number. Tagging SNPs (tSNPs) are selected based on haplotypes showing an estimated frequency higher than 5%. Haploview estimation of haplotype frequencies is displayed on the *right*

Discussion

In search of additional breast cancer susceptibility genes, we undertook analysis of the *GADD45A* gene in our cohort of high-risk, non-*BRCA1/2* breast cancer families and healthy controls based on the close involvement of *GADD45A* with the products of the known predisposition genes *TP53* and *BRCA1*. All individuals were drawn from the French Canadian founder population, thus theoretically decreasing the underlying genetic heterogeneity. In addition, breast cancer cases were purposely selected from high-risk families (one per family), which has been demonstrated to increase statistical power (Antoniou and Easton 2003).

Although no truncating mutations were found, a total of 12 sequence variations were identified in our data sets, including two variations not previously reported. No nucleotide change on its own seemed to be significantly associated with breast cancer risk, including c.384 + 118C/T and c.384 + 168T/C, which are both near but not directly within intronic regions of strong conservation observed in *GADD45A* third intron (from UCSC conservation data, data not shown). Despite this, we can assume they have no functional significance given that these variants displayed similar frequencies in both sample sets. The only noncoding variation present in a region of strong conservation was the rare c.498 + 27C/T nucleotide change located in the 3′ untranslated region (UTR), which displayed similar frequencies in both groups.

Further LD analyses of the genetic variants identified highlighted relatively strong LD between all *GADD45A* sequence variations. Among these, variants c.385-205A/G, c.492A/G, and c.498 + 27C/T were in complete LD (for both |D′| and r^2 measures), with variants c.492A/G and c.498 + 27C/T being observed in the same case and control individuals, therefore suggesting that these variations are present on the same allele. As for c.385-205A/G, we can expect this variant to be present on the other allele. Indeed, although it was present on the same control individual as c.492A/G and c.498 + 27C/T, it was not present on the breast cancer sample bearing the other two variants. HapMap data within this genomic region displayed a block

of strong LD encompassing *GADD45A* and nearly the entire adjacent *GNG12* gene (G-protein gamma 12 subunit, data not shown), which is approximately 14 kb apart in the opposite direction and is implicated in cellular signal transduction. On its 5′ end, *GADD45A*'s nearest neighbor is the *SERBP1* gene involved in the binding of messenger ribonucleic acid (mRNA) 3′ ends. However, as the *SERBP1* gene is located more than 260 kb apart, it would not be expected to share LD with *GADD45A* according to HapMap data, although we could not exclude the possibility of a greater extent of LD in a founder population, such as the French Canadian population (Vézina et al. 2005; Laberge et al. 2005).

Given that the association of a gene with disease can be specific to certain alleles, the haplotype diversity of *GADD45A* was first estimated with the use of the PHASE software. PHASE estimated that an ensemble of 18 haplotypes solves all case and control genotypes observed, but no significant differences in haplotype structure and frequency were estimated between both series. However, a closer inspection of the attributed haplotypes between both groups highlighted a difference in frequency of the common haplotype PH9. Therefore, to further refine our haplotype analysis, we used the WHAP program, which enables an estimation of a possible haplotype-specific association with breast cancer risk and which confirmed that WH2 (which is identical to PH9) was indeed overrepresented among the control group. An analysis of this same haplotype using only a subset of variants indicated that a strong component of this specific difference is dependent on the 5′ end of the *GADD45A* gene.

Although none of the identified variants of *GADD45A* were clearly causative, we cannot rule out the possible effect of yet unidentified variants in other regulatory portions of the gene other than those analyzed here. Indeed, *GADD45A* is subject to complex multilevel regulation in response to various extracellular signals, and the association of transcriptional inductors and repressors such as p53 and ZNF350/BRCA1 with specific regions of its genomic sequence tightly regulates its expression. However, analysis of the consensus sequences located in intron 3 in both sample sets did not lead to identification of any genomic variants potentially disrupting these motifs.

Nonetheless, *GADD45A* is also regulated through post-transcriptional events, mainly changes in its mRNA stability through the binding of stabilization (HuR, nucleolin) and destabilization (AUF1) proteins on the 3′ UTR distal sequence. In addition, *GADD45A* expression has been recently shown to be regulated through a cap-independent and internal ribosome entry site (IRES)-dependent mechanism in response to cellular stress signals such as arsenic-induced cytotoxic and genotoxic damage (Chang et al. 2007). This expression is dependent on an IRES element located in the 5′ UTR region proximal to the start codon. As mutations of IRES elements or impairment of binding between transacting factors and IRES elements have been associated with cancer (Chappell et al. 2000), we carefully screened both sample sets for deleterious mutations that could potentially affect the expression through the IRES element of *GADD45A*. No sequence variant was identified within the *GADD45A* IRES sequence. However, it has to be stated that the possibility remains that changes in mRNA stability could be eventually associated with breast cancer risk.

Although *GADD45A* is part of a cellular pathway that includes several genes demonstrated to be involved in breast tumorigenesis (*ATM*, *TP53*, *BRCA1*), it has not been subjected to an extensive mutation search. Blaszyk et al. (1996) analyzed a series of sporadic breast cancer tumors (with and without p53 mutations) for alterations in *GADD45A* p53-binding site ($n = 53$ tumors) and coding sequence ($n = 26$ tumors) and only identified one polymorphism in intron 3. To our knowledge, the only other study of *GADD45A* alterations in familial breast cancer was performed on tumors from individuals showing a high incidence of early onset breast cancer ($n = 34$ families), male breast cancer ($n =$ two families), or breast and ovarian cancer ($n =$ seven families), the majority of which were not previously screened for the presence of a BRCA1 or BRCA2 mutation (Sensi et al. 2004). In agreement with what we found in non-BRCA1/BRCA2 high-risk breast cancer families, they did not identify any alterations of coding sequence, exon/intron boundaries, or the p53-binding domain.

Therefore, although *GADD45A* represents a very attractive breast cancer susceptibility candidate gene, our analysis and the current knowledge of *GADD45A* alterations in breast cancer individuals, cell lines, and tumors do not support a strong involvement of this gene with breast cancer risk. Nonetheless, additional investigations will be definitely needed to further ascertain the involvement of variations within the regulatory region of *GADD45A* with regard to breast cancer. However, given the limited information on sequence variant data available in the HapMap database, this study provides identification of tSNPs, which will be useful for further testing the association of the *GADD45A* gene in larger cohorts or with another syndrome or disease.

Acknowledgments The authors are indebted to the participants and their families for their generosity and providing DNA samples. We thank Damian Labuda and Claudia Moreau at the Centre de Recherche de l'Hôpital Ste-Justine for providing control DNA samples. We thank Dr. Martine Dumont, Gilles Leblanc, Carolle Samson, and Martine Tranchant for sample management, mutation screening, and skillful technical assistance, and Tina Babineau, Nathalie Bolduc, Claire Brousseau, Marie-Andrée Lajoie, Pascale Léger, Hélène Malouin, Andrée McMillan, and Josée Rhéaume for genetic counseling and clinical data management at the Cancer Genomics Laboratory.

We thank Professor Bartha Maria Knoppers and her colleagues from the Centre de recherche en droit public de l'Université de Montréal for their valuable help with ELSI issues related to our research program. We also appreciate advice received from ethics committees. This work was supported by the Canadian Institutes of Health Research (CIHR) through the INHERIT BRCAs research program, the Fonds de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the CURE Foundation, and the Canadian Breast Cancer Research Alliance (CBCRA). SD holds a studentship from Fondation René Bussières, JS is chairholder of the Canada Research Chair in Oncogenetics, and FD is a recipient of a chercheur-boursier from the Fonds de la Recherche en Santé du Québec (FRSQ) and a Research Career Award in the Health Sciences from CIHR/Rx&D Health Research Foundation.

Appendix

Other members of INHERIT BRCAs involved in clinical aspects of the study:

Paul Bessette: Department of Obstetrics and Gynecology, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, Canada

Peter Bridge: Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary, Alberta, Canada

Jocelyne Chiquette: Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, Canada

Rachel Laframboise: Medical Genetics Division, Centre Hospitalier Universitaire de Québec, CHUL, Laval University, Québec, Canada

Jean Lépine: Haemato-Oncology Service, Centre Hospitalier Régional de Rimouski, Rimouski, Canada

Bernard Lespérance, Roxane Pichette: Department of Haemato-Oncology, Hôpital du Sacré-Coeur de Montréal, Montréal, Canada

Marie Plante: Gynecology Oncology Division, Hôtel-Dieu de Québec, Centre Hospitalier Universitaire de Québec, Laval University, Québec, Canada

References

Antoniou AC, Easton DF (2003) Polygenic inheritance of breast cancer: implications for design of association studies. Genet Epidemiol 25:190–202

Antoniou AC, Easton DF (2006) Models of genetic susceptibility to breast cancer. Oncogene 25:5898–5905

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Blaszyk H, Hartmann A, Sommer SS, Kovach JS (1996) A polymorphism but no mutations in the *GADD45* gene in breast cancers. Hum Genet 97:543–547

Campomenosi P, Hall PA (2000) *GADD45* mutations are uncommon in human tumour cell lines. Cell Prolif 33:301–306

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acid Res 31:3568–3571

Chang Q, Bhatia D, Zhang Y, Meighan T, Castranova V, Shi X, Chen F (2007) Incorporation of an internal ribosome entry site-dependent mechanism in arsenic-induced GADD45 alpha expression. Cancer Res 67:6146–6154

Chappell SA, LeQuesne JP, Paulin FE, deSchoolmeester ML, Stoneley M, Soutar RL, Ralston SH, Helfrich MH, Willis AE (2000) A mutation in the c-myc-IRES leads to enhanced internal ribosome entry in multiple myeloma: a novel mechanism of oncogene de-regulation. Oncogene 19:4437–4440

Desjardins S, Belleau P, Labrie Y, Ouellette G, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M, Durocher F (2008) Genetic variants and haplotype analyses of the *ZBRK1/ZNF350* gene in high-risk non BRCA1/2 French Canadian breast and ovarian cancer families. Int J Cancer 122:108–116

Devlin B, Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29:311–322

Ding K, Zhou K, He F, Shen Y (2003) LDA—a java-based linkage disequilibrium analyzer. Bioinformatics 19:2147–2148

Durocher F, Labrie Y, Soucy P, Sinilnikova O, Labuda D, Bessette P, Chiquette, Laframboise R, Lepine J, B. Lesperance B, G. Ouellette G, Pichette R, Plante M, Tavtigian SV, Simard J (2006) Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. BMC Cancer 6:230–251

Fan W, Jin S, Tong T, Zhao H, Fan F, Antinore MJ, Rajasekaran B, Wu M, Zhan Q (2002) BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. J Biol Chem 277:8061–8067

Harkin DP, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD, Haber DA (1999) Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. Cell 97:575–586

Hollander MC, Sheikh MS, Bulavin DV, Lundgren K, Augeri-Henmueller L, Shehee R, Molinaro TA, Kim KE, Tolosa E, Ashwell JD, Rosenberg MP, Zhan Q, Fernandez-Salguero PM, Morgan WF, Deng CX, Fornace AJ Jr (1999) Genomic instability in Gadd45a-deficient mice. Nat Genet 23:176–184

Hollander MC, Philburn RT, Patterson AD, Wyatt MA, Fornace AJ Jr (2005) Genomic instability in Gadd45a−/− cells is coupled with S-phase checkpoint defects. Cell Cycle 4:704–709

Jin S, Fan F, Fan W, Zhao H, Tong T, Blanck P, Alomo I, Rajasekaran B, Zhan Q (2001) Transcription factors Oct-1 and NF-YA regulate the p53-independent induction of the GADD45 following DNA damage. Oncogene 20:2683–2690

Laberge AM, Michaud J, Richter A, Lemyre E, Lambert M, Brais B, Mitchell GA (2005) Population history and its impact on medical genetics in Quebec. Clin Genet 68:287–301

Lewontin RC (1964) The interaction of selection and linkage. Ii. Optimum Models, Genetics 50:757–782

Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA (2002) Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31:33–36

Purcell S, Daly MJ, Sham PC (2007) WHAP: haplotype-based association analysis. Bioinformatics 23:255–256

Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. J Comput Biol 4:311–323

Sensi E, Tancredi M, Aretini P, Cipollini G, Collecchi P, Naccarato AG, Viacava P, Bevilacqua G, Caligo MA (2004) Clinicopathological significance of GADD45 gene alterations in human familial breast carcinoma. Breast Cancer Res Treat 87:197–201

Simard J, Dumont M, Moisan AM, Gaborieau V, Malouin H, Durocher F, Chiquette J, Plante M, Avard D, Bessette P, Brousseau C, Dorval M, Godard B, Houde L, Joly Y, Lajoie MA, Leblanc G, Lepine J, Lesperance B, Vezina H, Parboosingh J,

Pichette R, Provencher L, Rheaume J, Sinnett D, Samson C, Simard JC, Tranchant M, Voyer P, Easton D, Tavtigian SV, Knoppers BM, Laframboise R, BridgeP, Goldgar D (2007) Evaluation of BRCA1 and BRCA2 mutation prevalence, risk prediction models and a multi-step testing approach in French–Canadian high-risk breast and ovarian cancer families. J Med Genet 44:107–121

Smith ML, Kontny HU, Zhan Q, Sreenath A, O'Connor PM, Fornace AJ Jr (1996) Antisense GADD45 expression results in decreased DNA repair and sensitizes cells to u.v.-irradiation or cisplatin. Oncogene 13:2255–2263

Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum Mol Genet 15:2490–2508

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Tront JS, Hoffman B, Liebermann DA (2006) Gadd45a suppresses Ras-driven mammary tumorigenesis by activation of c-Jun NH2-terminal kinase and p38 stress signaling resulting in apoptosis and senescence. Cancer Res 66:8448–8454

Vezina H, Durocher F, Dumont M, Houde L, Szabo C, Tranchant M, Chiquette J, Plante M, Laframboise R, Lepine J, Nevanlinna H, Stoppa-Lyonnet D, Goldgar D, Bridge P, Simard J (2005) Molecular and genealogical characterization of the R1443X-BRCA1 mutation in high-risk French–Canadian breast/ovarian cancer families. Hum Genet 117:119–132

Wang XW, Zhan Q, Coursen JD, Khan MA, Kontny HU, Yu L, Hollander MC, O'Connor PM, Fornace AJ Jr, Harris CC (1999) GADD45 induction of a G2/M cell cycle checkpoint. Proc Natl Acad Sci USA 96:3706–3711

Zhan Q, Chen IT, Antinore MJ, Fornace AJ Jr (1998) Tumor suppressor p53 can participate in transcriptional induction of the GADD45 promoter in the absence of direct DNA binding. Mol Cell Biol 18:2768–2778

Zheng L, Pan H, Li S, Flesken-Nikitin A, Chen PL, Boyer TG, Lee WH (2000) Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1. Mol Cell 6:757–768