Research article

Open Access

A screen for germline mutations in the gene encoding CCCTC-binding factor (CTCF) in familial non-*BRCA1/BRCA2* breast cancer

Xiao-Lei Zhou, Barbro Werelius and Annika Lindblom

Department of Molecular Medicine, Karolinska Institutet, Stockholm, Sweden

Corresponding author: Annika Lindblom (e-mail: annika.lindblom@cmm.ki.se)

Received: 7 Aug 2003 Revisions requested: 28 Aug 2003 Revisions received: 4 Feb 2004 Accepted: 12 Feb 2004 Published: 9 Mar 2004

Breast Cancer Res 2004, **6**:R187-R190 (DOI 10.1186/bcr774)

Abstract

Introduction: The CCCTC-binding factor (CTCF), known as a versatile transcription factor and chromatin insulator and to be involved in X inactivation, has also been suggested to be a tumour suppressor on 16q. We investigated 153 patients with familial non-*BRCA1/BRCA2* breast cancer for germline mutations in the *CTCF* gene.

Methods: Mutation screening of *CTCF* was performed by denaturing high-performance liquid chromatography followed by cycle sequencing.

Results: We found two sequence variants, ^{240}G→A in the 5′ untranslated region and ^{1455}C→T (S388S) in exon 4, in five familial breast cancer cases. Three of these five cases had both variants. Cases and controls showed the same prevalence for the two variants, which were found in linkage disequilibrium in most cases and controls.

Conclusion: The present study suggests that germline mutations in *CTCF* are not important as a risk factor for breast cancer.

Keywords: CTCF, familial breast cancer, mutation screening

Introduction

The CCCTC-binding factor (CTCF), best known as a versatile transcription factor and chromatin insulator (reviewed in [1]), has been suggested as a risk factor in familial breast cancer, possibly acting as a tumour suppressor. The gene was mapped to 16q22–24, a region frequently showing loss of heterozygosity (LOH) in sporadic and familial breast cancer [2–4], and LOH of this region has been shown to correlate to increased survival and late distant metastasis [4–6]. A tumour-specific rearrangement of *CTCF* exons was first reported in one primary breast cancer patient [7], and four additional *CTCF* somatic mutations were subsequently observed in a set of 130 cases of breast, prostate and Wilms tumours in another study [8].

Skewed X inactivation denotes that the choice of which of the two X chromosomes is inactivated in females is non-random [9]; it has been implicated in cancer development. In a study of patients with ovarian cancer, a higher frequency of skewed X inactivation was found in patients with invasive cancer than in patients with borderline cancer and healthy controls [10]. It was also found that young patients with breast cancer have a higher frequency of skewed X inactivation in blood cells than controls of the same age [11]. This finding indicated that skewed X inactivation might be a risk factor in breast cancer development, and suggested the involvement of as yet unknown X-linked genes, or genetic factors involved in the X inactivation process in the development of breast cancer in young females. Recently, CTCF was shown to be a candidate *trans*-acting factor for X inactivation choice [12], suggesting that it might therefore have a role in breast development.

In the present study we investigated *CTCF* for germline mutations in 153 cases from 139 non-*BRCA1/BRCA2* breast cancer families to determine whether *CTCF*

CTCF = CCCTC-binding factor; DHPLC = denaturing high-performance liquid chromatography; LOH = loss of heterozygosity; RT–PCR = reverse transcriptase–polymerase chain reaction.

mutations could have a role in breast cancer predisposition. Of the 139 families, 28 had previously been included in a genome-wide linkage analysis. The linkage analysis results showed that all the 28 families could share a putative predisposing gene in 16q22–24 (Luo *et al.*, unpublished data). In addition, 26 tumours from 26 of the 139 families included in the present study had previously been analysed for LOH [13]. Of the 26 tumours analysed, 16 were informative for at least one microsatellite marker at 16q. Only three tumours showed LOH at 16q and were investigated for germline mutations in the E-cadherin gene, and no pathogenic mutation was found [14]. Thus E-cadherin is not likely to be involved in our breast cancer cases.

Materials and methods

Patients

For *CTCF* analysis, 153 breast cancer patients were recruited from 139 families. In our previous genome-wide linkage analysis, 28 of the families were shown to share a common haplotype in 16q22–24, where the *CTCF* gene resides. All the families included were recruited through a clinicogenetic counselling procedure and were considered *BRCA1* and *BRCA2* negative [15,16]. In total, 90 cases were ascertained from high-risk families in which there were three or more first-degree affected relatives with breast cancer over at least two generations. The mean number of cases in each family was 3.3 and the mean age of diagnosis was 53 years of age. The other 63 cases were ascertained from low-risk families in which there were only two first-degree affected relatives with breast cancer, with a mean age of onset of 47 years. All the cases in this study were included in accordance with guidelines approved by the Ethics Committee at Karolinska Institutet. As control population we used 190 unrelated healthy relatives of patients recruited at Department of Clinical Genetics, Karolinska Hospital.

Denaturing high-performance liquid chromatography (DHPLC)

DHPLC analysis was performed with automated instrumentation equipped with a DNASep column, the Wave® nucleic acid fragment analysis system (Transgenomic, Santa Clara, CA), on the reverse transcriptase polymerase chain reaction (RT–PCR) product. The entire coding sequence of the gene *CTCF* was divided into six overlapping fragments; the primers used for RT–PCR amplification of each segment are given in Table 1. Total RNA of each sample was extracted from EBV-transformed lymphocytes using TRIzol® Total RNA Isolation Reagent kit (Invitrogen, Carlsbad, CA) and was reverse transcribed with random hexamers using the GeneAmp® RNA PCR kit (PE Biosystems, Foster City, CA) to generate cDNA. A 2 µl aliquot of cDNA was used in PCR amplification in 50 µl reaction volumes containing 10 pmol of sense and antisense primers for each fragment, 1.5 mM $MgCl_2$, 100 µM dNTPs, 1.25 U of AmpliTaq Gold polymerase (PE Biosystems) and 1 × PCR buffer supplied by the manufacturer. A universal Touchdown PCR protocol was employed to improve PCR specificity and to minimize PCR optimization. It consisted of an initial incubation of 95°C for 10 min, a first round of 6 cycles of 95°C for 30 s, 58°C for 45 s with 1°C decrement per cycle, 72°C for 45 s, a second round of 26 cycles of 95°C for 30 s, 53°C for 45 s, 72°C for 45 s, and a final extension step of 72°C for 7 min. PCR products were then denatured at 95°C for 5 min and cooled slowly in a PCR instrument at a rate of 1.5°C per cycle for 40 cycles. Each PCR product (10 µl) was then loaded on the Wave® instrument and eluted with a linear acetonitrile gradient consisting of buffer A (0.1 M triethylamine acetate; TEAA) and buffer B (0.1 M TEAA in 25% acetonitrile) at a constant flow rate of 0.9 ml/min. Specific values of the gradient ranges and the column temperature required for optimal resolution of each amplicon were determined by the WaveMaker software (Transgenomic) based on the sequence. A Mutation Standard (Transgenomic) was run with the samples analysed, to ensure the best performance of the DHPLC instrument for mutation detection. The elution profiles were recorded and analysed by HSM 7000 software (Transgenomic).

Cycle sequencing

Samples with an altered DHPLC profiles were reamplified and purified with QIAEX II gel extraction kit (Qiagen, Hilden, Germany). The bi-directional sequencing reaction was performed with ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kits version 2.0 (PE Biosystems) in accordance with the manufacturer's instructions, with the primers used in RT–PCR amplification as sequencing primers (Table 1).

Pyrosequencing

Pyrosequencing was adopted in this study to determine the frequency of the *CTCF* variants found in familial breast cancer cases in normal controls. Primers used in the pyrosequencing analysis are listed in Table 2. The conditions of PCR amplification were the same as those used in the DHPLC analysis except that genomic DNA rather than cDNA was used as template and the number of cycles was increased from 35 to 50. Biotin-labelled amplicons (30 µl) were mixed with 25 µl of BW buffer (10 mM Tris-HCl, 2 M NaCl, 1 mM EDTA and 0.1 Tween 20, pH 7.6) and immobilized on 20 µl of streptavidin-coated super paramagnetic beads (Dynabeads®, M-280–streptavidin; Dynal AS, Oslo, Norway) by incubation at 65°C for 15 min, with shaking. Single-stranded DNA was obtained by incubating the immobilized amplicons in 50 µl of 0.5 M NaOH for 5 min using a PSQ 96 Sample Prep Tool (Pyrosequencing AB, Uppsala, Sweden). Each sample (well) was washed twice with 100 µl of 1 × annealing buffer (200 mM Tris acetate and 50 mM magnesium

Table 1

List of the primers used for amplification and cycle sequencing in the polymerase chain reaction

Fragments	Sense (5′–3′) positions of primers[a]	Antisense (5′–3′) positions of primers	Size (bp)
1	^{162}GGAGAATGATTACGGACCTG^{181}	^{622}CTATGTTTATGGGCTGTTCCTC^{601}	461
2	^{519}AGTAATGGAGGGCACAGTG^{537}	^{1017}CGCATTAACCTCTGATAGCA^{998}	499
3	^{940}GAGGGCAAAGATGTAGATGTG^{960}	^{1415}CCAGTATGAGAGCGAATGTGA^{1395}	476
4	^{1366}GCCAGTGTAGAAGTCAGCAA^{1385}	^{1811}TCCTGTCTACAAGCGTAATCA^{1791}	446
5	^{1768}AAGCGCTTTAAGTGTGACCAG^{1788}	^{2184}TTCTACGGCAGGCTCCTC^{2167}	417
6	^{2044}GGGGAAAAGGAGGAGAA^{2061}	^{2520}TAAACACAGCCCAGAGAAGTC^{2500}	477

[a]Position according to GenBank reference cDNA sequence NM_006565. bp, base pairs.

Table 2

List of the primers used for pyrosequencing

Exons	Sense (5′–3′) positions of primers[a]	Antisense (5′–3′) positions of primers[b]	Sequencing positions of primers
5′ UTR	^{188}AAGAACAAGATGCGCTAG^{205}	Biotin-^{259}CTCCGTGGCTGCAAAGTCAGTT^{280}	^{216}GCTGACCAGGGGCTTGAGAGCTGGG^{239}
4	^{723}TCACATTCGCTCTCATACTGG^{742}	Biotin-^{968}CGGAGAAGCATTATCAATTC^{949}	^{758}AGTGCAGTTTGTGCAGTTATGC^{779}

[a]Position according to GenBank reference cDNA sequence NM_006565. [b]Position according to GenBank reference DNA sequence AF145470. UTR, untranslated region.

acetate). The immobilized strand was resuspended in 45 µl of 1 × annealing buffer containing 2 pmol of sequencing primer. Hybridization was performed by incubation at 80°C for 2 min, followed by cooling to room temperature. Finally, real-time pyrosequencing was performed on an automated 96-well pyrosequencer instrument with a PSQ SNP Reagent Kit (including all required enzymes and substrates) provided by the manufacturer (Pyrosequencing AB). The base sequence interpretation of the chromatograms and SNP genotype recognition were implemented automatically by the software and checked manually.

Results and discussion

In 153 familial non-*BRCA1*/*BRCA2* breast cancer cases, we found two sequence variants, ^{240}G→A in the 5′ untranslated region and ^{1455}C→T (S388S) in exon 4. Each variant was identified in four cases (4 of 153; 2.6%). The method used for studying this gene was selected to minimize the risk of missing a mutation for technical reasons. The DHPLC technique is known to be very sensitive; the fragment length and optimal conditions for the DHPLC assay were carefully designed and the RNA template used was expressed in sufficient amounts. Thus, the low mutation frequency obtained in the present study is not due to the methods used.

The ^{240}G→A alteration was identified in 12 of 186 normal controls (6.5%), and the ^{1455}C→T alteration in 10 of 188 normal controls (5.3%). Each variant had similar prevalences in controls and in familial breast cancer cases ($P = 0.10$ and $P = 0.21$, respectively; χ^2 test).

In addition, both variants were found to occur together in three of the cases. They might be in linkage disequilibrium and constitute a haplotype on the same chromosome. Alternatively, they might be on two different chromosomes. If so, they could each contribute a causative effect, acting in a recessive mode, and the concurrence of the two variants should be observed more often in familial cases than in normal controls as a result of genetic selection. On examining the normal controls, 10 of the 12 individuals with sequence alteration turned out to be carriers of both the variants, which was similar to the familial cases. Thus, these two variants were most probably inherited together on the same chromosome – that is, in linkage disequilibrium – in most breast cancer cases and controls. The concurrence of the two variants is not caused by genetic selection, and accordingly both the variants are considered non-pathogenic.

Conclusion

The present study suggests that germline mutations in the *CTCF* gene are not important as a risk factor in familial breast cancer. However, the *CTCF* gene could still have a role in cancer development. A tumour-specific truncating 14 base pair insertion in the *CTCF* gene was recently identified in one invasive ductal breast cancer case. The mutation resulted in silencing of the wild-type allele and loss of protein expression [17]. Moreover, CTCF is likely to participate in loss of imprinting of the gene encoding insulin-like growth factor II (IGF2) in colorectal cancer and Wilms tumours [18,19]. Further studies are therefore needed to elucidate the role of *CTCF* in breast cancer.

Competing interests

None declared.

Acknowledgements

We thank the families for their cooperation, and Victor Lobanenkov for valuable discussions. The study was financed by the Swedish Cancer Society, the Stockholm Cancer Society and the Gustav V Jubilee Foundation.

References

1. Ohlsson R, Renkawitz R, Lobanenkov V: **CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease.** *Trends Genet* 2001, **17:**520-527.
2. Cleton-Jansen AM, Moerland EW, Kuipers-Dijkshoorn NJ, Callen DF, Sutherland GR, Hansen B, Devilee P, Cornelisse CJ: **At least two different regions are involved in allelic imbalance on chromosome arm 16q in breast cancer.** *Genes Chromosomes Cancer* 1994, **9:**101-107.
3. Dorion-Bonnet F, Mautalen S, Hostein I, Longy M: **Allelic imbalance study of 16q in human primary breast carcinomas using microsatellite markers.** *Genes Chromosomes Cancer* 1995, **14:** 171-181.
4. Lindblom A, Rotstein S, Skoog L, Nordenskjold M, Larsson C: **Deletions on chromosome 16 in primary familial breast carcinomas are associated with development of distant metastases.** *Cancer Res* 1993, **53:**3707-3711.
5. Caligo MA, Polidoro L, Ghimenti C, Campani D, Cecchetti D, Bevilacqua G: **A region on the long arm of chromosome 16 is frequently deleted in metastatic node-negative breast cancer.** *Int J Oncol* 1998, **13:**177-182.
6. Hansen LL, Yilmaz M, Overgaard J, Andersen J, Kruse TA: **Allelic loss of 16q23.2–24.2 is an independent marker of good prognosis in primary breast cancer.** *Cancer Res* 1998, **58:**2166-2169.
7. Filippova GN, Lindblom A, Meincke LJ, Klenova EM, Neiman PE, Collins SJ, Doggett NA, Lobanenkov VV: **A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers.** *Genes Chromosomes Cancer* 1998, **22:**26-36.
8. Filippova GN, Qi CF, Ulmer JE, Moore JM, Ward MD, Hu YJ, Loukinov DI, Pugacheva EM, Klenova EM, Grundy PE *et al*: **Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity.** *Cancer Res* 2002, **62:**48-52.
9. Van den Veyver IB: **Skewed X inactivation in X-linked disorders**. *Semin Reprod Med* 2001, **19:**183-191.
10. Buller RE, Sood AK, Lallas T, Buekers T, Skilling JS: **Association between nonrandom X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer.** *J Natl Cancer Inst* 1999, **91:**339-346.
11. Kristiansen M, Langerod A, Knudsen GP, Weber BL, Borresen-Dale AL, Orstavik KH: **High frequency of skewed X inactivation in young breast cancer patients.** *J Med Genet* 2002, **39:**30-33.
12. Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT: **CTCF, a candidate *trans*-acting factor for X-inactivation choice.** *Science* 2002, **295:**345-347.
13. Lindblom A, Skoog L, Rotstein S, Werelius B, Larsson C, Nordenskjold M: **Loss of heterozygosity in familial breast carcinomas.** *Cancer Res* 1993, **53:**4356-4361.
14. Salahshor S, Haixin L, Huo H, Kristensen VN, Loman N, Sjoberg-Margolin S, Borg A, Borresen-Dale AL, Vorechovsky I, Lindblom A: **Low frequency of E-cadherin alterations in familial breast cancer.** *Breast Cancer Res* 2001, **3:**199-207.
15. Arver B, Borg A, Lindblom A: **First BRCA1 and BRCA2 gene testing implemented in the health care system of Stockholm.** *Genet Test* 2001, **5:**1-8.
16. Chen J, Hedman MZ, Arver BW, Sigurdsson S, Eyfjord JE, Lindblom A: **BRCA2 germline mutations in Swedish breast cancer families.** *Eur J Hum Genet* 1998, **6:**134-139.
17. Aulmann S, Blaker H, Penzel R, Rieker RJ, Otto HF, Sinn HP: **CTCF gene mutations in invasive ductal breast cancer.** *Breast Cancer Res Treat* 2003, **80:**347-352.
18. Cui H, Niemitz EL, Ravenel JD, Onyango P, Brandenburg SA, Lobanenkov VV, Feinberg AP: **Loss of imprinting of insulin-like growth factor-II in Wilms' tumor commonly involves altered methylation but not mutations of CTCF or its binding site.** *Cancer Res* 2001, **61:**4947-4950.
19. Nakagawa H, Chadwick RB, Peltomaki P, Plass C, Nakamura Y, de La Chapelle A: **Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of H19-associated CTCF-binding sites in colorectal cancer.** *Proc Natl Acad Sci USA* 2001, **98:**591-596.

Correspondence

Annika Lindblom, Department of Molecular Medicine, Karolinska Institutet, S 171 76 Stockholm, Sweden. Tel: +46 8 51775248; fax: +46 8 51773620; e-mail: annika.lindblom@cmm.ki.se