



HAL
open science

Mutation and association analysis of in breast cancer susceptibility

Clare Turnbull, Sarah Hines, Anthony Renwick, Deborah Hughes, David Pernet, Anna Elliott, Sheila Seal, Margaret Warren-Perry, D. Gareth Evans, Diana Eccles, et al.

► **To cite this version:**

Clare Turnbull, Sarah Hines, Anthony Renwick, Deborah Hughes, David Pernet, et al.. Mutation and association analysis of in breast cancer susceptibility. *Breast Cancer Research and Treatment*, 2010, 124 (1), pp.283-288. 10.1007/s10549-010-0949-1 . hal-00541154

HAL Id: hal-00541154

<https://hal.science/hal-00541154>

Submitted on 30 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mutation and association analysis of *GEN1* in breast cancer susceptibility

Clare Turnbull^{1*}, Sarah Hines^{1*}, Anthony Renwick¹, Deborah Hughes¹, David Pernet¹, Anna Elliott¹, Sheila Seal¹, Margaret Warren-Perry¹, D. Gareth Evans², Diana Eccles³, Breast Cancer Susceptibility Collaboration (UK), Michael R. Stratton^{1,4} and Nazneen Rahman¹

¹ Section of Cancer Genetics, the Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK, ²Regional Genetic Service. St Mary's Hospital, Manchester, UK, ³Wessex Clinical Genetics Service, Princess Ann Hospital, Southampton, UK, ⁴Cancer Genome Project, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

*These two authors contributed equally

Keywords: Breast cancer, Genetic susceptibility, DNA repair, Cancer genes.

Corresponding Author: Nazneen Rahman, Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, United Kingdom.

Phone: 44-208-7224026; Fax: 44-208-7224359; E-mail: nazneen.rahman@icr.ac.uk.

ABSTRACT

Purpose: GEN1 was recently identified as a key Holliday junction resolvase involved in homologous recombination. Somatic truncating *GEN1* mutations have been reported in two breast cancers. Together these data led to the proposition that *GEN1* is a breast cancer predisposition gene. In this paper we have formally investigated this hypothesis.

Methods: We performed full-gene mutational analysis of *GEN1* in 176 *BRCA1/2*-negative familial breast cancer samples and 159 controls. We genotyped six SNPs tagging the thirty common variants in the transcribed region of *GEN1* in 3,750 breast cancer cases and 4,907 controls.

Results: Mutation analysis revealed one truncating variant, c.2515_2519delAAGTT, which was present in 4% of cases and 4% of controls. We identified control individuals homozygous for the deletion, demonstrating that the last 69 amino acids of GEN1 are dispensable for its function. We identified 17 other variants, but their frequency did not significantly differ between cases and controls. Analysis of in 3,750 breast cancer cases and 4,907 controls demonstrated no evidence of significant association with breast cancer for six SNPs tagging the thirty common *GEN1* variants.

Conclusions: These data indicate that although it also plays a key role in double-strand DNA break repair, *GEN1* does not make an appreciable contribution to breast cancer susceptibility by acting as a high- or intermediate-penetrance breast cancer preposition gene like *BRCA1*, *BRCA2*, *CHEK2*, *ATM*, *BRIP1* and *PALB2* and that common *GEN1* variants do not act as low penetrance susceptibility alleles analogous to SNPs in *FGFR2*. Furthermore, our analyses demonstrate the importance of undertaking appropriate genetic investigations, typically full gene screening in cases and controls together with large-scale case-control association analyses, to evaluate the contribution of genes to cancer susceptibility.

INTRODUCTION

Breast cancer is twice as common in women with an affected first degree relative and germline mutations in the known breast cancer predisposition genes account for less than 30% of this excess familial risk. Inactivating mutations in *BRCA1* and *BRCA2* are high-penetrance breast cancer susceptibility alleles accounting for ~16%, whilst mutations in the functionally-related DNA repair genes *CHEK2*, *PALB2*, *ATM* and *BRIP1* are of intermediate penetrance, and account for less than 3% [1-7]. Genome wide association studies have identified 18 common variants which have been classed as low-penetrance breast cancer predisposition alleles. When combined these SNPs account for approximately 8% of familial disease risk [8-15].

The breast cancer predisposition genes *BRCA1*, *BRCA2*, *CHEK2*, *ATM*, *BRIP1* and *PALB2* are involved in double-strand break repair via the homologous recombination pathway [16-18]. This pathway repairs breaks caused by ionising radiation and mutagenic chemicals by utilising the homologous chromosome as a template for repair [19]. During this process a covalent link between each pair of homologous chromatids is formed and is known as a Holliday junction. Once repair is complete, these junctions are resolved by symmetrical nicking of the DNA strands, followed by separation and ligation to form two separate duplex molecules. Holliday junction resolvases mediate the transition from four covalently bonded chromatids to two separate duplex chromosomes [20]. In 2008, *GEN1* was identified as the gene encoding a human Holliday junction resolvase with a key a role in this recombinational repair pathway [21].

Through genome-wide sequencing of the exome in breast cancer cell lines and primary tumours, two somatic frameshift mutations in *GEN1* were identified [22,23]. This, together with recognition of the role of *GEN1* in DNA repair, led to the conclusion that constitutional *GEN1* mutations would confer susceptibility to breast

cancer in a fashion analogous to some other DNA repair genes [21]. However, to date, no data to support this conclusion have been published. In order to investigate formally the contribution of *GEN1* to breast cancer susceptibility, we have undertaken mutational analysis of the full gene in 192 breast cancer cases and 184 controls and an association analysis of common variants in the vicinity of *GEN1* in constitutional DNA from 3,750 breast cancer cases and 4,907 controls.

MATERIALS AND METHODS

Samples

Cases were unrelated individuals with breast cancer and a family history of breast cancer that were recruited through cancer genetics clinics in the UK, through the Genetics of Familial Breast Cancer Study. Informed consent was obtained from all family members and the research was approved by the London Multicentre Research Ethics Committee (MREC/01/2/18). Samples from non-Caucasian UK ethnic groups were excluded. The extent of family history was quantified using a Family History Score, defined by the number of relatives with breast cancer, weighted by their degree of relatedness to the index case. A score of 1.0 is assigned to the index case, with an additional 0.5 for each affected 1st degree relative, and an additional 0.25 for each affected 2nd degree relative. Where an individual has bilateral cancer their score is doubled. In the *GEN1* mutation screen we utilised 192 samples with a median Family History Score of 2.75 (range was 1.75-4). All cases were negative for *BRCA1* and *BRCA2* mutations and large deletions / duplications. In the *GEN1* association study we utilised 3,750 cases with a median Family History Score of 1.75 (range 1-5.25). *BRCA1/2* mutations had either been excluded (3,304) or the status was unknown (446).

Controls were obtained from the 1958 Birth Cohort Collection, an ongoing follow-up of persons born in Great Britain in one week in 1958 [24]. Informed consent has been

obtained for all blood samples in this collection to be used as a genetic resource. Additional controls for the genome wide association study were obtained from the United Kingdom Blood Services Collection of Common Controls established for the Wellcome Trust Case Control study, a collection of DNA samples from consenting blood donors of the English, Scottish and Welsh Blood Services [25]. Individuals of self-reported white ethnicity and representative of gender and each geographical region were selected.

GEN1 mutation analyses

We screened genomic DNA from the familial breast cancer case and control samples through the 13 coding exons and intron-exon boundaries of *GEN1* (Q17RS7) in 15 PCR fragments (Supplementary table 1). Following PCR, we carried out uni-directional sequencing using BigDyeTerminator Cycle sequencing kit and 3730XL automated sequencer (ABI). All variants and mutations were confirmed by separate bi-directional sequencing in a different aliquot of native DNA. We analysed the coding sequence and ten intronic flanking bases of each exon using Mutation Surveyor software version 3.20 (SoftGenetics) and visual inspection. Only samples successfully analysed through at least 90% of the *GEN1* coding sequence were included; we successfully mutationally screened 176/192 cases and 159/184 controls. We assessed the likely pathogenicity of variants using Polyphen (<http://genetics.bwh.harvard.edu/pph/>), SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>) and NNSplice (http://fruitfly.org:9005/seq_tools/splice.html) software *in silico* software.

To further evaluate the *GEN1* truncating variant c.2515_519delAAGTT, we extended mutation analysis of exon 13 to 536 cases and 525 controls in total. To investigate whether the variant causes nonsense mediated RNA decay, we extracted RNA from EBV transformed lymphoblastoid cell lines from two cases and two controls

heterozygous for the variant. We used SuperScript II Reverse Transcriptase (Invitrogen) to generate cDNA which was amplified, sequenced and analysed as described above using primers Forward–AAGGAGACCAGCTGCTTCAA and Reverse–GGAAGAGGGCTATCCAACA.

Statistical analyses

We performed comparisons of the frequencies between cases and controls of variants detected through mutational screening using a two-sided Fisher's exact test. We carried out a genome wide association study for breast cancer susceptibility alleles genotyping 3,960 breast cancer cases on a custom Illumina Infinium 670k array. Genotype frequencies were compared with those obtained on 5,069 controls genotyped on an Illumina Infinium 1.2M array, utilising data on 594,375 SNPs that were successfully genotyped on both arrays. We excluded closely related individuals (IBS probability > 0.86), individuals with > 15% non-European ancestry (by computing IBS scores between participants and individuals in HapMap and using multi-dimensional scaling) and restricted analyses to individuals that were called on > 97% of successfully genotyped SNPs. After these exclusions, 3,750 cases and 4,907 controls were used in the final analysis [9].

The transcribed region of *GEN1* extends from 17,798,661bp to 17,830,113bp on chromosome 2 (<http://genome.ucsc.edu/>) and contains 30 single nucleotide polymorphisms of minor allele frequency > 0.05 (<http://hapmap.ncbi.nlm.nih.gov/>). Linkage disequilibrium (LD) in the region was evaluated in 90 HapMap CEU individuals using a sliding window of 1,000kB and 10,000 SNPs. These LD data were used to select six SNPs from our dataset which tag these 30 SNPs in *GEN1* at $r^2 > 0.8$ (supp table 2). We undertook association testing using a 1df Cochran-Armitage test and a general 2df chi-squared test. Analyses were performed using Stata10 (State College, Texas, US) and PLINK (v1.06) software [26].

RESULTS

We successfully analysed the full coding sequence and intron-exon boundaries of *GEN1* in 176 individuals with familial breast cancer and 159 controls (Table 1). We identified one truncating variant, c.2515_2519delAAGTT, a five base pair deletion in the final exon of the coding sequence. We extended the analysis of this mutation which demonstrated that it was present in similar frequencies in case and control chromosomes (47/1047 cases vs 47/1050 controls) and both cases and controls homozygous for the deletion were identified (Fig 1A-C). This mutation is predicted to cause protein truncation generating a product lacking 69 amino acids (~8% of the protein) from the c-terminus. The mutation is in the last exon of *GEN1* and would be anticipated to escape nonsense-mediated RNA decay [27]. This was confirmed by analysis of cDNA from cases and controls heterozygous for c.2515_2519delAAGTT, which demonstrated equal proportions of the mutant and wild-type transcripts.

We also identified four synonymous and 13 non-synonymous *GEN1* variants. 14 variants were detected at similar frequencies in cases and controls including five common variants (frequency >0.05). Two rare non-synonymous variants were found in cases but not controls and two rare non-synonymous variants were found in controls but not cases. None of the variants were predicted to affect splicing. Only one variant, c.2692C>T p.R898C, was predicted to be deleterious by both Polyphen and SIFT algorithms but this was present in similar numbers of case (3/372) and control (6/360) chromosomes (Table 1).

We compared the frequency between 3,750 familial breast cancer cases and 4,907 controls of six SNPs which tag the 30 common variants in the genomic region encompassing *GEN1* (Supplementary Table 2). There was no evidence of significant association for any of these tag SNPs with breast cancer (Table 2).

DISCUSSION

GEN1 was recently identified as a Holliday junction resolvase with a key role in repair of DNA double-strand breaks. This function, together with the report of somatic *GEN1* mutations in two breast cancers led to the proposition that *GEN1* would act as a breast cancer susceptibility gene, similar to some other DNA repair genes [1-5,7]. In these recognised breast cancer susceptibility genes, *BRCA1*, *BRCA2*, *CHEK2*, *ATM*, *BRIP1* and *PALB2*, inactivating, primarily truncating, mutations confer high or intermediate risks of breast cancer. We identified a single *GEN1* truncating mutation, c.2515_2519delAAGTT. However, this deletion was present at equal frequency in cases and controls, indicating that it is not associated with appreciable increased risk of breast cancer. The deletion is in the final exon of the gene, results in truncation of less than 10% of the protein and mutant transcripts are not subjected to nonsense-mediated decay. Moreover, we identified several control individuals homozygous for the deletion, demonstrating that the last 69 amino acids of the GEN1 protein are dispensable for its function. This is consistent firstly with findings of Ip *et al.* who reported that a truncated form of GEN1, lacking the C-terminal, is sufficient for Holliday junction resolvase activity [21] and secondly with phylogenetic evidence which demonstrates strong conservation between GEN1 and its yeast homolog *yen1* over the first 480 amino acids, but very little in the C-terminal regions [21]. Our mutation screen did not identify any additional truncating mutations and there was no evidence that non-truncating variants are likely to be pathogenic.

Within recent years, common variants conferring small risks of breast cancer have been identified using large case-control series via genome-wide analyses of single nucleotide polymorphisms [8-15,28]. Of the 18 common, low penetrance breast cancer susceptibility alleles identified to date, none have been in regions containing DNA repair genes. We evaluated 30 common SNPs in the vicinity of *GEN1* by comparing the frequencies of six tag SNPs in 3,750 breast cancer cases and 4,907

controls and found no evidence to suggest that any common variants in this region are associated with breast cancer.

Our mutational screening data indicate that *GEN1* does not make an appreciable contribution to breast cancer predisposition by acting as a high-penetrance breast cancer predisposition gene akin to *BRCA1* and *BRCA2* or intermediate-penetrance breast cancer predisposition gene, similar to *ATM*, *BRIP1*, *CHEK2* or *PALB2*. The association analysis finds no evidence that common variation targeting *GEN1* confers susceptibility to breast cancer. Overall, these data strongly suggest that constitutional *GEN1* variation does not contribute to breast cancer predisposition. In addition our analyses demonstrate the importance of undertaking appropriate genetic investigations, typically full gene screening in cases and controls together with large-scale case-control association analyses, to evaluate the contribution of genes to cancer susceptibility.

ACKNOWLEDGMENTS

We thank all the patients and families that participated in the research. We gratefully acknowledge the clinicians and counsellors from the Breast Cancer Susceptibility Collaboration UK (BCSC) who coordinated recruitment and collection of the FBSC samples: A. Ardern-Jones, G. Attard, K. Bailey, C. Bardsley, J. Barwell, L. Baxter, R. Belk, J. Berg, N. Bradshaw, A. Brady, S. Brant, C. Brewer, G. Brice, G. Bromilow, C. Brooks, A. Bruce, B. Bulman, L. Burgess, J. Campbell, B. Castle, R. Cetnarskyj, C. Chapman, C. Chu, N. Coates, A. Collins, J. Cook, S. Coulson, G. Crawford, D. Cruger, C. Cummings, R. Davidson, L. Day, L. de Silva, B. Dell, C. Dolling, A. Donaldson, A. Donaldson, H. Dorkins, F. Douglas, S. Downing, S. Drummond, J. Dunlop, S. Durrell, D. Eccles, C. Eddy, M. Edwards, E. Edwards, J. Edwardson, R. Eeles, F. Elmslie, G. Evans, B. Gibbens, C. Giblin, S. Gibson, S. Goff, S. Goodman, D. Goudie, L. Greenhalgh, J. Greer, H. Gregory, R. Hardy, C. Hartigan, T. Heaton, C. Higgins, S. Hodgson, T. Homfray, D. Horrigan, C. Houghton, L. Hughes, V. Hunt, L. Irvine, L. Izatt, L. Jackson, C. Jacobs, S. James, M. James, L. Jeffers, I. Jobson, W. Jones, S. Kenwick, C. Kightley, C. Kirk, L. Kirk, E. Kivuva, A. Kumar, F. Lalloo, N. Lambord, C. Langman, P. Leonard, S. Levene, S. Locker, P. Logan, M. Longmuir, A. Lucassen, V. Lyus, A. Magee, S. Mansour, D. McBride, E. McCann, V. McConnell, M. McEntagart, K. McDermot, L. McLeish, D. McLeod, L. Mercer, C. Mercer, Z. Miedzybrodzka, J. Miller, P. Morrison, J. Myring, J. Paterson, P. Pearson, G. Pichert, K. Platt, M. Porteous, C. Pottinger, S. Price, L. Protheroe, L. Protheroe, S. Pugh, C. Riddick, V. Roffey-Johnson, M. Rogers, S. Rose, S. Rowe, A. Schofield, G. Scott, J. Scott, A. Searle, S. Shanley, S. Sharif, J. Shaw, J. Shea-Simonds, L. Side, J. Sillibourne, K. Simon, S. Simpson, S. Slater, K. Smith, L. Snadden, J. Soloway, Y. Stait, B. Stayner, M. Steel, C. Steel, H. Stewart, D. Stirling, M. Thomas, S. Thomas, S. Tomkins, H. Turner, E. Tyler, E. Wakeling, F. Waldrup, L. Walker, L. Walker, C. Watt, S. Watts, A. Webber, C. Whyte, J. Wiggins, E. Williams, L. Winchester. We thank Anita Hall and Darshna Dudakia for assistance in recruitment and Katrina Spanova and Bernadette Ebbs for running the sequencers. This work was funded by Cancer Research UK (C8620_A8372); US Military ACQ Activity, Era of Hope Award (W81XWH-05-1-0204) and the Institute of Cancer Research (UK). We acknowledge NHS funding to the NIHR Biomedical Research Centre. This study makes use of data generated by the Wellcome Trust Case-Control Consortium 2. A full list of the investigators who contributed to the generation of the data is available from <http://www.wtccc.org.uk>. We acknowledge use of DNA from the British 1958 Birth Cohort collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02.

REFERENCES

1. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al (2002) Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 31:55-9
2. Miki Y, Swensen J, Shattuck-Eidens D, et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266:66-71
3. Rahman N, Seal S, Thompson D, et al (2007) PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 39:165-7
4. Renwick A, Thompson D, Seal S, et al (2006) ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet* 38:873-5
5. Seal S, Thompson D, Renwick A, et al (2006) Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 38:1239-41
6. Stratton MR, Rahman N (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet* 40:17-22
7. Wooster R, Neuhausen SL, Mangion J, et al (1994) Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* 265:2088-90
8. Easton DF, Pooley KA, Dunning AM, et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447:1087-93
9. Turnbull C, Ahmed S, Morrison J, et al (2010) Identification of five new breast cancer susceptibility loci through a genome-wide association study. *Nat.Genet.* (In press)
10. Ahmed S, Thomas G, Ghoussaini M, et al (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat.Genet.* 41:585-590

11. Hunter DJ, Kraft P, Jacobs KB, et al (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat.Genet.* 39:870-874
12. Stacey SN, Manolescu A, Sulem P, et al (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat.Genet.* 39:865-869
13. Stacey SN, Manolescu A, Sulem P, et al (2008) Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat.Genet.* 40:703-706
14. Thomas G, Jacobs KB, Kraft P, et al (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat.Genet.* 41:579-584
15. Zheng W, Long J, Gao YT, et al (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat.Genet.* 41:324-328
16. Turnbull C, Rahman N (2008) Genetic predisposition to breast cancer: past, present, and future. *Annu Rev Genomics Hum Genet* 9:321-45
17. Venkitaraman AR (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 108:171-82
18. Zhang F, Ma J, Wu J, et al (2009) PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol* 19:524-9
19. Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* 27:247-54

20. Lilley DM, White MF (2001) The junction-resolving enzymes. *Nat Rev Mol Cell Biol* 2:433-43
21. Ip SC, Rass U, Blanco MG, et al (2008) Identification of Holliday junction resolvases from humans and yeast. *Nature* 456:357-61
22. Forbes SA, Bhamra G, Bamford S, et al (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet* Chapter 10:Unit 10 11
23. Wood LD, Parsons DW, Jones S, et al (2007) The genomic landscapes of human breast and colorectal cancers. *Science* 318:1108-13
24. Weblink: National Child Development Study,
<http://www.cls.ioe.ac.uk/studies.asp?section=000100020004>, 7-10-2008
25. WTCC (2007) Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661-78
26. Purcell S, Neale B, Todd-Brown K, et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-75
27. Lejeune F, Maquat LE (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr Opin Cell Biol*. 17:309-315
28. Cox A, Dunning AM, Garcia-Closas M, et al (2007) A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet*. 39:352-358

Table 1. Coding *GEN1* variants in breast cancer cases and controls

| Variant | dbSNP ¹ | Allele frequencies ² | | P-Value for association ³ |
|---------------------|--------------------|---------------------------------|----------|--------------------------------------|
| | | Cases | Controls | |
| c.274T>A; p.S92T | rs1812152 | 195/358 | 226/346 | 0.1 |
| c.428A>G; p.N143S | rs16981869 | 22/366 | 21/364 | 0.9 |
| c.566G>A; p.S189N | - | 6/362 | 6/328 | 0.9 |
| c.607A>G; p.I203V | rs10177628 | 3/362 | 0/328 | 0.1 |
| c.905G>A; p.R302H | - | 2/382 | 1/344 | 0.6 |
| c.988G>A; p.E330K | - | 0/380 | 1/340 | 0.3 |
| c.1341A>G; p.A447A | rs16983864 | 4/362 | 1/356 | 0.2 |
| c.1526C>G; p.S509W | - | 1/372 | 3/358 | 0.3 |
| c.1638T>A; p.S546S | - | 6/372 | 5/358 | 0.8 |
| c.1971A>G; p.E657E | rs300168 | 189/384 | 189/350 | 0.5 |
| c.2039C>T; p.T680I | rs300169 | 233/384 | 228/350 | 0.6 |
| c.2445C>T; p.Y815Y | - | 3/382 | 1/360 | 0.3 |
| c.2449A>G; p.T817A | - | 0/382 | 1/360 | 0.3 |
| c.2515_2519delAAGTT | - | 47/1072 | 47/1050 | 0.9 |
| c.2567C>T; p.S856F | - | 1/382 | 0/360 | 0.3 |
| c.2619T>G; p.S873R | rs57936182 | 4/372 | 1/360 | 0.2 |
| c.2644A>G; p.K882E | - | 6/372 | 7/360 | 0.7 |
| c.2692C>T; p.R898C | rs17315702 | 3/372 | 6/360 | 0.3 |

¹ www.ncbi.nlm.nih.gov/projects/SNP² The denominator for each variant indicates the number of chromosomes successfully sequenced³ P-value for 2-sided Fishers exact test (1df),

Table 2. Association with breast cancer of six SNPs tagging common *GEN1* variants

| Illumina tagSNP | Minor allele | MAF¹ cases | MAF¹ controls | P-value² |
|------------------------|---------------------|------------------------------|---------------------------------|----------------------------|
| rs7556886 | T | 0.19 | 0.19 | 0.97 |
| rs6761104 | A | 0.10 | 0.10 | 0.49 |
| rs300168 | A | 0.46 | 0.47 | 0.69 |
| rs300169 | G | 0.36 | 0.36 | 0.42 |
| rs17315736 | A | 0.09 | 0.09 | 0.49 |
| rs13031876 | C | 0.35 | 0.35 | 0.60 |

¹MAF: minor allele frequency

²Cochran Armitage trend test (1 degree of freedom), unadjusted for multiple testing

FIGURE LEGEND

Figure 1. Sequence traces for wildtype, deletion heterozygote and deletion homozygotes. Reverse sequencing chromatograms of the sequence encompassing the c.2515_2519delAAGTT deletion showing the wildtype sequence (A) deletion heterozygote sequence (B) and deletion homozygote sequence (C). The five deleted bases are indicated by the red square in wildtype sequence.

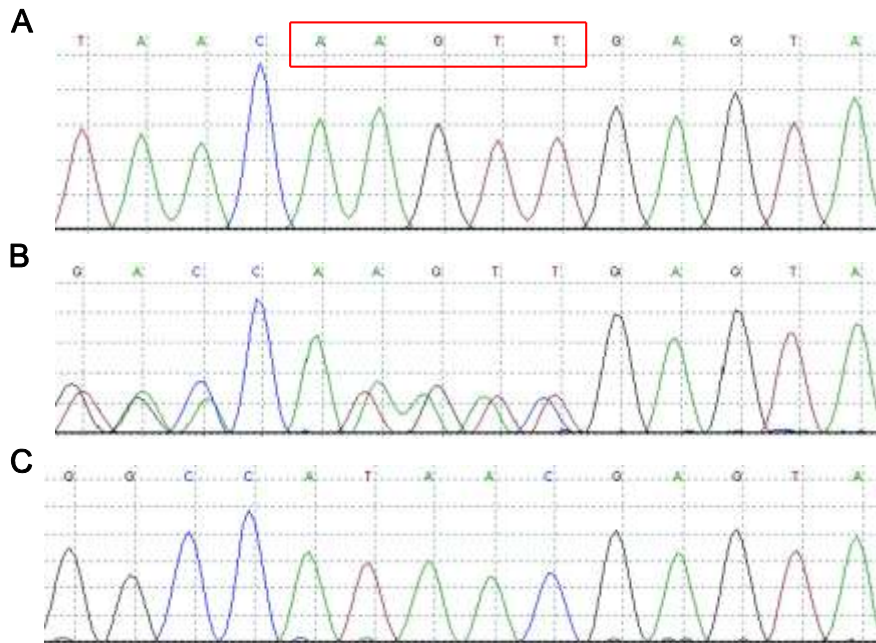


Figure 1 Sequence traces for wildtype, deletion heterozygote and deletion homozygote.