

AD _____

Award Number: DAMD17-00-1-0328

TITLE: Characterization of BRCA2 Mutation in a Series of
Functional Assays

PRINCIPAL INVESTIGATOR: Fergus J. Couch, Ph.D.

CONTRACTING ORGANIZATION: Mayo Clinic Rochester
Rochester, MN 55905

REPORT DATE: May 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20050916 176

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 01-05-2005		2. REPORT TYPE Final		3. DATES COVERED (From - To) 1 May 2000 - 30 Apr 2005	
4. TITLE AND SUBTITLE Characterization of BRCA2 Mutation in a Series of Functional				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER DAMD17-00-1-0328	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Fergus J. Couch, Ph.D.				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mayo Clinic Rochester Rochester, MN 55905				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Abstract is on page 4 of the report.					
15. SUBJECT TERMS Breast Cancer, BRCA2, missense mutation, functional assay					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			55

Table of Contents

Cover.....	1
SF 298.....	2
Table of Contents.....	3
Abstract.....	4
Introduction.....	5
Body.....	5
Key Research Accomplishments.....	8
Reportable Outcomes.....	8
Conclusions.....	9
References.....	9
Appendices.....	10

Abstract

Mutations in the BRCA2 gene account for approximately 20% of all hereditary breast cancer. Many individuals undergo expensive clinical testing for mutations in the BRCA2 gene in order to provide information to their family members about risk of breast cancer. The majority of mutations identified during clinical testing result in truncation of the protein, while approximately 20% are missense mutations. The influence of these missense mutations on BRCA2 function is not known. We proposed to develop a series of functional assays that could determine if particular missense mutations disrupt BRCA2 function and are disease associated. A series of assays that test the various functions of BRCA2 have now been successfully established and applied to a number of BRCA2 missense mutations. A number of these mutations have now been classified as either disease causing or neutral/benign based in part on these assays.

INTRODUCTION

The BRCA2 breast and ovarian cancer predisposition gene was identified in 1995. Mutations in the gene account for approximately 20% of all hereditary breast cancer and perhaps 2% of all breast cancer cases. Many individuals undergo expensive clinical testing for mutations in the BRCA2 gene in order to provide information to their family members about risk of breast cancer. The majority of mutations identified during clinical testing result in truncation of the protein, while approximately 45% are missense mutations. The affect of these missense mutations on BRCA2 function is not known. Thus, these mutations are termed unclassified variants and women carrying these mutations are informed that their results are inconclusive. To improve upon this situation we have proposed to develop a series of functional assays that can be used to determine if particular missense mutations disrupt BRCA2 function and are disease associated or not.

Body

Aim #1: To assess the role of BRCA2 in cell growth.

We proposed to establish Capan-1 BRCA2 mutant cells that stably express wildtype BRCA2. However, using multiple different transfection techniques we were unable to achieve greater than 0.01% transfection efficiency, which is insufficient for the outlined experiments. To overcome this problem we proposed to overexpress BRCA2 in MCF7 breast cancer cells because the cells were readily transfected and because overexpression of mutant forms could competitively inhibit wildtype BRCA2 function. We focused our efforts on the use of MCF7 cells stably expressing wtBRCA2, vector alone, and 6174delT-BRCA2. Expression was confirmed by western blot (Task 1). We also tried to generate MCF7 Tet-On and MCF7 Tet-Off inducible BRCA2 cell lines to no avail (Task 2 and 6). The stably expressing MCF7 cells were subsequently assessed for alterations in growth rate (MTT assays), anchorage independence (soft agar assays), colony formation, and cell cycle (FACS) (Task 3). In MTT assays two independent wtBRCA2 expressing cell lines grew significantly slower than vector transfected cells, while cells expressing mutant BRCA2 grew the most rapidly. In soft agar and colony formation assays more colonies formed from mutant or vector expressing cells. Cell cycle analysis revealed no change in the cell cycle over time. Overall, the data suggested that BRCA2 regulates cell growth and transformation in a cell cycle independent manner.

A total of 8 BRCA2 mutant constructs containing 7 missense mutations and 1 truncating mutation (K3326X) in different domains of BRCA2 were generated. The Y42C mutation is located in a putative transactivation domain, while E462G, P655R, 4812C>T, and K1690N are all located in the BRC repeat region that has been associated with DNA repair activity, and the D2723H, T2515I and R2659K variants are located in the DNA binding domain of BRCA2. Evaluation of these mutants showed that D2723H, R2659K, and P655R had no effect on cell growth or colony formation while the other mutations had effects similar to wildtype BRCA2 (Tasks 3 and 5). However, it must be noted that these cells already express wildtype BRCA2. Thus, the relevance of adding more BRCA2 to the endogeneous BRCA2 already present was unclear.

We subsequently obtained the chinese lung fibroblast cell line (VC-8) that represents the XRCC11 DNA repair deficient complementation group that is deficient in BRCA2. We established transfection conditions but were unable to generate stable cell lines for BRCA2 or BRCA2 mutants. This appears to be a common problem in all BRCA2 research laboratories, and suggests that continuous over-expression of wildtype or subtly mutated forms (missense mutations) of BRCA2 may induce cell death. With this in mind we returned to the use of transient transfection approaches using both BRCA2 deficient VC8 cells. Green fluorescent protein tagged forms of wildtype BRCA2 and mutant BRCA2 containing mutations in functional domains, including the recently established RAD51 binding domains and the DNA binding domain, were cloned into mammalian expression vectors and transiently transfected into VC8 cells. Cells were flow sorted for GFP signal and selected cells were replated and evaluated for cell growth over 5 days. The wildtype BRCA2 construct grows far slower than the GFP vector alone. However, the T2515I variant grew at similar rates to vector suggesting that this missense mutation affects the growth regulatory properties of BRCA2. Interestingly, this mutation is

located at the beginning of the DNA binding domain of BRCA2. We subsequently evaluated the effects of the 8 missense mutations on cell growth rate. The results mirrored those from the MCF7 cell lines.

However, other discoveries regarding the function of BRCA2 clearly established its primary function as a DNA repair protein. Thus, the effect on cell growth and colony formation appears to result from toxicity associated with overexpression. We did not undertake nude mouse experiments to assess effects of BRCA2 on tumorigenesis because the non-specific growth effects indicated that this would have represented an improper use of animals for experiments with no possible useful outcome (Task 4 and 7).

Aim #2: To test the effect of missense mutations and polymorphisms on BRCA2 function in a series of functional assays.

We generated 34 different GFP tagged full-length BRCA2 mutant cDNA expression constructs using site directed mutagenesis. A total of 28 of the mutants were missense mutations and 6 were in frame single codon deletions or insertions. Nine of the missense mutations were located in the BRC RAD 51 binding motifs and 17 were located in the DNA binding domain of BRCA2.

BRCA2 expression and localization

We transiently transfected each of the 34 constructs into 293T cells and verified expression by immunoprecipitation/western blot with anti-BRCA2 antibody. In addition, we evaluated the localization of these proteins in these cells by visualization of the GFP tag on a fluorescence microscope. Interestingly, the D2723H variant was localized exclusively to the cytoplasm whereas wildtype BRCA2 and polymorphic forms of BRCA2 were located predominantly in the nucleus. Evaluation of other variants has identified 8 missense mutations that cause BRCA2 protein to be restricted to the cytoplasm. This effect may be due to structural alterations caused by the amino acid changes rather than alteration of a specific nuclear localization domain. This location of the protein is expected to inactivate all nuclear functions and suggests that each of these 9 variants inactivate BRCA2 and predispose to cancer. Interestingly, all of these mutations are located in residues of BRCA2 that are perfectly conserved across one billion years of evolution, suggesting that the residues are critical for BRCA2 function. Thus, the sequence conservation data strongly supports the functional assay data.

BRCA2 enhances cell survival following DNA damage.

To evaluate the effect of BRCA2 expression on DNA repair we used clonogenic survival assays. BRCA2 deficient VC8 cells were transiently transfected with GFP-tagged wildtype *BRCA2*, vector alone, and missense mutant forms of BRCA2. GFP positive cells were selected by flow sorting and subsequently used in mitomycin c (MMC) dependent clonogenic survival assays. Colonies were enumerated after 5 days. These studies were repeated using cell counting by trypan blue exclusion to determine cell viability. In both cases, wildtype BRCA2 enhanced cell survival while the known deleterious mutant 6174delT showed no ability to promote cell survival relative to GFP vector alone. Furthermore, the Y42C and E462G isoforms behaved similarly to wildtype, while several other missense mutant forms (D2723H, R2659K, P655R, T2515I) did not suggesting that they inactivate BRCA2 function (Wu et al., 2005).

BRCA2 promotes homologous recombination.

We established a homologous recombination assay (Moynahan et al., 2001) in VC8 BRCA2 deficient cells in order to measure the influence of BRCA2 mutations on the ability of BRCA2 to mediate homologous recombination repair of double strand DNA breaks. VC8 cells with a stably integrated single copy of the DR-GFP reporter construct were generated. These cells were transiently transfected with BRCA2 constructs in combination with an I-SceI construct expressing an enzyme that induces a unique double strand break in the reporter plasmid. In this system, cells with active recombination due to the presence of wildtype BRCA2 express GFP while cells expressing inactive mutant BRCA2 display only background levels of GFP. The number of GFP positive cells can be enumerated by fluorescence activated cell sorting (FACS). Expression of wildtype BRCA2 in these VC8 cells promoted recombination relative to vector. Likewise the Y42C and E462G variants again functioned like wildtype. However, the 6174delT control and several missense mutants failed to induce recombination suggesting that these represent functionally inactive forms of BRCA2 (Wu et al., 2005).

Mutant BRCA2 is associated with centrosome amplification.

We found that the BRCA2 protein localizes to the centrosome and that truncated forms of BRCA2 cause centrosome amplification and aneuploidy in a dominant negative manner when expressed in 293 cells (data not shown) while wildtype BRCA2 can rescue centrosome amplification in BRCA2 deficient VC8 cells. These studies suggest that quantitation of centrosome amplification is a useful method for evaluating the ability of BRCA2 to regulate chromosomal instability. We performed immunofluorescence for the centrin and pericentrin components of the centrosome on cells transfected with wildtype and mutant BRCA2. The numbers of centrosomes in 200 GFP positive wildtype and mutant expressing cells were enumerated and compared. We found that ectopic expression of the D2723H BRCA2 mutant caused centrosome amplification in 293 cells while wildtype BRCA2 has no effect. Similarly all of the mislocalized mutants also caused centrosome amplification in 293T cells while variants that had no effect on localization or homologous recombination had no effect on centrosome number (Wu et al., 2005). Indeed, the results from the centrosome amplification assays were consistent with those from the DNA repair assays for all variants.

Evaluations of missense mutations in BRCA2 by genetic analysis.

While the functional assays appear to measure BRCA2 function, it is not yet clear if the results from these assays correlate with cancer risk. The only way to do this is to evaluate a number of mutations with known cancer risk. As very few BRCA2 missense mutations have been defined in terms of their associated cancer risk this approach to validation will not be possible.

Instead, we have decided to try to evaluate the specificity and sensitivity of the assays and the association between the assay results and cancer risk by comparing the results with co-segregation data for the mutations. Essentially, through a collaboration with Dr. David Goldgar at the University of Utah, the extent to which the missense mutations co-segregate with cancer in the families that are known to carry these mutations could be examined. This approach which is based on the model of Thompson and colleagues (Thompson et al., 2003), estimates the likelihood that the mutation is disease causing. As this measure is based on the observation of the disease in families it can be considered a gold standard. Thus, for some of the mutations evaluated in our functional assays, Dr. Goldgar provided a likelihood estimate of causality based on co-segregation. We were then able to combine this likelihood with similar likelihood calculations based on sequence conservation, the physico-chemical properties of the amino acid changes caused by the mutations, and the co-occurrence of the mutations with other known deleterious mutations in patient samples. The overall likelihood that a missense mutation predisposes to cancer was calculated for a number of mutations. Here we show some of the results from these analyses (Goldgar et al., 2004).

BRCA2 D2723H. This variant has been observed 24 times, but never with a proven deleterious mutation. We have analyzed 10 pedigrees and observed complete co-segregation with breast and ovarian cancer, yielding overall odds of 13723:1 in favor of causality. The Aspartate residue is completely conserved as far back in evolution as *tetraodon nigroviridis* (puffer fish). This mutation showed identical disrupted DNA repair capacity after exposure to gamma irradiation and mitomycin-c as the truncating mutation 6174delT. Moreover, the BRCA2 protein with D2723H showed aberrant cellular localization compared to wild type protein. Thus, for the D2723H variant all the sources of information are in agreement and thus this variant can be preliminarily classified as a deleterious *BRCA2* allele.

BRCA2 Y42C. Observed 92 times and 5 times with other deleterious mutations, giving overwhelming evidence against causality. Combined co-segregation analysis from 13 pedigrees gives odds against causality of $10^{5.5}$. The Tyrosine residue is conserved as far as chicken but is deleted in fish. Suggesting that it is not important for BRCA2 function. The Y42C BRCA2 protein displays similar function as wild-type BRCA2 protein in all assays described above. We conclude that the Y42C variant is not a high-risk mutation.

Note that we have not specifically addressed a number of the stated Tasks in Aim #2. However, it should be clear from this report that this is a result of improved understanding and knowledge of the function of BRCA2 that has occurred since the proposal was written. Importantly, we have conducted a large number of functional

studies in the spirit of the Aim and have succeeded in classifying the disease relevance of a number of mutations. Thus, Tasks 8-12 from Aim #2 have essentially been completed.

Aim #3: To test the effect of missense mutations and polymorphisms of BRCA2 on BRCA2 function in a yeast expression system.

In this Aim we proposed to evaluate the effect of BRCA2 mutants relative to wildtype on yeast cells. However, the finding that the BRCA2 gene is not present in yeast suggests that any efforts in this regard would be artificial and extremely difficult to interpret in terms of human disease. In fact, similar studies in yeast models of BRCA1 mutations have been met with some skepticism by the scientific community and the researchers involved have been required to repeat the studies in human cell lines, in a similar manner to our own approach. Thus, we decided that we would not attempt any of the experiments in Aim #3 but instead would focus our efforts on the cell line studies reported in Aims #1 and #2.

In summary, we have now established a small number of assays that we believe can discriminate between benign/neutral variants and disease associated variants. In the course of the studies we developed methods for improving ectopic BRCA2 expression levels in cell lines, improving BRCA2 plasmid transfection efficiency in cell lines, determining *BRCA1* expression levels, and evaluating BRCA2 function in both BRCA2 deficient and wildtype cell lines. Based on our results we propose that BRCA2 mutations can be evaluated by analysis of clonogenic survival, centrosome amplification, in vitro homologous recombination, and RAD51 focus formation. We have also determined that the combination of segregation data from families carrying these mutations with the results of functional assays is an effective method for classifying the disease relevance of BRCA2 and *BRCA1* missense mutations.

Key Research Accomplishments

- 1) Generated 34 different full length BRCA2 mutant cDNA expressions constructs and verified expression in 293T cells.
- 2) Evaluated expression of 34 BRCA2 constructs in 293T cells by immunofluorescence and determined the cellular localization of the ectopically expressed mutant proteins.
- 3) Established clonogenic mitomycin c survival assays of BRCA2 function in VC8 cells.
- 4) Established an assay that measures centrosome amplification induced by mutant BRCA2.
- 5) Established an assay that measures the homologous recombination repair activity of BRCA2 in VC8 cells.
- 6) Compared the results from the functional assays for a small number of common mutations with co-segregation of the mutations with disease in families.

Reportable Outcomes

Three manuscripts have been generated based on this work:

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ, and the Breast Cancer Information Core (BIC) Steering Committee. Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to BRCA1 and BRCA2. *Am J Hum Genet* 2004, 75(4):535-44.

Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer Res.* 2005, 65(2):417-426.

Lovelock PK, Healey S, Au W, Sum E, Tesoriero A, Wong EM, Hinson S, Brinkworth R, Bekessy A, Diez O, Izatt L, Solomon E, Jenkins M, Renard H, Hopper J, Waring P, Kconfab K, Tavtigian SV, Goldgar D, Lindeman GJ, Visvader JE, Couch FJ, Henderson BR, Southey M, Chenevix-Trench G, Spurdle AB, Brown MA. Genetic, functional and histopathological evaluation of two C-terminal *BRCA1* missense variants. *J Med Genet.* 2005, May 27, [Epub].

One grant has been awarded on the basis of this work:

ACS CCE-107497 Research Scholar
American Cancer Society

Couch (PI)

6/1/04-5/31/08

Cancer Predisposing BRCA1 and BRCA2 missense mutations

The goal of this award is to determine which mutations in *BRCA1* and *BRCA2* predispose to cancer by combining data on co-segregation of mutations with cancer in families, on co-occurrence of mutations with other deleterious mutations, inter-species sequence conservation, and functional assays.

Conclusions

Our conclusions at this point are that it is possible to identify biologically relevant missense mutations in the *BRCA2* gene using a series of functional assays. In this work, we have shown that *BRCA2* regulates centrosome amplification, that *BRCA2* has a direct suppression effect on cell growth, that *BRCA2* regulates survival in response to DNA damage, and that *BRCA2* mediates homologous recombination repair of double strand DNA breaks. We have also shown that certain missense mutations in *BRCA2* can be discriminated from wildtype *BRCA2* using assays that measure these functions of *BRCA2*. In addition, we have established the clinical relevance of two *BRCA2* missense mutations by combining data from the functional assays with co-segregation data from high-risk breast cancer families. We are now poised to evaluate a large number of different missense mutations in *BRCA2* using these assays in an effort to generate sufficient data to validate the assays.

References

1. Moynahan ME, Pierce AJ, Jasin M. *BRCA2* is required for homology-directed repair of chromosomal breaks. *Mol Cell*. 2001, 7(2):263-72.
2. Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet*. 2003, 73(3):652-5.
3. Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. Functional evaluation and cancer risk assessment of *BRCA2* unclassified variants. *Cancer Res*. 2005, 65(2):417-426.
4. Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ, and the Breast Cancer Information Core (BIC) Steering Committee. Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to *BRCA1* and *BRCA2*. *Am J Hum Genet* 2004, 75(4):535-44.

Personnel supported by award

Kangjian Wu, M.D.
Patricia Wendt, M.S.
Jianmin Wu, Ph.D.
Fergus J. Couch, Ph.D.

Appendices

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ, and the Breast Cancer Information Core (BIC) Steering Committee. Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to BRCA1 and BRCA2. *Am J Hum Genet* 2004, 75(4):535-44.

Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer Res.* 2005, 65(2):417-426.

Lovelock PK, Healey S, Au W, Sum E, Tesoriero A, Wong EM, Hinson S, Brinkworth R, Bekessy A, Diez O, Izatt L, Solomon E, Jenkins M, Renard H, Hopper J, Waring P, Kconfab K, Tavtigian SV, Goldgar D, Lindeman GJ, Visvader JE, Couch FJ, Henderson BR, Southey M, Chenevix-Trench G, Spurdle AB, Brown MA. Genetic, functional and histopathological evaluation of two C-terminal *BRCA1* missense variants. *J Med Genet.* 2005, May 27, [Epub].

Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to *BRCA1* and *BRCA2*

David E. Goldgar,¹ Douglas F. Easton,² Amie M. Deffenbaugh,³ Alvaro N. A. Monteiro,⁴ Sean V. Tavtigian,¹ Fergus J. Couch,⁵ and the Breast Cancer Information Core (BIC) Steering Committee*

¹International Agency for Research on Cancer, Lyon, France; ²Cancer Research UK, Genetic Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom; ³Myriad Genetics Laboratories, Salt Lake City; ⁴H. Lee Moffitt Cancer Center, Tampa, FL; and ⁵Mayo Clinic College of Medicine, Rochester, MN

Many sequence variants in predisposition genes are of uncertain clinical significance, and classification of these variants into high- or low-risk categories is an important problem in clinical genetics. Classification of such variants can be performed by direct epidemiological observations, including cosegregation with disease in families and degree of family history of the disease, or by indirect measures, including amino acid conservation, severity of amino acid change, and evidence from functional assays. In this study, we have developed an approach to the synthesis of such evidence in a multifactorial likelihood-ratio model. We applied this model to the analysis of three unclassified variants in *BRCA1* and three in *BRCA2*. The evidence strongly suggests that two variants (C1787S in *BRCA1* and D2723H in *BRCA2*) are deleterious, three (R841W in *BRCA1* and Y42C and P655R in *BRCA2*) are neutral, and one (R1699Q in *BRCA1*) remains of uncertain significance. These results provide a demonstration of the utility of the model.

Introduction

The identification of specific genes involved in a number of common diseases has resulted in the integration of genetic testing into clinical practice. For many of these genes, the sequence variants that are identified include known deleterious (often protein-truncating) mutations, recognized polymorphisms assumed to be neutral in terms of disease risk, and other variants (usually with missense changes) of uncertain clinical relevance. The last category poses problems for genetics counseling, since tested individuals and their families are given a seemingly ambiguous result, unless sufficient evidence is available that a given missense change is deleterious. In the case of the breast cancer susceptibility genes *BRCA1* (MIM 113705) and *BRCA2* (MIM 600185), these so-called unclassified variants (UCVs) account for approximately half of all unique variants detected (other than common polymorphisms) (see Breast Cancer Information Core [BIC] database Web site) and were identified

in 13% of all women tested in one study (Frank et al. 2002). Thus, if one accepts that more rigorous screening and/or other preventive measures are useful in lowering morbidity and mortality in individuals who carry a high-risk deleterious mutation in these genes, a relatively large number of them could be helped by the classification of these variants as neutral or deleterious. Although the present article focuses on *BRCA1* and *BRCA2*, similar issues occur in genetic testing for other common disorders for which major susceptibility genes have been identified.

To address this important clinical problem, various types of evidence may help to classify such variants as deleterious or neutral, with respect to the disease of interest. These include frequency of the variant in cases and controls, co-occurrence of the variant with a known deleterious mutation in one or more tested individuals (under the assumption that either homozygosity for true deleterious mutations is embryonically lethal or homozygotes will at least have a clearly recognizable phenotype), cosegregation of the variant with disease in families, occurrence of disease in relatives of index cases with a given variant, the nature and position of the amino acid substitution, the degree of conservation of amino acids among species, and the results of functional assays. Each of these sources of evidence has particular strengths and limitations in addressing the general problem of causality of sequence variants. These lines of evidence are summarized in table 1.

We and others have examined such classification

Received June 25, 2004; accepted for publication July 8, 2004; electronically published August 2, 2004.

Address for correspondence and reprints: Dr. David E. Goldgar, Unit of Genetic Epidemiology, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69008 Lyon, France. E-mail: goldgar@iarc.fr

* Members of the BIC Steering Committee are listed in the Acknowledgments section.

© 2004 by The American Society of Human Genetics. All rights reserved. 0002-9297/2004/7504-0002\$15.00

Table 1
Types of Evidence Potentially Useful for UCV Classification

Line of Evidence	Advantage(s)	Disadvantage(s)
Frequency in cases and controls	Provides a direct estimate of associated cancer risk	Variants are rare, so such studies would need to be prohibitively large (10,000+)
Co-occurrence (in <i>trans</i>) with deleterious mutations	If homozygotes and compound heterozygotes are assumed to be embryonically lethal (or vanishingly rare), we can often classify a variant as neutral on the basis of a single observation	Much less power to show causality; quantification is dependent on the assumed fitness of the homozygous genotype, which is not known with precision
Cosegregation with disease in pedigrees	Easily quantifiable and directly related to disease risk; not susceptible to uncertainties in mutation frequencies or population stratification	Requires sampling of additional individuals in the pedigrees (particularly additional cases), which may be difficult to achieve
Family history	Usually available for most variants without additional data or sample collection; potentially very powerful	Dependent on family ascertainment scheme; could be biased in stratified populations with heterogeneous ascertainment, so not as robust as cosegregation; power may be low for infrequent variants
Species conservation and amino acid-change severity	Can be applied to every possible missense change in the <i>BRCA1</i> and <i>BRCA2</i> genes; does not require extensive family history; complete conservation is predictive if enough evolutionary time sequence is available	Only indirectly related to disease risk; the magnitude of odds ratios is not sufficient to classify variants without additional information
Functional studies	Can evaluate biologically the variant's effect on the protein's ability to perform some key cellular functions	May only be relevant for variants in certain domains of the protein; the function tested may not be related to cancer causation
Loss of heterozygosity	Straightforward to quantify as an adjunct to cosegregation data; robust	Requires tumor material
Pathological classification	Potentially powerful for <i>BRCA1</i> tumors in which the pathological characteristics are quite distinct; quantifiable	Prediction is weak when routine pathology data are used; systematic evaluation requires tumor material; weakly predictive for <i>BRCA2</i>

schemes by use of a variety of approaches. For example, in terms of cosegregation of variants within a pedigree, Thompson et al. (2003) provided a method of calculating odds of causality for UCVs by use of complete pedigree data. Petersen et al. (1998) performed a similar study but used a more restricted approach. In terms of conservation of amino acids across species, a number of studies have been done. Miller and Kumar (2001) validated the hypothesis that missense variants at highly conserved/invariant residues would more often be deleterious, whereas highly variable changes would more likely be neutral. With regard to *BRCA1*, Fleming et al. (2003) and Abkevich et al. (2004) used the conservation of *BRCA1* residues in a variety of mammalian and non-mammalian species to make preliminary classifications of 139 (Fleming et al. 2003) and 146 (Abkevich et al. 2004) putative missense mutations. In terms of functional assays, examination of *BRCA1* has been limited to two functional domains: the RING finger (residues 24–64) and the BRCT domain (residues 1642–1863). Functional mammalian and yeast-based assays have focused on transcriptional activation by the BRCT domain (Vallon-Christersson et al. 2001). Recently, Mirkovic et al. (2004) used the three-dimensional protein structure to develop a rule-based system for the classification of variants, applying this approach to 57 observed constitutional missense variants in the BRCT domain of *BRCA1*. For *BRCA2*, analyses of functional domains have focused on the DNA-binding region between amino acids 2373 and 3256 (Yang et al. 2002) and on the eight 40–amino-acid BRC repeats in exon 11 that are associated with interaction of *BRCA2* with the RAD51-recombination and DNA-repair protein (Wong et al. 1997; Chen et al. 1999; Davies et al. 2001).

A comprehensive model is needed, in which all these sources of evidence can be used together to create a combined assessment of a particular sequence variant of interest. In this comprehensive model, both quantitative and qualitative evidence would be properly weighted to arrive at a final classification. Ideally, the end result would be the overall odds of causality—that is, the ratio of the likelihood of the observed data under the hypothesis of causality to that under the hypothesis of neutrality. If all of the various types of evidence were quantifiable in the same way, this would be straightforward. However, each type of evidence depends on different models and underlying assumptions, and some are more suitable to quantification and formulation as a likelihood ratio than others. Here, we focus primarily on the relevant data that can be evaluated directly on a genetic/epidemiological basis, as these data are easily quantifiable in terms of likelihood ratios; moreover, they are most directly related to the clinical outcome of interest—that is, the risk of developing cancer for a carrier of the particular sequence variant under consideration.

Methods

For clarity, we assumed that all variants in the gene of interest can be classified into two categories: “mutations” (M) that predispose to a high risk of breast and ovarian cancer and “neutral variants” (V) that cause no risk. Thus, we make the important simplifying assumption that variants do not have an intermediate risk. Almost all protein-truncating variants are known, with high probability, to be mutations. The aim is to determine whether or not other variants are likely to be deleterious mutations. These include amino acid substitutions, in-frame deletions, silent mutations, and some intronic changes. We would like to determine statistically the posterior probability that each variant (V) is a mutation (M), given the available data:

$$\Pr(M|\text{Data}) = \frac{\Pr(\text{Data}|M)\Pr(M)}{\Pr(\text{Data}|M)\Pr(M) + \Pr(\text{Data}|V)\Pr(V)}$$

The statistical analysis focuses on the likelihood ratio ($\Pr[\text{Data}|M]/\Pr[\text{Data}|V]$). The choice of an appropriate prior probability ($\Pr[M]$) that a new variant is a mutation is uncertain. However, given that there is a high frequency of such variants and that only a few of the variants can be unequivocally classified as mutations, it is clear that the probability is low. At least 70% of the families with breast or ovarian cancer that exhibit clear linkage to *BRCA1* or *BRCA2* have been shown to harbor deleterious mutations; a significant fraction of the remaining families (at least those linked to *BRCA1*) harbor large-scale rearrangements. We believe that the prior probability of a given UCV being deleterious is <10% and may be closer to 1%. This suggests that the appropriate likelihood threshold for declaring a variant to be deleterious should be at least 1,000:1. The appropriate threshold for declaring against causality is not as critical, since this decision does not affect genetics counseling. For the purposes of classification in the BIC, we suggest a likelihood ratio of 100:1 against causality as a useful criterion. Of course, the choice of threshold in each clinical situation will vary according to the particular circumstances.

Specific Contributions of Individual/Family Data Components

Co-occurrence with deleterious mutations.—A variety of mouse studies (Gowen et al. 1996; Liu et al. 1996; Hohenstein et al. 2001) have indicated that homozygosity for *Brca1* is embryonically lethal. This finding is reinforced by the clear deficit of *BRCA1* homozygotes and compound heterozygotes, compared with expected

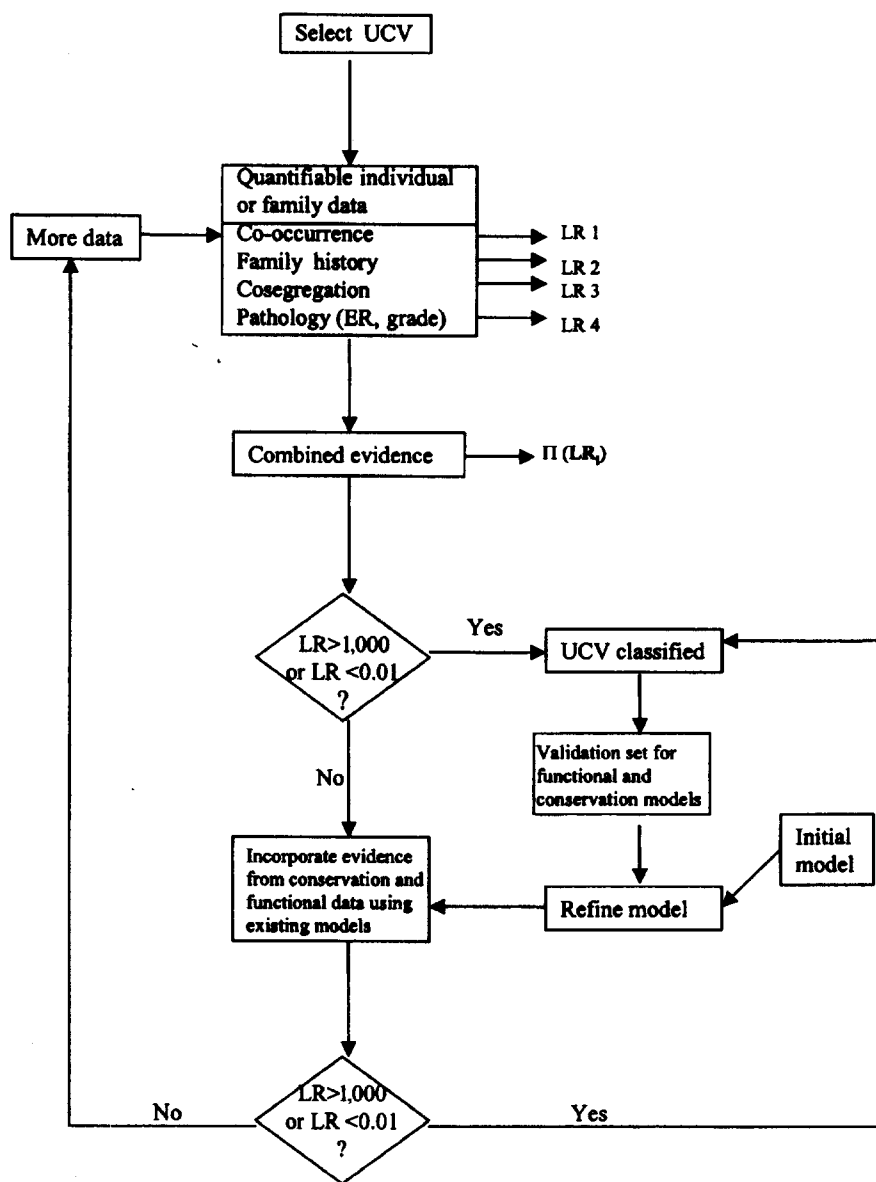


Figure 1 Flowchart of the procedure for classification of sequence variants of unknown clinical significance. ER = estrogen receptor status; LR = likelihood ratio.

numbers, among a series of individuals with the founder mutations 185delAG and 5382insC (0 observed vs. 6.5 expected) (Frank et al. 2002; Abkevich et al. 2004). For each *BRCA1* variant under consideration, we first examined the frequency of the mutation in the Myriad Genetics Laboratories database, which contains complete full-sequence data for both *BRCA1* and *BRCA2* from >20,000 individuals, as well as rudimentary family and patient history. In the following analysis, we assumed that individuals homozygous for a deleterious mutation in *BRCA1* or *BRCA2* are extremely rare. If

the variant is neutral, the probability of an individual with the variant also carrying (in *trans*) a deleterious mutation, p_1 , can be roughly estimated as half the overall frequency of deleterious mutations in the population being studied. If the variant is deleterious, this probability becomes

$$p_2 = \frac{\text{Pr}(\text{Individual carries deleterious mutation})}{\text{Pr}(\text{Individual carries variant and individual phenotype})}$$

Thus, if one observes the variant n times, k of which are in conjunction with a deleterious mutation, the appropriate likelihood ratio is given by the following binomial likelihood ratio:

$$\frac{(p_2)^k(1 - p_2)^{n-k}}{p_1^k(1 - p_1)^{n-k}}$$

For *BRCA1*, the frequency of deleterious mutations in the Myriad Genetics Laboratories database is 0.088 (1,765 known *BRCA1* deleterious mutations in 20,000 tests). Taking into account the evidence that *BRCA1* homozygotes and compound heterozygotes are vanishingly rare and quite likely to be embryonically lethal, we assumed $p_2 = 0.0001$ for these calculations.

For *BRCA2*, the corresponding estimate for the frequency of deleterious mutations is 0.059. The fitness issue here is slightly more problematic, since *BRCA2* compound heterozygotes have been found among individuals with the rare recessive disease Fanconi anemia type D1 (Howlett et al. 2002; Wagner et al. 2004). However, it is reasonable to assume that compound heterozygotes for deleterious mutations in *BRCA2* are extremely rare in adults, since the Fanconi anemia phenotype usually leads to death in early childhood. Taking into account the additional uncertainty associated with *BRCA2* homozygosity, we assumed $p_2 = 0.001$ for these calculations. One complication that arises in these data is the distinction between mutations occurring in *cis* and those in *trans*. Although the parental origin of the mutations is rarely known, mutations occurring in *cis* can often be recognized by recurrent observation of the same mutation/variant combination, and we have ignored these instances in our calculations.

The frequency of variants in groups of individuals, classified by likelihood of being a mutation carrier (i.e., family history).—A substantial amount of family history information is available for *BRCA1* and *BRCA2*. The most important source, given the scope and completeness of the genotyping, is the data obtained from sequencing by Myriad Genetics Laboratories. The rationale here is that mutation prevalence is known to be strongly dependent on certain key factors (disease status of the proband, age at diagnosis, and number and age of relatives with breast or ovarian cancer), so these characteristics should also predict the prevalence of a new disease-causing variant, whereas the prevalence of a neutral variant should be independent of family history. As a “proof of principle,” we have examined the confirmed deleterious missense mutation *BRCA1* C61G, for which there are 57 occurrences in the Myriad Genetics Laboratories database with family history information available. We compared the family histories of these 57 index cases with those of all known deleterious mutations in the database by use of a multinomial likelihood-ratio model, resulting in odds in favor

of causality of $>1,000,000:1$, showing the potential utility of this approach, at least for relatively frequent variants.

Cosegregation data.—To assess causality from the cosegregation data, we used the statistical model described by Thompson et al. (2003). For these calculations, we assumed an allele frequency of the variant of 0.0001 and used the *BRCA* penetrance estimates that were based on the recent meta-analysis of 22 population-based studies (Antoniou et al. 2003), with pooling across age groups, if necessary, depending on the level of detail of the family history information. Although family-based estimates might be more appropriate, we preferred to use these estimates, since the criteria for testing differ markedly among testing centers and the use of the population data would, if anything, be conservative. We do not, at present, allow for the possibility that a variant observed in the proband is a *de novo* mutation, although this could easily be incorporated into the model. Because, in many cases, complete pedigree data were unavailable, we relied on crude family history information and constructed complete pedigrees by creating individuals of unknown phenotype and genotype to connect the individuals in the pedigree. Note that, since analysis of cosegregation is conditional on the phenotypes in the family, the data on cosegregation can be considered independent of the data on family history (FH). The data from the co-occurrence of the variant with deleterious mutations are independent of the other information as well, so that these three likelihood ratios can be evaluated independently and multiplied:

$$\frac{\Pr(\text{Data}|\text{M})}{\Pr(\text{Data}|\text{V})} = \frac{\Pr(\text{FH}|\text{M})}{\Pr(\text{FH}|\text{V})} \times \frac{\Pr(\text{Cosegregation}|\text{M})}{\Pr(\text{Cosegregation}|\text{V})} \\ \times \frac{\Pr(\text{Co-occurrence}|\text{M})}{\Pr(\text{Co-occurrence}|\text{V})}$$

Incorporation of the Data on Sequence Conservation, Nature of Substitution, and Functional Characteristics

These data are more difficult to evaluate statistically than the data described above, since there is no direct link between these data and cancer risk. Our approach was to start with an initial model that was based on the limited number of already-classified missense variants for which data are available, and then, using the individual-specific data described above, we iteratively refined the estimated parameters as variants were classified into either deleterious or neutral categories. We describe below, in more detail, some initial models for this process.

Severity of the amino acid substitution.—The idea here is to use a score for the type of substitution and to derive the likelihood ratio on the basis of the distribution of this score in known neutral variants and known dele-

terious mutations. One approach is to use the chemical-difference matrix proposed by Grantham (1974) to produce a score (Grantham matrix score [GMS]) for the observed substitution in the variant that is being investigated (GMS_{UV}). We then determined the probability density function of the two distributions of scores, $f(GMS; \theta_M)$ and $f(GMS; \theta_V)$, where the form and parameterization, θ , of $f()$ depends on the distribution of the data. The likelihood ratio for these data is then given by

$$\frac{f(GMS_{UV}; \theta_M)}{f(GMS_{UV}; \theta_V)}$$

As a preliminary strategy for incorporating these data, we calculated the mean and SD of the GMS in known deleterious *BRCA1* missense mutations (excluding those that are known to be splice mutations), as well as that for known missense changes that are clearly neutral (e.g., common polymorphisms). For true deleterious missense mutations, the mean and SD were 133 and 65, respectively, whereas, for neutral variants, the corresponding values were 65 and 39. Given the apparent relationship between the mean and SD, we assumed that the distribution of $f(GMS; \theta)$ was lognormal, although, at present, there are insufficient numbers of known deleterious and neutral variants available to test the fit to this (or any other) distribution. This approach assumes that the mechanism of action in cancer causation is the change in the protein associated with the missense UCV. For variants near the intron/exon boundary, however, this assumption may not be valid, and the variant may be associated with disease through alternative splicing. To avoid this problem, such variants could be evaluated for their potential effect on splicing by use of a predictive algorithm, such as that used in the Berkeley *Drosophila* Genome Project (see Berkeley *Drosophila* Genome Project Web site). If possible, these variants were assessed through evaluation of alternative splicing by use of mRNA from blood samples of patients carrying the variant.

Conservation of the variant amino acid across species.—Although mutations at fully conserved amino acids are plausibly likely to be deleterious, it is not known whether such mutations are invariably associated with an increased cancer risk. Using sequence data from the genes orthologous to human *BRCA1* and *BRCA2* in six and four additional species, respectively, Abkevich et al. (2004) derived a mathematical model for *BRCA* sequence variation in which they postulated two types of amino acid substitutions: one under functional constraint and therefore slowly substituting (SS), and the other under no selective pressure and therefore fast substituting (FS). Thus, a UCV that results in an amino acid substitution at an SS position might be expected to be deleterious, whereas a UCV that occurs at an FS position is more

likely to be neutral. On the basis of the observed multiple sequence alignments and a mathematical model, the relative fraction of the two types of changes can be estimated for each possible number of different residues seen in the multiple sequence alignment, and the relative odds of a variant being of either type can be calculated under the model. For example, under this model, a UCV in *BRCA1* that changes a completely conserved amino acid is 10.4 times (125:12) more likely to be of the SS variety (and, hence, more likely to be deleterious). If this classification were completely concordant with the risk classification, these would also be the odds in favor of causality. For *BRCA2*, a similar procedure can be used, although the limited number of species for which sequence data are available reduces the discriminatory power. As more *BRCA* sequence data become available, these models will undoubtedly be improved.

Functional data.—These data are perhaps the most difficult to put into a likelihood-based framework. This is because there are a number of functional assays, each of which potentially tests a different function of the protein. To incorporate these data into the model, it will be necessary to have a larger set of variants with both (1) clear classification (according to the specified thresholds) of the deleterious and neutral categories and (2) functional data from a variety of different assays. For this reason, we have used functional data as qualitative supporting evidence, without directly incorporating these data into the likelihood-based evaluation.

On the basis of the data for which we have good initial models relevant to cancer risk, we can easily combine the relevant odds of causality. Those variants that are classified with high probability (i.e., with odds for or against causality reaching predefined thresholds) can then be used to evaluate and refine statistical models relating to functional or sequence-conservation data. As more variants are classified, these models will become more discriminating and, hence, more useful in the classification of variants for which there is insufficient family history and cosegregation data to achieve a clinically useful level of evidence for or against causality. This process is detailed in the flowchart in figure 1.

Results

To illustrate the model, we have selected three UCVs in *BRCA1* and three in *BRCA2* for analysis with the approaches described above. The likelihood ratios for each of the components in the analysis, as well as the combined odds for each of the six variants analyzed, are discussed below and are summarized in table 2.

Table 2

Odds in Favor of Each Variant Being Deleterious for the Six Variants Discussed in the Text, for Each Source of Information and Overall

DATA SOURCE	ODDS IN FAVOR OF CAUSALITY FOR					
	BRCA1			BRCA2		
	C1787S	R1699Q	R841W	Y42C	P655R	D2723H
Co-occurrence	1.2	1.4	.028	8.9×10^{-11}	.007	2.0
Cosegregation	1,694	2.84	4×10^{-9}	6.7×10^{-7}	.48	13,731
GMS	1.5	.48	1.31	3.49	1.35	.98
Conservation	10.4	10.4	.006	.194 ^a	.004 ^a	5.0
Overall odds	31,692	20	8.7×10^{-13}	4×10^{-17}	.00002	134,563

^a Deleted residue counted as a substitution.

BRCA1

C1787S.—This variant has been observed four times, but it has not been detected in any individual who also carried a clear deleterious mutation. Two available families show evidence of cosegregation with disease-yielding combined odds in favor of causality from the cosegregation data of 1,694:1. Incorporation of the data on co-occurrence yields overall odds in favor of causality of 2,032:1. Thus, on the basis of the family data alone, this variant could be classified as a disease-associated mutation. The cysteine residue is completely conserved, including in *Xenopus* and in the pufferfish *Tetraodon*. The substitution to serine is associated with a GMS of 112, compared with the average GMS for known polymorphisms of 60, the expected GMS value of 78 for a random missense change, and a GMS of 133 for 16 previously characterized deleterious missense mutations. The genomic data give odds of 15.5:1 in favor of the variant being deleterious, consistent with the pedigree data. This sequence variant has not yet been characterized functionally, but its effect on the three-dimensional protein structure has been modeled, and it is predicted to impact protein function (Mirkovic et al. 2004). It should be mentioned that the C1787S variant is always seen (presumably in *cis*) with an additional variant, G1788D.

R1699Q.—This mutation has been observed seven times in the Myriad Genetics Laboratories database, but it has never been detected in an individual with a deleterious mutation. This provides odds of 1.4:1 in favor of it being a deleterious mutation. Three small families with multiple individuals who were tested for this variant were available for analysis, leading to an overall cosegregation-based odds ratio of 2.8:1 in favor of causality for this variant. The combined odds from these two sources are 4:1 in favor of causality, and, therefore, this variant cannot be classified on the basis of this evidence alone. As with C1787S, the arginine residue is completely conserved. However, the change from arginine to glutamine yields a GMS of 43, lower than many of the known polymorphic substitutions. The combined odds ratio from the

genomic data is 4.99:1, again slightly in favor of causality. In mammalian cells, this sequence variant showed clear loss of transcriptional activation capability (Vallon-Christersson et al. 2001). It should be noted that another alteration in this same codon, R1699W, is considered by Myriad Genetics Laboratories to be a deleterious mutation, on the basis of both functional (Koonin et al. 1996; Vallon-Christersson et al. 2001) and cosegregation data.

R841W.—This variant has been observed in the Myriad Genetics Laboratories database 57 times, with 1 of those observations occurring in an individual who also carried a known deleterious mutation. Analysis of cosegregation in six pedigrees with multiple individuals tested showed quite convincing evidence against this variant being a high-risk allele (250,000,000:1 against causality). Thus, this variant can be unequivocally assigned to the neutral/nondisease-associated sequence variant category.

In contrast to the previous two *BRCA1* UCVs discussed above, this residue shows considerable variation among the various orthologues, with three alternative amino acids present and no conservation other than in the *Pan troglodytes* sequence. The amino acid associated with this variant, tryptophan, is not found in any of the five other species with sequence data available. The sequence and substitution data give odds of ~130:1 against causality, which supports the genetic data. Barker et al. (1996) have suggested that this variant may be associated with a modest increased risk of breast cancer. Since our approach considers only the hypotheses that the variant is high penetrance or is neutral, we cannot exclude the possibility that R841W is associated with a more moderate risk.

BRCA2

Y42C.—This mutation has been observed 144 times, 8 of which were in patients who also carried a known *BRCA2* deleterious mutation in *trans* with Y42C. We analyzed 17 pedigrees with this UCV and the overall odds against causality from these data were ~1,500,000:

1. Thus, on the basis of the pedigree cosegregation data alone, the odds are overwhelming against causality, and the co-occurrence data provides, if anything, even stronger evidence against causality.

For this variant, the tyrosine residue is conserved in chicken but is deleted in the *Tetraodon* sequence. The change from a tyrosine to a cysteine is one of the most severe changes, as measured by the GMS (194). Thus, the evidence based on sequence conservation and severity of the amino acid substitution is equivocal (combined odds, 1.3:1 against causality). However, as noted above, the co-occurrence and cosegregation data are overwhelmingly against Y42C being a deleterious *BRCA2* allele.

P655R.—This variant has been detected 63 times, twice with a known deleterious mutation. Ten pedigrees were analyzed for this variant and, taken together, exhibited weak evidence against causality (2:1). The combined evidence from the pedigree and co-occurrence data is 298:1 against causality, which would exceed our suggested threshold for classifying this as a neutral variant. This residue is conserved in rat and dog but is deleted in chicken and *Tetraodon*. The proline-to-arginine change is associated with a GMS of 103, a score that is between the average value for neutral changes and the value for *BRCA1* deleterious mutations.

D2723H.—This variant has been observed in the Myriad Genetics Laboratories database 24 times and has never appeared with a proven deleterious mutation. The variant yields odds in favor of causality under the *BRCA2* co-occurrence model of 2.0:1. All 10 pedigrees with multiple individuals tested for this variant showed complete cosegregation with breast and ovarian cancer, yielding overall odds of 13,731:1 in favor of causality. Thus, the pedigree data provide odds of ~57,000:1 in favor of causality—more than sufficient to classify the variant as deleterious by use of the suggested threshold of 1,000:1. The aspartate residue is completely conserved as far out as *Tetraodon*, although the GMS for this substitution is only 81. A *BRCA2* protein carrying this variant showed disrupted DNA-repair capacity after exposure to gamma irradiation and mitomycin-c, similar to the deleterious truncating mutation 6174delT. Moreover, in 293T human embryonal kidney cancer cells, the *BRCA2* protein with D2723H showed aberrant cellular localization, compared with the wild-type protein (K. Wu, S. Hinson, A. Ohashi, S. Tavtigian, A. Deffenbaugh, D. Goldgar, and F. Couch, unpublished data). Thus, in our view, the *BRCA2* D2723H variant can be classified unequivocally as a deleterious *BRCA2* allele.

Discussion

Although we have focused on the *BRCA1* and *BRCA2* genes, many of the methods described here are quite general and can be used for any hereditary disease in

which the genes responsible are characterized by many sequence variants for which it is difficult to assess a clear association with disease. As genetic testing for common, multifactorial diseases moves into clinical practice, the problems associated with the interpretation of sequence variants of unknown significance will result in psychological stress for patients and families and an increased burden on genetics counselors. In addition to the obvious clinical utility of developing and implementing a rigorous classification procedure for UCVs, the process could raise interesting questions about the biological basis of the disease predisposition conferred by the gene being studied. If, for example, a particular sequence variant shows conclusive evidence of causality on the basis of epidemiological data but functions normally in a specific assay, this would lead us to infer that the function being tested is perhaps not relevant to the disease process.

In addition to the main factors discussed extensively above, a number of other pieces of data could aid the classification of unknown sequence variants. These might be somewhat dependent on the disease and the gene being studied. For example, for *BRCA1*, we could take advantage of the fact that there is strong evidence that the pathology of *BRCA1* tumors differs from that of tumors in noncarriers of the same age (Breast Cancer Linkage Consortium 1997; Lakhani et al. 1998). Provided that one assumes that the pathological characteristics of tumors are not dependent on other familial factors, the odds based on pathological characteristics can be multiplied across all tumors carrying that specific germline UCV.

Another piece of information that could potentially be incorporated into such models, at least for many tumor-suppressor genes, is loss of heterozygosity (LOH) in tumors carrying the putative causal variant. For example, in *BRCA1*, ~85% of tumors exhibit LOH at *BRCA1*, compared with ~30% of breast cancers in noncarriers. Moreover, the LOH invariably involves the wild-type chromosome (Cornelis et al. 1995). Similar arguments apply to *BRCA2* and to several other cancer-predisposition genes. Methods for incorporating LOH data into linkage analysis have been developed (Rebeck et al. 1994), and this approach could be used to extend the cosegregation analysis.

For almost all the lines of evidence we have considered, it is clearly easier to obtain high odds in favor of neutrality than it is to show causality. This is similar to the situation in linkage analysis in which a single recombinant event is sufficient to exclude tight linkage but a much larger number of events is required to provide significant evidence in favor of linkage. It should be emphasized that our classification evaluation is based on the relative likelihood of the observed data under two specific hypotheses: that of complete neutrality of the variant (i.e., it confers no increased risk of disease)

and that of what we have termed as "causality" (i.e., the risk of disease conferred by the variant under consideration is comparable to the risk conferred by known mutations). Whether this is appropriate for all disease genes (and for which ones) is an important consideration in the application of this approach to a particular problem. If, for example, a particular missense mutation were associated with an intermediate risk, it might be classified in the deleterious or the neutral category, depending on the type of data available. Clearly, in this situation, more sophisticated models will be required. One of the expectations for such intermediate-risk variants is that they will prove difficult to classify, in spite of a substantial amount of data. This is a result of the potential for conflicting data from the various sources, which would make it difficult to achieve the specified thresholds for classification as either neutral or deleterious. If sufficient pedigrees are available for such variants and if these pedigrees have a reasonable number of individuals typed for the variant, it may be possible to estimate directly (through pedigree or case-control studies) the risk associated with the variant, although the derived estimate is likely to have wide confidence limits.

The classification of variants should ideally be based on clinical observations, since these are directly related to cancer risk and, hence, are the most relevant and also the most straightforward to quantify. On the basis of clinical information and our assumed models and thresholds, we were able to classify five of the six variants we studied as either deleterious or neutral. The additional value of the genomic data in these cases was less clear, but, in general, the genomic data supported the clinical data. It is interesting to note that, in each case, the score derived from species conservation pointed in the same direction as the clinical data, although the odds were much weaker. The GMSs, however, were inconsistent, giving odds in the opposite direction in four of the five classifiable cases and calling into question the utility of this measure in the classification process.

In summary, we believe that this multidisciplinary approach to evaluation of sequence variants of unknown significance provides a system of checks and balances and avoids overreliance on one source of information. This should result in more-reliable classification of such variants, which in turn will improve the clinical utility of genetic tests now being offered to patients and their families. The work presented here represents only the first step in an ongoing process. Additional work remains to be done, including the examination of the robustness of the method against violations of basic assumptions, the incorporation of this uncertainty into the model, the validation of each of the individual components through the accumulation of large amounts of additional data, and

the development of other approaches to the integration of the various components into a comprehensive model.

Acknowledgments

Members of the BIC Steering Committee are (in alphabetical order) Merete Bjørnslett (Department of Cancer Genetics, Norwegian Radium Hospital, Oslo), Larry Brody (National Human Genome Research Institute, Bethesda), Georgia Chenevix-Trench (Queensland Institute of Medical Research, Brisbane, Australia), Fergus J. Couch (Mayo Clinic College of Medicine, Rochester, MN), Amie M. Deffenbaugh (Myriad Genetics Laboratories, Salt Lake City), Peter Devilee (Department of Human Genetics, Leiden University, Leiden, Netherlands), Douglas Easton (Cancer Research UK, Genetic Epidemiology Unit, Strangeways Research Laboratory, Cambridge, United Kingdom), Charis Eng (Clinical Cancer Genetics Program, Division of Human Genetics, Ohio State University, Columbus), William Foulkes (McGill University, Montreal), David Goldgar (Unit of Genetic Epidemiology, International Agency for Research on Cancer, Lyon, France), Kathi Malone (Fred Hutchinson Cancer Research Center, Seattle), Alvaro N. A. Monteiro (H. Lee Moffitt Cancer Center, Tampa, FL), Kate Nathanson (Department of Medicine, Medical Genetics, University of Pennsylvania School of Medicine, Philadelphia), Susan Neuhausen (Epidemiology Division, Department of Medicine, University of California at Irvine, Irvine), Sharon Plon (Department of Pediatrics, Baylor College of Medicine, Houston), Csilla Szabo (Lyon, France), Sean Tavtigian (Unit of Genetic Cancer Susceptibility, International Agency for Research on Cancer, Lyon, France), and Tom Walsh (Department of Genetics, University of Washington, Seattle). This work was supported by National Institutes of Health awards CA92309 (to A.N.A.M.), CA81203 (to D.E.G.), and CA82267 (to F.J.C.). F.J.C. also acknowledges support from the Breast Cancer Research Foundation and the U.S. Army Medical Research and Materiel Command (DAMD17-1-00-0328). D.F.E. is a Principal Research Fellow of Cancer Research UK. D.F.E. and D.E.G. received support for this work from the Canadian Institutes of Health Research through the INHERIT BRCA research program. The authors gratefully acknowledge the technical assistance of Helene Renard, Colette Bonnardel, and Yvette Granjard.

Electronic-Database Information

The URLs for data presented herein are as follows:

- Berkeley *Drosophila* Genome Project, http://www.fruitfly.org/seq_tools/splice.html (for splice-site analysis)
- Breast Cancer Information Core (BIC) database, <http://research.nhgri.nih.gov/bic/>
- Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for *BRCA1* and *BRCA2*)

References

- Abkevich V, Zharkikh A, Deffenbaugh A, Frank D, Chen Y, Shattuck D, Skolnick M, Gutin A, Tavtigian S (2004) Anal-

- ysis of missense variation in human *BRCA1* in the context of interspecific sequence variation. *J Med Genet* 41:492-507
- Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, et al (2003) Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72:1117-1130
- Barker DF, Almeida ER, Casey G, Fain PR, Liao SY, Masunaka I, Noble B, Kurosaki T, Anton-Culver H (1996) *BRCA1* R841W: a strong candidate for a common mutation with moderate phenotype. *Genet Epidemiol* 13:595-604
- Breast Cancer Linkage Consortium (1997) Pathology of familial breast cancer: differences between breast cancers in carriers of *BRCA1* or *BRCA2* mutations and sporadic cases. *Lancet* 349:1505-1510
- Chen CF, Chen PL, Zhong Q, Sharp ZD, Lee WH (1999) Expression of BRC repeats in breast cancer cells disrupts the *BRCA2*-Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. *J Biol Chem* 274:32931-32935
- Cornelis RS, Neuhausen SL, Johansson O, Arason A, Kelsell D, Ponder BA, Tonin Hamann U, et al (1995) High allele loss rates at 17q12-q21 in breast and ovarian tumours from *BRCA1*-linked families. The Breast Cancer Linkage Consortium. *Genes Chromosomes Cancer* 13:203-210
- Davies AA, Masson JY, McIlwraith MJ, Stasiak AZ, Stasiak A (2001) Role of *BRCA2* in control of the RAD51 recombination and DNA repair protein. *Mol Cell* 7:273-282
- Fleming MA, Potter JD, Ramirez CJ, Ostrander GK, Ostrander EA (2003) Understanding missense mutations in the *BRCA1* gene: an evolutionary approach. *Proc Natl Acad Sci USA* 100:1151-1156
- Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpfer KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC (2002) Clinical characteristics of individuals with germline mutations in *BRCA1* and *BRCA2*: analysis of 10,000 individuals. *J Clin Oncol* 20:1480-1490
- Gowen LC, Johnson BL, Latour AM, Sulik KK, Koller BH (1996) *BRCA1* deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat Genet* 12:191-194
- Grantham R (1974) Amino acid difference formula to help explain protein evolution. *Science* 185:862-864
- Hohenstein P, Kielman MF, Breukel C, Bennett LM, Wiseman R, Krimpenfort P, Cornelisse C, van Ommen GJ, Devilee P, Fodde R (2001) A targeted mouse *BRCA1* mutation removing the last BRCT repeat results in apoptosis and embryonic lethality at the headfold stage. *Oncogene* 20:2544-2550
- Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, D'Andrea AD (2002) Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 297:606-609
- Koonin EV, Altschul SF, Bork P (1996) *BRCA1* protein products...functional motifs.... *Nat Genet* 13:266-268
- Lakhani S, Jacquemier J, Sloane JP, Gusterson BA, Anderson TJ, van de Vijver MJ, Farid LM, et al (1998) Multifactorial analysis of differences between sporadic breast cancers and cancers involving *BRCA1* and *BRCA2* mutations. *J Natl Cancer Inst* 90:1138-1145
- Liu CY, Flesken-Nikitin A, Li S, Zeng Y, Lee WH (1996) Inactivation of the mouse *BRCA1* gene leads to failure in the morphogenesis of the egg cylinder in early postimplantation development. *Genes Dev* 10:1835-1843
- Miller MP, Kumar S (2001) Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* 10:2319-2328
- Mirkovic N, Marti-Renom M, Weber B, Sali A, Monteiro A (2004) Structure-based assessment of missense mutations in human *BRCA1*: implications for breast and ovarian cancer predisposition. *Cancer Res* 64:3790-3797
- Petersen GM, Parmigiani G, Thomas D (1998) Missense mutations in disease genes: a Bayesian approach to evaluate causality. *Am J Hum Genet* 62:1516-1524
- Rebbeck TR, Lustbader ED, Buetow KH (1994) Somatic allele loss in genetic linkage analysis of cancer. *Genet Epidemiol* 11:419-429
- Thompson D, Easton DF, Goldgar DE (2003) A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 73:652-655
- Vallou-Christersson J, Cayan C, Haraldsson K, Loman N, Bergthorsson JT, Brondum-Nielsen K, Gerdes AM, Moller P, Kristofferson U, Olsson H, Borg A, Monteiro AN (2001) Functional analysis of *BRCA1* C-terminal missense mutations identified in breast and ovarian cancer families. *Hum Mol Genet* 10:353-360
- Wagner JE, Tolar J, Levran O, Scholl T, Deffenbaugh A, Sa-tagopan J, Ben-Porat L, Mah K, Batish SD, Kutler DI, Mac-Millan ML, Hanenberg H, Auerbach AD (2004) Germline mutations in *BRCA2*: shared genetic susceptibility to breast cancer, early onset leukemia and Fanconi anemia. *Blood* 103:3226-3229
- Wong AK, Pero R, Ormonde PA, Tavtigian SV, Bartel PL (1997) RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *BRCA2*. *J Biol Chem* 272:31941-31944
- Yang H, Jeffrey PD, Miller J, Kinnucan E, Sun Y, Thoma NH, Zheng N, Chen PL, Lee WH, Pavletich NP (2002) *BRCA2* function in DNA binding and recombination from a *BRCA2*-DSS1-ssDNA structure. *Science* 297:1837-1848

Functional Evaluation and Cancer Risk Assessment of BRCA2 Unclassified Variants

Kangjian Wu,¹ Shannon R. Hinson,¹ Akihiro Ohashi,¹ Daniel Farrugia,¹ Patricia Wendt,¹ Sean V. Tavtigian,² Amie Deffenbaugh,³ David Goldgar,² and Fergus J. Couch¹

¹Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, Minnesota; ²Cancer Susceptibility Unit and Genetic Epidemiology Unit, IARC, Lyon, France; and ³Myriad Genetics Laboratories, Inc., Salt Lake City, Utah

Abstract

The influence of germ line *BRCA2* unclassified variants (UCV), including missense mutations and in-frame deletions and insertions on *BRCA2* function and on cancer risk, has not been defined although these mutations account for 43% of all identified *BRCA2* sequence alterations. To investigate the effects of UCVs on *BRCA2* function, we compared mutant and wild-type forms of *BRCA2* using assays of cellular survival and viability, homologous recombination repair, and genome instability. We confirm that the effects of known deleterious mutations can be distinguished from neutral polymorphisms and wild-type *BRCA2* in these assays, and we characterize the influence of a series of UCVs on *BRCA2* function. We also describe how the results from the assays can be combined with data from analysis of cosegregation of the UCVs with cancer, co-occurrence of the UCVs with other deleterious mutations, and interspecies sequence variation in a comprehensive framework in an effort to better distinguish between disease predisposing and neutral UCVs. This combined approach represents a useful means of addressing the functional significance and cancer relevance of UCVs in *BRCA2*. (Cancer Res 2005; 65(2): 417-26)

Introduction

Germ line mutations in the *BRCA2* gene on chromosome 13q12-13 (1) that truncate the *BRCA2* protein are associated with a 60% to 85% lifetime risk of breast cancer, a 15% to 30% lifetime risk of ovarian cancer (2), and predisposition to pancreatic cancer (3). These mutations are readily classified as cancer predisposing because they uniformly truncate *BRCA2* prior to the COOH-terminal nuclear localization signals (4), resulting in exclusion of the mutant proteins from the nucleus and inactivation of all associated nuclear functions. However, 43% of all sequence alterations that have been detected in the *BRCA2* gene during clinical mutation screening, excluding common polymorphisms, do not truncate the encoded protein (5). These unclassified variants (UCV) are predominantly missense mutations and in-frame deletions and have been detected in 13% of women undergoing clinical screening for mutations (5). Indeed, >800 unique *BRCA2* UCVs have been described (Breast Cancer Information Core, <http://research.nhgri.nih.gov/bic/>). To date, it has proved difficult to discriminate the disease predisposing/causing UCVs from the neutral/benign UCVs. This inability to determine

which mutations are disease causing has led to significant problems in risk evaluation, counseling, and preventative care of thousands of carriers of these variants.

A number of methods for discriminating deleterious/high-risk from neutral/low-risk UCVs in *BRCA2* and other genes have been proposed. Analysis of cosegregation of the UCVs with cancer in families (6) is a powerful approach based on the cancer phenotype. Approaches based on sequence conservation and the physicochemical properties of amino acid changes (7, 8) have also been used. Similarly, functional studies that measure the influence of mutations on the wild-type activity of a protein, such as ATM or BRCA1, in assays based on the known functions of the protein have been used with some success (9-11). Whereas all of these approaches have limitations, the combination of these approaches is likely to provide some insight into which missense mutations are cancer causing (6).

Importantly, little has been done in terms of functional analysis of *BRCA2* missense mutations. Although the function of *BRCA2* is not fully defined at the biochemical level, there is strong support for a role in DNA damage repair and in maintenance of genomic integrity (12). Specifically, *BRCA2* binds to the Rad51 DNA recombination enzyme and regulates formation of the rad51 nucleoprotein filament that is required for homologous recombination repair of DNA damage (13, 14). Indeed, *BRCA2* null cells show a reduced efficiency of homologous recombination mediated double-strand break repair (15), and *brca2* null or mutant cells exhibit hypersensitivity to DNA damage (15-24). *BRCA2* also seems to have a distinct role in mediation of chromosomal integrity, as evidenced by its association with BRAF35 chromatin complexes during chromosome condensation (25), and the observation that *brca2* null or mutant cells display a high frequency of broken and deformed chromosomes, micro-nuclei, and centrosome amplification (18, 20, 23, 24, 26).

Here we describe for the first time, the evaluation of the effects of *BRCA2* UCVs on *BRCA2* function using a series of *in vitro* assays based on these known activities of *BRCA2*. We show that the assays can clearly discriminate between known deleterious mutations and neutral variants/polymorphisms. In addition, we report on the evaluation of a number of UCVs and the finding that certain UCVs alter *BRCA2* function. Furthermore, we combine the results from these assays with data from a disease causality prediction algorithm for the UCVs that includes genetic segregation data from high-risk breast cancer families, observations of co-occurrence of UCVs with other deleterious mutations, and interspecies sequence variation (6), to confirm that the results from the functional analysis of *BRCA2* UCVs are consistent with this "genetic" data.

Materials and Methods

Mutation Frequency. DNA from 476 controls with no personal or family history of cancer were genotyped for each of 11 *BRCA2* mutations by dHPLC analysis. Briefly, exons containing the relevant mutations were

Note: Supplementary data for this article are available at *Cancer Research* Online (<http://cancerres.aacrjournals.org/>).

Requests for reprints: Fergus J. Couch, Mayo Clinic College of Medicine, 200 First Street Southwest, Rochester, MN 55905. Phone: 507-284-3623; Fax: 507-266-0824; E-mail: couch.fergus@mayo.edu.

©2005 American Association for Cancer Research.

PCR amplified using intronic primers, products were denatured and heteroduplexed, and evaluated for the presence of the mutations by dHPLC using product-specific melting and solvent conditions. Details are available from the authors.

Site-Directed Mutagenesis. Nucleotide changes were incorporated into five partial *BRCA2* cDNA pCR3.1 and pcDNA3.1 subclones, defined by unique restriction enzyme sites in the *BRCA2* cDNA, using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). Primer sequences and conditions are available from the authors. Mutant partial cDNA fragments were subcloned into the green fluorescent protein (GFP)-tagged full-length *BRCA2* cDNA in the EGFP-C2 plasmid and the untagged *BRCA2* cDNA in pCR3.1. The presence of the mutations was confirmed by sequencing and the stability of the constructs was established by a series of restriction enzyme digestions (data not shown).

Cell Culture, Transfection, and Stable Cells. *BRCA2*-deficient VC8 cells (23) were maintained in DMEM-F12 (BioWhittaker, Walkersville, MD) supplemented with 10% bovine calf serum (HyClone, Logan, UT), 2 mmol/L L-glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin. 293T and 293 cells were obtained from the American Type Culture Collection (Rockville, MD) and cultured in DMEM (Life Technologies, Gaithersburg, MD) with 10% bovine calf serum. Transfections were done using FuGENE-6 reagents according to the manufacturer's protocol (Roche, Indianapolis, IN). VC8 *BRCA2* stable cells were selected with G418 (400 µg/mL).

Immunoprecipitation and Immunoblotting. Cell lysates were prepared, *BRCA2* protein was immunoprecipitated with anti-*BRCA2* Ab-1 antibody (Oncogene Research, Boston, MA), and *BRCA2* protein was immunoblotted using the anti-*BRCA2* Ab-2 antibody (Oncogene Research) as described previously (27).

Immunofluorescence and Confocal Microscopy. Cells were cultured on glass coverslips and transfected with GFP-*BRCA2* constructs. For subcellular localization experiments, the cells were fixed with cold methanol, stained with 1 µg/mL Hoechst (Molecular Probes, Eugene, OR), and imaged 48 hours post-transfection. For indirect immunofluorescence, cells were fixed, permeabilized, and stained with anti-pericentrin (1:800) antibodies, Texas red goat anti-rabbit antibodies (1:800), and Hoechst 4 days after transfection (28). Centrosome numbers were counted in ~300 total cells from two independent experiments. All images were acquired with a Zeiss LSM510 confocal microscope.

Clonogenic Survival and Trypan Blue Viability Assays. Forty hours after transfection with GFP-tagged constructs, VC8 cells were flow sorted for GFP-positive cells. For clonogenic survival assay, 2,000 GFP-positive cells or stable VC8 cells were seeded into 10-cm dishes, incubated overnight, and treated with 1, 1.75, or 2.5 ng/mL mitomycin C (MMC) for 72 hours. In parallel, 500 cells were incubated in the absence of MMC. Cells were rinsed with fresh medium incubated for 4 days, stained with 0.2% Coomassie brilliant blue-R250 solution and visible colonies were enumerated. Stable cell lines were also irradiated with γ -irradiation (0, 2, 4, or 6 Gy) using a ³⁵Cs source, incubated, and stained as before. For trypan blue exclusion assays, 4,000 GFP-positive cells or stable VC8 cells were plated, treated with MMC (0, 1, 2.5, or 5 ng/mL) for 72 hours, harvested, resuspended in 50% trypan blue solution, and the living cells were enumerated using a hemocytometer. Survival was calculated by comparing the number of surviving colonies or cells from MMC or radiation treated and untreated cells.

Homologous Recombination Assay. Stable incorporation of DR-GFP (15) in VC8 cells was established by selection with 2.5 µg/mL puromycin. Southern blotting of genomic DNA from clonal isolates with a DR-GFP probe was done and a low copy number clone (VC8-DR-GFP) was selected for use in the assay. VC8-DR-GFP cells were transfected with either PCR3.1 vector, pcBasce1 vector containing the *SceI* restriction endonuclease gene, or pcBasce1 plus various *BRCA2*/PCR3.1 constructs. The percentage of cells that were GFP positive was quantitated by flow cytometric analysis 5 days after transfection.

Multiple Sequence Alignment. A multiple sequence alignment of full-length *BRCA2* cDNA sequences was generated using the T-coffee program (29). The alignment was restricted to *BRCA2* cDNA sequences derived from vertebrates in an effort to avoid using sequences that are evolutionarily

related but are not direct homologues of the *BRCA2* gene. The Tetraodon *BRCA2* sequence was obtained from unpublished sources,⁴ and other *BRCA2* protein sequences were obtained from Genbank.

Segregation Analysis. For the analysis of cosegregation, we assumed an allele frequency of the variant of 0.0001 and a penetrance model with separate age-specific risks of breast and ovarian cancer for *BRCA2* based on the meta-analysis estimates of Antoniou et al. (30), with pooling across age groups (6). For *BRCA2* variants, we assumed relative risks in carriers relative to noncarriers of 11.5 for breast cancer, independent of age, whereas for ovarian cancer these relative risks were 4.8 for ovarian cancer below age 50 and 13.1 for ovarian cancer at ages ≥ 50 years. The corresponding cumulative risk of breast or ovarian cancer was 51% by age 70. In cases where full pedigrees were unavailable, we relied on family history information and reconstructed pedigrees by creating individuals of unknown phenotype and genotype to connect the individuals in the pedigree. These assumptions were applied to the model of Thompson et al. (31) for estimation of the likelihood of disease causality associated with *BRCA2* UCVs, as recently outlined (6).

Results

Frequency in Control Populations. The goal of this study was to show how functional assays can discriminate between disease-causing and neutral UCVs in the *BRCA2* gene and to outline how data from functional and genetic studies can be combined to predict the disease causality of *BRCA2* UCVs. For the purposes of the study, we selected 11 *BRCA2* mutations (Table 1). Three encoded amino acid alterations in the NH₂-terminal region (Y42C, N372H, and E462G), two were in the BRC repeats (T1302del and E1382del), and four were in the COOH-terminal DNA binding domain (T2515I, R2659K, D2723H, and V2908G). The *BRCA2* truncating 6174delT Ashkenazi Jewish founder mutation associated with a breast cancer risk of 70% by age 70 (32) was also included as a positive/inactivating control, whereas the K3326X truncating variant that does not segregate with disease (33) was used as a negative/wild-type control.

Initially, the frequency of all 11 mutations in normal controls was assessed to determine if any of the mutations were frequently observed in normal controls and were therefore likely to be neutral/low-risk mutations. Germ line DNAs from 476 healthy Caucasians of varying ages (18-65 years) undergoing routine colonoscopy at the Mayo Clinic were evaluated for the presence of the mutations using dHPLC analysis. All UCVs were at a frequency of 0 in 476 except for N372H and K3326X, which were found in 183 and 4 individuals, respectively. As only these two mutations detected in the control population can be excluded as inactivating and deleterious, the majority of the UCVs under study here remain as candidate deleterious mutations based on these data.

In vitro Mutagenesis of *BRCA2* cDNA. In order to conduct *in vitro* assays to distinguish missense mutations and in-frame point deletions/insertions (IFD/I) that alter *BRCA2* function from those with no effect, full-length *BRCA2* cDNA expression constructs encoding the wild-type and the 11 mutant forms of the protein were generated by site directed mutagenesis (Table 1; Fig. 1A). Of note, the R2659K UCV has recently been shown to cause an in-frame 171-bp deletion of exon 17 of *BRCA2* (34). In the studies described here, we have used a *BRCA2* expression construct containing a 171-bp in-frame deletion to correctly reflect this exon skipping.

⁴ S.V. Tavtigian, unpublished data.

Table 1. Sequence conservation and GMS for BRCA2 UCVs

Nucleotide variant	Amino acid variant	Exon	GMS UCV	Rat	Mouse	Chick	Cat	Dog	Puffer fish	Alignment score	Odds alignment	Odds GMS
353A > G	Y42C	3	104	Y	Y	Y	Y	Y	Δ	—	0.19	3.49
1342C > A	N372H	10	68	H	S	G	N	H	Δ	—	—	—
1613A > G	E462G	10	98	E	E	K	E	E	Δ	—	0.01	1.26
4132delACT	T1302del	11	—	T	T	H	T	T	Δ	—	0.01	—
4382delAGA	E1382del	11	—	E	E	E	E	E	E	0	5.10	—
6174delT	6174delT	11	—	—	—	—	—	—	—	—	—	—
7772C > T	T2515I	15	89	S	S	T	T	T	S	58	0.01	0.91
8204G > A	R2659K	17	26	R	R	R	R	R	R	0	5.10	0.33
8395G > C	D2723H	18	81	D	D	D	D	D	D	0	5.00	0.98
8951T > G	V2908G	21	109	V	V	I	V	V	V	29	0.19	1.46
10204A > T	K3326X	27	—	F	F	R	L	R	R	—	—	—

NOTE: Δ, deleted; —, not applicable. Rat *brca2* cDNA (AB107955.1), mouse *brca2* cDNA (MMU82270), chicken *brca2* cDNA (AB066374.1 and AY083934.1), cat *brca2* cDNA (AB107955.1), and dog *brca2* cDNA (AB043895.2).

BRCA2 Expression and Localization. Ectopic expression of the GFP-tagged BRCA2 proteins in 293T cells after transient transfection was verified by immunoblotting with anti-BRCA2 antibodies (Fig. 1B). Immunoblotting of nontagged versions of the BRCA2 proteins detected similar levels of protein, suggesting that the GFP tag did not have a significant effect on protein stability. Protein stability analysis in the presence of cycloheximide also showed that the mutations had no effect on the half-life of the BRCA2 protein (data not shown).

Subcellular localization of the GFP-tagged BRCA2 proteins was evaluated by fluorescence confocal microscopy. The localization of BRCA2 and 6174delT to the nucleus and cytoplasm (Table 2; Fig. 1C), respectively, was consistent with the C-terminal location of the human BRCA2 nuclear localization signals (4). As expected, the Y42C, N372H, E462G, T1302del, E1382del, V2908G, and K3326X BRCA2 variants localized to the nucleus in at least 95% of transfected cells. However, the D2723H and R2659K BRCA2 mutants were localized predominantly in the cytoplasm in >90% of transfected cells (Table 2; Fig. 1C). Interestingly, the T2515I mutant was found in the nucleus (58%), the cytoplasm (26%), and both (16%). The influence of this partial relocalization of the protein on function is unclear. Mislocalization of these mutant proteins relative to the other mutants and wild-type BRCA2 was confirmed in HeLa cells (data not shown). As these mutations do not directly disrupt the known nuclear localization signal motifs (4), the results suggest an effect either on previously undetected sites that mediate BRCA2 localization or an effect on BRCA2 protein folding.

MMC Hypersensitivity. The BRCA2 deficient VC8 (23) cell line is extremely sensitive to cross-linking agents such as MMC. Stable reintroduction of the *BRCA2* gene by chromosomal transfer or by bacterial artificial chromosome significantly reduces this MMC sensitivity (23), suggesting that it is possible to discriminate between wild-type BRCA2 and inactivated mutant BRCA2 based on the MMC sensitivity of VC8 cells. To begin to evaluate this possibility, we generated clonal isolates of VC8 cells stably transfected with wild-type *BRCA2* and confirmed BRCA2 expression by immunoblotting of immunoprecipitated BRCA2 (Fig. 2A). The integrity of wild-type BRCA2 function in these cells was established by demonstrating that reconstituted BRCA2 facilitated

the formation of Rad51 nuclear foci in response to γ -irradiation (Supplementary Fig. 1). Subsequently, we used clonogenic survival assays to show that constitutive expression of wild-type BRCA2 in VC8 cells reduced MMC sensitivity (Fig. 2B). In addition, we showed that wild-type BRCA2 expressing VC8 cells were less sensitive than vector transfected cells to γ -irradiation (Fig. 2B). In light of these observations, we also evaluated the effects of BRCA2 UCVs on MMC sensitivity. Stable expression of E462G, D2723H, and T2515I BRCA2 in VC8 cells was established and verified by immunoprecipitation/Western blot (Fig. 2A), and the cells were used in MMC-dependent clonogenic survival assays. The E462G variant suppressed hypersensitivity to MMC and γ -irradiation similarly to wild-type BRCA2, whereas the T2515I and D2723H variants seemed to inactivate BRCA2 and had no ability to reduce the cellular sensitivity to these agents (Fig. 2B).

As we were unable to generate stable cell lines constitutively expressing other BRCA2 UCVs, we developed a strategy for evaluation of MMC and radiation sensitivity of VC8 cells based on transient transfection of VC8 cells with wild-type and mutant forms of BRCA2. VC8 cells transiently transfected with GFP vector, GFP-tagged wild type, and GFP-tagged 6174delT human *BRCA2* expression constructs were flow sorted for GFP-positive cells and BRCA2 expression was verified by immunoprecipitation/Western blot. The selected cells were evaluated for MMC sensitivity using clonogenic survival assays and cell viability assays. Wild-type BRCA2 significantly decreased cellular sensitivity to MMC relative to vector or the 6174delT truncation mutant (Fig. 2C). Similarly, wild-type BRCA2 enhanced cell viability relative to controls in response to MMC (Fig. 2C). Subsequent analysis of flow-sorted VC8 cells expressing GFP-tagged Y42C, N372H, E462G, V2908G, and K3326X showed 5-fold enhanced survival (Fig. 2D) and 12- to 14-fold improved viability (Fig. 2D) in response to MMC, similarly to wild-type BRCA2 expressing cells. In contrast, GFP-tagged T2515I, R2659K, D2723H, T1302del, and E1382del mutants did not reduce sensitivity to MMC. As these mutants all expressed equivalently to wild type, the suggestion is that these UCVs inactivate BRCA2 function and that transiently transfected and flow sorted GFP-BRCA2 expressing cells can be used in MMC sensitivity assays to evaluate the influence of UCVs on BRCA2 function.

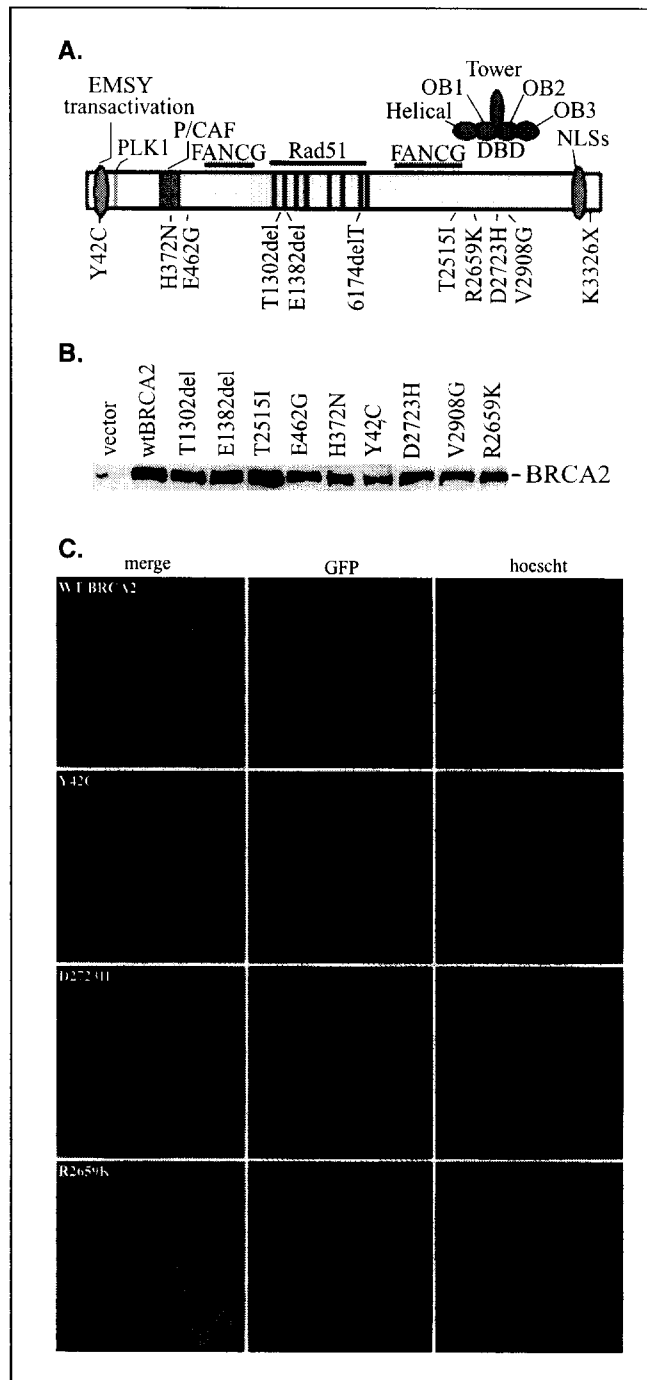


Figure 1. Expression of BRCA2 wild-type and mutant proteins. *A*, schematic of the BRCA2 protein: location of specific motifs and the BRCA2 UCVs. *B*, immunoblots of ectopically expressed wild-type (*wtBRCA2*) and mutant BRCA2 protein from 293T cells. *C*, subcellular localization of ectopically expressed wild-type and mutant BRCA2 protein in 293T cells. GFP-tagged wild-type and mutant protein was visualized by confocal microscopy.

Homology-Directed Repair. Direct evidence for a role of BRCA2 in promoting homology-directed repair (HDR) in BRCA2 mutant Capan-1 cells and mouse embryo fibroblasts was recently obtained using an I-SceI-dependent DR-GFP reporter assay (15, 35). In this reporter, a *GFP* gene is inactivated by the introduction of an I-SceI recognition site, whereas an adjacent *GFP* gene is differentially

mutated. After the introduction of a DNA double-strand break at the I-SceI site, the *GFP* gene can be reconstituted by HDR using the downstream-inactivated *GFP* gene as a template. To evaluate the influence of UCVs on HDR using this system, the DR-GFP reporter construct was first stably integrated into VC8 cells and the presence of the construct at low copy number in the genome was verified by Southern blotting. These DR-GFP VC8 cells were not GFP positive in the absence of I-SceI indicating that spontaneous gene conversion in DR-GFP-VC8 cells was rare (Fig. 3A). Similarly, spontaneous recombination in the presence of I-SceI was only detected in 0.35% of cells (Fig. 3A). In contrast, coexpression of I-SceI and non-GFP tagged wild-type *BRCA2* resulted in 8-fold more GFP-positive cells (Fig. 3A and B), indicating that wild-type BRCA2 can promote recombinational repair of SceI-mediated DNA double-strand breaks. This approach has also been successfully applied to the evaluation of the role of the 53BP1 protein in HDR of double-strand breaks (36). Subsequent evaluation of the effects of BRCA2 mutants on the ability of BRCA2 to mediate recombinational DNA repair revealed that the Y42C, N372H, E462G, T1302del, and K3326X mutations all functioned similarly to wild-type BRCA2 (Fig. 3B). In contrast, levels of HDR equivalent to those obtained with vector alone were associated with expression of the E1382del, R2659K, D2723H, and 6174delT mutants, whereas T2515I showed a slight reduction in HDR relative to wild-type (Fig. 3B). In parallel, confocal microscopy and semiquantitative reverse transcription-PCR were used to verify that all constructs had similar transfection efficiencies and levels of expression in the DR-GFP-VC8 cells (Fig. 3C). Taken together, the data indicate that transient transfection of *BRCA2* expression constructs into DR-GFP-VC8 cells can be used to evaluate the effects of BRCA2 UCVs on the HDR activity of BRCA2.

Induction of Centrosome Amplification. BRCA2 mutant mouse embryo fibroblasts, Capan-1, and FANCD1 cells, and BRCA2-deficient VC8 cells all display extensive centrosome amplification (23, 24, 26), suggesting that the chromosomal instability observed in BRCA2 mutant cells (23, 26) either in part results from aberrant centrosome function and abnormal numbers of centrosomes, or causes centrosome amplification. Alternatively, BRCA2-associated centrosome amplification may result from uncoupling of DNA replication and centrosome duplication in S phase. In either case, centrosome amplification is a marker of cell cycle disruption and chromosomal instability and can be used as a measure of BRCA2 function. Indeed, Kraakman-van der Zwet et al. (23) showed that reconstitution of wild-type *BRCA2* through chromosomal transfer or bacterial artificial chromosome clones in VC8 cells caused a significant reduction in the proportion of cells with centrosome amplification and chromosomal instability.

To investigate the influence of BRCA2 UCVs on BRCA2 function using this approach, we quantified centrosome amplification in VC8 *BRCA2* wild-type and UCV stable cell lines. Cells containing a vector control or D2723H BRCA2 did not have a reduced frequency of centrosome amplification, whereas the proportion of wild-type BRCA2 and the E462G UCV expressing cells with amplification was reduced from 80% to 40% to 50% (Fig. 4A). These findings indicate that plasmid based expression of BRCA2 can rescue centrosome amplification and that centrosome amplification can be used to evaluate the influence of UCVs on BRCA2 function. However, because of the difficulty in establishing VC8 BRCA2 stable cell lines, we also developed a transient transfection-based approach. We used 293T cells because they are readily transfected by *BRCA2* expression constructs, have a normal complement of centrosomes, and display centrosome amplification in response to mutant forms

of BRCA2 (Fig. 4B and C). Indeed, only 5% to 6% of 293T cells ectopically expressing GFP-tagged wild-type BRCA2 or GFP vector have centrosome amplification (Fig. 4B), whereas 18% to 20% of cells expressing the 6174delT truncation mutant display centrosome amplification. Similarly, 17% to 20% of 293T cells ectopically expressing GFP-tagged T1302del, R2659K, and D2723H contained amplified centrosomes, whereas expression of several other UCVs had no effect on centrosome number (Fig. 4B). Similar results were obtained in 293 and HeLa cells (data not shown), suggesting that mutant forms of BRCA2 can induce centrosome amplification and that the T1302del, R2659K, and D2723H UCVs all inactivate BRCA2 function. It should be noted that these effects are likely the result of competitive dominant negative effects of the overexpressed mutants. However, the results are consistent with those from stable cell lines and seem to reflect the role of BRCA2 in regulation of cytokinesis (37) and centrosome number (data not shown).

Cross-Species Multiple Sequence Alignment. Interspecies multiple protein sequence alignments provide an evolutionary perspective to identification of important functional motifs and amino acids within proteins (38, 39). In addition, the Grantham chemical difference matrix (40) quantifies physicochemical differences between mutant and wild-type amino acids in humans and can be used to correlate alterations with changes in phenotype and function, as recently shown for BRCA1 (7, 8, 40). Here, we use multiple sequence alignment and Grantham matrix scores (GMS) to predict whether the BRCA2 UCVs are important for BRCA2 function, as previously described (6). As shown in Table 1, the E1382, R2659, and D2723 residues are fully conserved across all species including the relatively distantly related pufferfish, *Tetraodon nigriviridis*. This strong conservation results in odds of ~5:1 in favor of the UCVs being deleterious (Table 1) and suggests that these UCVs are important for BRCA2 function, although the GMSs for the associated UCVs in humans were not high. The Y42, N372, E462, T1302, T2515, and V2908 residues were divergent across species and even deleted from the tetraodon BRCA2

sequence and are likely not important for function (Table 1). However, it remains possible that BRCA2 has acquired functions during evolution and that the NH₂-terminal residues that are not present in tetraodon but are present in mammals are important for function. Overall, the odds of causality based on alignment and the GMS (Table 1) are insufficient for classification of the disease causality of the UCVs, but they do provide support for the outcomes of the functional assays.

Cosegregation Analysis. To further evaluate the ability of the functional assays to discriminate between deleterious and neutral UCVs, we compared the results from the functional assays with the predicted disease causality of the UCVs. Disease causality of the UCVs was defined using a recently developed likelihood model that establishes the odds of disease causality for mutations in BRCA1 and BRCA2 based on available family data (6, 31). Information on segregation of the UCVs in families was provided by Myriad Genetics Laboratories, Inc. (Salt Lake City, UT) and the odds in favor or against disease causality for each UCV were determined (Table 2). Segregation analyses for the N372H and K3326X mutations were not done because both mutations have been detected many times in combination with a large variety of truncating BRCA2 mutations and it is clear that these variants are associated with little or no risk of cancer. We assign odds of $<1 \times 10^{-3}$ for disease causality for these variants in Table 2 to reflect the known neutrality of the mutations. In contrast, segregation was evaluated in 10 of the 24 known families that carry the D2723H UCV. All pedigrees showed complete cosegregation with breast and ovarian cancer, yielding overall odds of 13,723:1 in favor of disease causality (6). Similarly, T2515I and R2659K were evaluated in 9 and 3 pedigrees, respectively. Information was limiting and the odds in favor of disease causality were low. Similarly, one pedigree with the E1382del variant generated odds in favor of causality of 1.65:1. In contrast, 17 pedigrees containing the Y42C variant and 20 containing the E462G variant yielded highly significant odds against causality (Table 2). Whereas it was only possible to classify

Table 2. Results from functional assays and likelihood of causality analysis of BRCA2 UCVs

Amino acid change	Cellular localization	Survival	HDR	Centrosome	Pedigrees*	Co-segregation †	Co-occurrence ‡	Combined odds of causality	Disease relevance
Y42C	N	+	+	+	17	6.70×10^{-7}	2.5×10^{-11}	1.1×10^{-17}	Low/neutral
N372H	N	+	+	+	>1,000	$<1 \times 10^{-3}$	$<1 \times 10^{-3}$	—	Low /neutral
E462G	N	+	+	+	20	3.60×10^{-5}	0.08	2.88×10^{-8}	Low /neutral
T1302del	N	—	+	—	0	ND	ND	—	Undefined
E1382del	N	—	—	+	1	1.65	1.27	10.69	Undefined
T2515I	N/C	—	+/-	+/-	9	1.09	16.63	0.16	Low /neutral
R2659K	C	—	—	—	3	2.66	2.05	9.18	Deleterious
D2723H	C	—	—	—	10	1.37×10^4	4.2	2.8×10^5	Deleterious
V2908G	N	+	+	+	1	ND	1.2	0.33	Low /neutral
K3326X	N	+	+	+	>300	$<1 \times 10^{-3}$	$<1 \times 10^{-3}$	—	Low /neutral
6174delT	C	—	—	—	>500	$>1 \times 10^3$	$>1 \times 10^3$	—	Deleterious

NOTE: Abbreviations: HDR, homology directed repair; ND, not done; N, nuclear; C, cytoplasmic.

+ , similar to wild type; - , similar to 6174delT mutant.

*No. pedigrees available for analysis.

†Likelihood of disease causality based on cosegregation with cancer.

‡Likelihood of disease causality based on co-occurrence with a deleterious mutation.

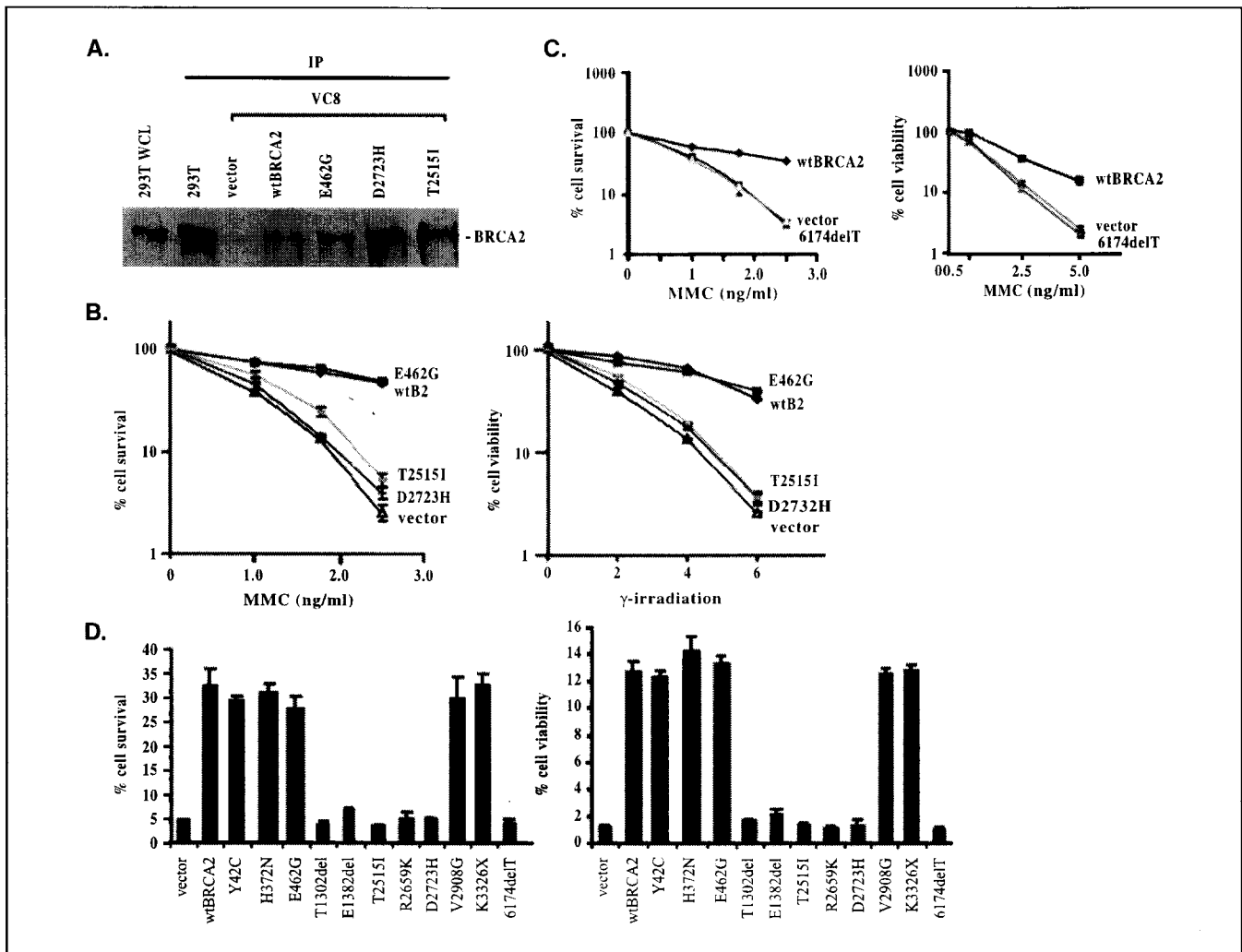


Figure 2. Comparison of the effects of wild-type and mutant BRCA2 on MMC response in VC8. *A*, immunoblots of BRCA2 protein from VC8 cells stably expressing wild-type (*wtBRCA2*) and mutant (E462G, D2723H, and T25151) forms of BRCA2. Protein was immunoprecipitated and compared by immunoblot with endogenous BRCA2 from 293T cells. Six-fold less protein was used in the 293T IP. *B*, clonogenic survival of VC8 cells stably expressing wild-type and mutant BRCA2 in response to MMC and γ -irradiation. *C*, ectopic expression of *wtBRCA2* enhances cell survival relative to vector and truncated BRCA2. VC8 cells transiently transfected with GFP-tagged *wtBRCA2* and 6174delT BRCA2 and flow sorted for GFP were evaluated for MMC sensitivity by clonogenic survival assay and trypan blue exclusion. % surviving colonies and % viable cells in treated relative to untreated cells was plotted against MMC concentration. *D*, BRCA2 UCVs influence VC8 MMC sensitivity. MMC sensitivity of flow-sorted VC8 cells expressing BRCA2 UCV mutants was evaluated by clonogenic survival assay and by trypan blue exclusion viability assay.

Y42C, N372H, E462G, D2723H, and K3326X by this method, it is important to note that the results matched the outcome from the functional assays (Table 2).

Co-occurrence with Deleterious Mutations. Given that biallelic inactivation of BRCA2 is associated with either embryonic lethality or Fanconi anemia, it is unlikely that UCVs that inactivate BRCA2 function can be found in *trans* with other truncating BRCA2 mutations in non-Fanconi anemia adults. Similarly, two inactivating *BRCA2* mutations have never been found on the same allele. This suggests that BRCA2 UCVs that co-occur with known deleterious BRCA2 mutations are unlikely to inactivate BRCA2 function or confer a high risk of cancer. By applying co-occurrence data from 20,000 individuals tested at Myriad Genetics Laboratories to a likelihood model that predicts disease causality based on co-occurrence in families (6) we evaluated the disease causality of the UCVs described above (Table 2). The N372H and K3326X variants co-occurred with

many truncating mutations suggesting that these variants do not alter BRCA2 function sufficiently to predispose to cancer. Likewise, the Y42C UCV co-occurred with eight different BRCA2 truncating mutations and can be excluded as disease causing, whereas E462G UCV co-occurred with a Y3098X truncating mutation, and is likely not a deleterious mutation. In contrast, the T25151, D2723H, and V2908G missense mutations and the T1302del, E1382del, and R2659K IFDs did not co-occur with other deleterious mutations (Table 2). However, the absence of co-occurring mutations does not verify that the UCVs are disease causing because the limited number of families with these UCVs restricts the possibility of observing co-occurrences. Whereas this model is most useful for frequently observed UCVs, the odds from the model can be combined with those from cosegregation, sequence conservation, and Grantham matrix analysis to give an overall likelihood of disease causality (Table 2; ref. 6); thus, it remains useful to conduct these analyses. Using odds of 1,000:1

in favor and 100:1 against as thresholds for disease causality, we noted that the functional assay data mirrors the results from the individual and combined likelihood analyses.

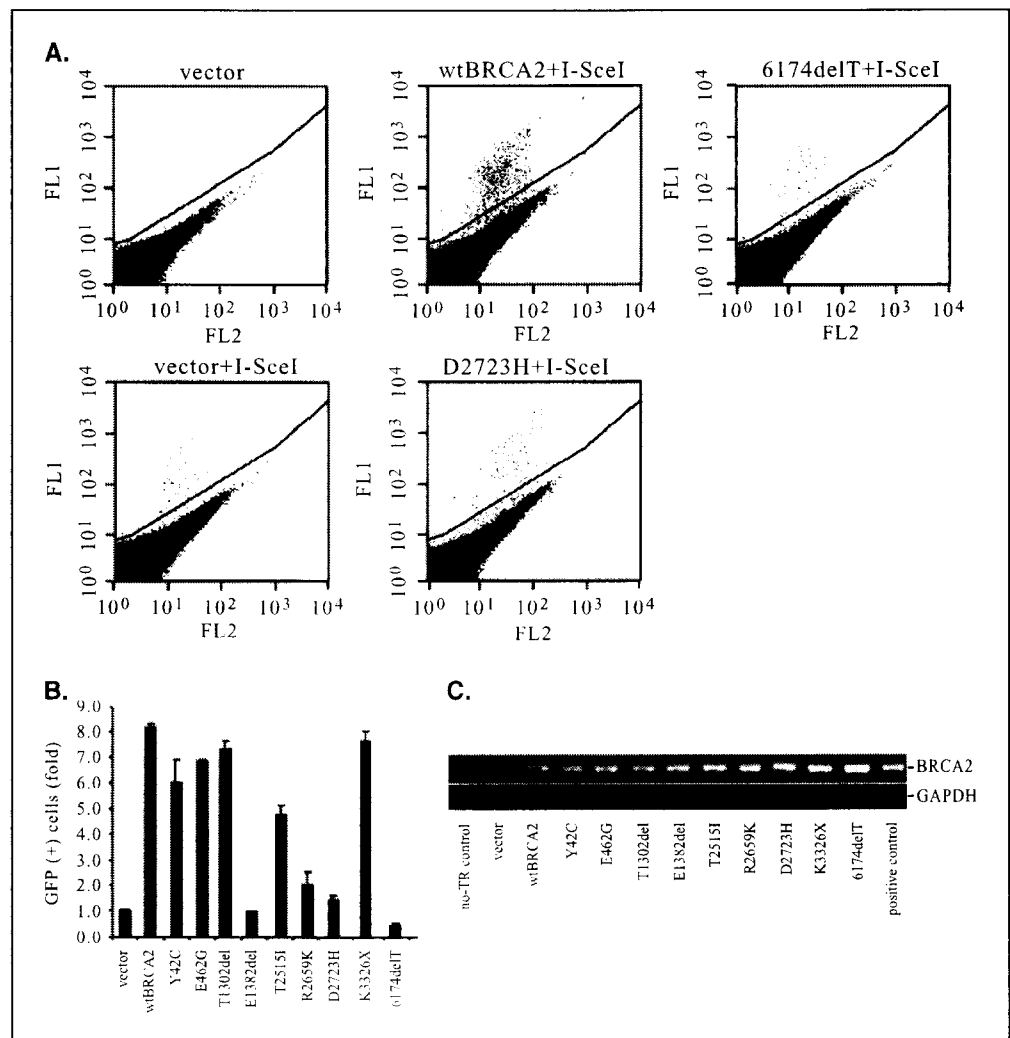
Discussion

Although many missense mutations and in-frame insertions and deletions in *BRCA2*, collectively termed unclassified variants, have been identified in high-risk breast and ovarian cancer patients (affected with breast or ovarian cancer and with a family history of breast or ovarian cancer), it is not known if any of these alter *BRCA2* function sufficiently to predispose carriers of these mutations to cancer. Thus, the carriers of these mutations receive limited cancer risk assessment. A number of approaches involving analysis of high-risk breast cancer families carrying these mutations have been used to address the relevance of these mutations to cancer. These include analysis of cosegregation of variants with cancer in families and evaluation of the co-occurrence of variants with other known deleterious mutations (6, 41). However, these approaches can be limited by the frequency of the variant and the availability of family data. Thus, several lines of investigation must be

considered to classify *BRCA2* variants as deleterious/high risk or neutral/low risk (6, 41, 42).

In this study, we describe the analysis of the effects of several UCVs on *BRCA2* function using *in vitro* assays that evaluate *BRCA2* subcellular localization, MMC sensitivity, homologous recombinational DNA repair, and centrosome amplification. We show that each assay can discriminate between the *BRCA2* 6174delT truncation mutant and wild-type *BRCA2*, and that a series of *BRCA2* UCVs can be separated into groups that inactivate or have no effect on the cellular response to *BRCA2*. The results from the functional studies were generally consistent and led to the prediction that D2723H and R2659K are inactivating and disease causing mutations. Whereas these mutations inactivated the cell survival, homologous recombination repair, and centrosome regulatory functions of *BRCA2*, their ability to inactivate these putatively independent functions of the protein in unison was unexpected, as each mutation alters only a single residue in *BRCA2*. However, the finding that both D2723H and R2659K are excluded from the cell nucleus and cannot perform the predominantly nuclear functions of *BRCA2* provides a likely explanation for these effects. The mechanism by which these mutant forms of *BRCA2* are restricted to the cytoplasm is

Figure 3. HDR analysis of *BRCA2* UCVs. **A**, fluorescence-activated cell sorting analysis to detect GFP-positive VC8 cells following repair of I-SceI-induced DNA breaks. GFP expressing cells are located above the diagonal line. **B**, quantitation of HDR in VC8 cells in response to transiently transfected wild-type and mutant forms of *BRCA2*. The proportion of GFP-positive cells for each construct relative to vector control is shown. Columns, means; bars, \pm SE. **C**, reverse transcription-PCR analysis of *BRCA2* and *GAPDH* in VC8 cells transiently transfected with *BRCA2* constructs confirms *BRCA2* expression. Equal amounts of RNA from transfected cells were used for semiquantitative reverse transcription-PCR analysis.



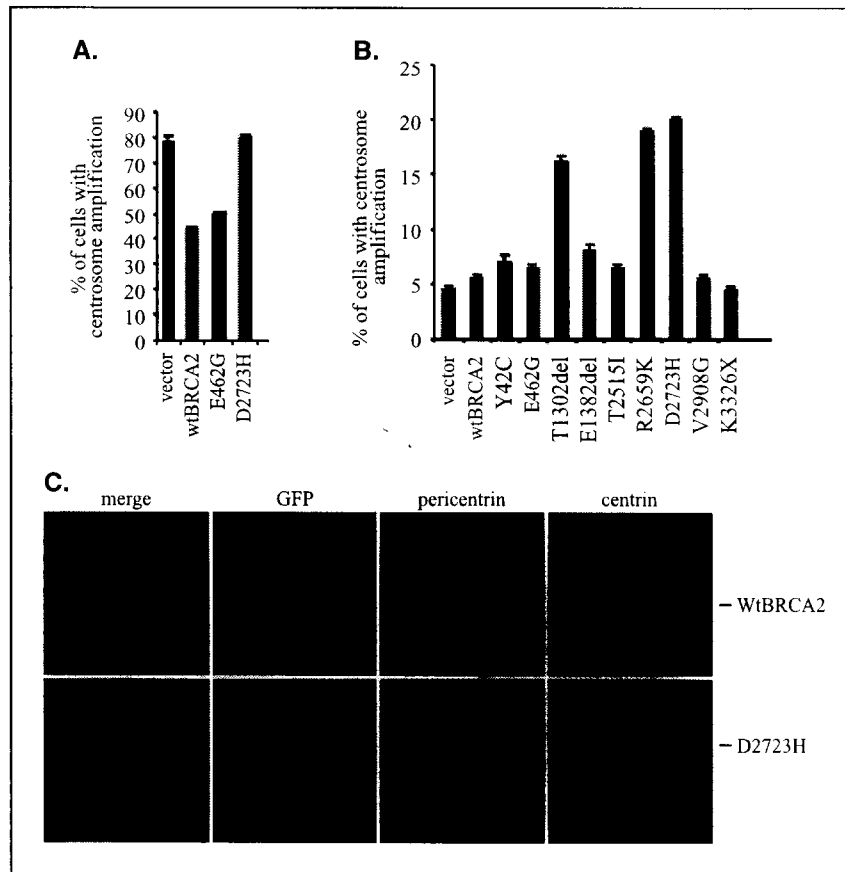


Figure 4. Centrosome amplification is induced by BRCA2 UCVs. **A**, centrosome amplification in VC8 cells is reduced by BRCA2. % VC8 cells stably expressing wild-type and mutant BRCA2 that had >2 pericentrin signals. **B**, quantitation of centrosome amplification in 293T cells ectopically expressing GFP-tagged wild-type and mutant BRCA2. % GFP-positive cells with >2 pericentrin signals. *Columns, means; bars, ± SE.* **C**, immunofluorescence of centrin and pericentrin in the centrosomes of 293T cells ectopically expressing wild-type and D2723H mutant BRCA2. Centrosome amplification is reflected by >4 centrioles (*centrin*) or >2 centrosomes (*pericentrin*).

unknown. However, the nuclear exclusion of these proteins and the associated inactivation of all measured BRCA2 functions suggests that evaluation of the cellular localization of ectopically expressed UCV mutants of BRCA2 may be useful as a prescreen for identifying inactivating UCVs. We also identified a number of UCVs that influenced a subset of BRCA2 functions or only partially inactivated BRCA2 function. The T2515I UCV localized to both the cytoplasm and the nucleus and seemed to partially inactivate the homologous recombination and centrosome regulatory functions of BRCA2 and completely ablated the cell survival activity of BRCA2. The partial inactivation of these functions by T2515I may be in response to the lower levels of BRCA2 in the nucleus resulting from mislocalization of a significant amount of BRCA2 to the cytoplasm, but clearly suggests that T2515I has a relatively subtle effect on BRCA2 function. Likewise, the E1382del variant located in the BRC repeats of BRCA2 inactivated BRCA2 in the cell survival and HDR assays suggesting that the UCV specifically alters Rad51 binding by BRCA2 resulting in loss of DNA repair activity. However, it would seem that E1382del does not confer a conformational change on BRCA2 that leads to inactivation of all BRCA2 functions because the UCV had no influence on centrosome number. This supports the notion that BRCA2 is a multifunctional protein that is not only involved in DNA repair but may also contribute to cell cycle and/or mitotic regulation in an as yet undefined manner. Similarly, the T1302del in-frame deletion influences specific functions of BRCA2. In this case, only cell survival, as previously reported (43), and centrosome amplification effects were observed, whereas the variant had no effect on HDR. This suggests that the influence of BRCA2 on

survival and viability in response to MMC damage is independent of its homologous recombination activity. In contrast, the Y42C, E462G, and V2908G UCVs and the frequently observed N372H and K3326X variants had no effect on BRCA2 function in any assay. Importantly, when considering all of the data, it is clear that homologous recombination, cell survival, and centrosome regulation are all independent functions of BRCA2. This observation validates our use of three independent assays for assessment of BRCA2 function in this study.

We also used this series of assays to evaluate BRCA2 function because they measure the cellular, as opposed to biochemical effect, of BRCA2 UCVs and may be more relevant to cancer development. Another important feature of this study is that full-length BRCA2 proteins were used to assess the effects of the UCVs on BRCA2 function. Whereas partial proteins containing specific functional domains can be useful for evaluation of protein function in certain circumstances, we believe that it is best to place mutations in the context of a complete protein especially when the protein is multifunctional and when cellular rather than biochemical assays are being applied.

In an effort to apply the results from the functional assays to an evaluation of the disease causality/predisposition associated with the BRCA2 UCVs we combined these data with results from genetic analyses of the UCVs (in a comprehensive approach to classification of BRCA2 UCVs (6). Specifically, we used the likelihood of disease causality based on cosegregation with cancer in families (Table 2), evidence of co-occurrence with other deleterious mutations (Table 2), and sequence conservation

of the relevant amino acids (Table 1), and physicochemical differences in wild-type and mutant amino acids (Table 1) as previously outlined by Goldgar et al. (6). The results from the functional assays were remarkably consistent with those from the individual likelihood models and with the combined overall odds of causality and suggest that these assays can discriminate between deleterious/high-risk and neutral/low-risk UCVs. Indeed, when combining these data, it is apparent that the D2723H UCV is a disease-causing mutation in BRCA2 (6). This is an important finding when considering that 24 families are known to carry this mutation. The data are less compelling for the R2659K UCV because the limited number of families carrying the mutation result in equivocal odds of causality in the cosegregation model and in the combined odds model. In this situation, the functional assays are the only means of defining the disease causality associated with the variant, and they suggest that R2659K is disease causing. The likelihood models also are consistent with the functional assays in defining Y42C (6) and E462G as neutral/low-risk variants. In both cases, the combined odds against causality are highly significant (Table 2), and agree completely with the finding that both variants have no effect on BRCA2 function. Whereas we cannot completely rule out that the UCVs are associated with a low risk of cancer, it is possible to classify them as not clinically relevant. Similar statements can be made about the common K3326X and N372H variants. Conversely, whereas the V2908G UCV might be classified as neutral/low risk based on a high GMS, evolutionary divergence, and limited influence on BRCA2 function in the functional assays, it remains possible that not all functions of BRCA2 are being evaluated in these assays. Thus, the definition of its role in cancer predisposition at a clinical level should await further experience in interpretation of these assays. Similarly, the

T1302del, E1382del, and T25151 variants must remain unclassified in terms of disease causality because of equivocal odds of causality in the likelihood model due to limited numbers and availability of pedigrees, and because it is not yet clear that inactivation of specific functions of BRCA2 as defined by the functional assays can increase predisposition to cancer.

The power in combining various UCV classification methods (6) is evident from our ability to establish the role of the D2723H, R2659K, Y42C, and E462G mutations in cancer predisposition. This suggests that many other BRCA2 UCVs might be classified in much the same manner leading to improved risk evaluation for UCV carriers and enhanced identification of individuals at risk or not at elevated risk of cancer. Importantly, this study also suggests that functional assays may also prove useful for classification of BRCA2 UCVs when cosegregation or co-occurrence data are limiting, as is the case for the majority of identified UCVs. However, a more complete understanding of all of the functions of BRCA2 will be needed so that mutations are not excluded as deleterious on the basis of incomplete information. In addition, a careful determination of the sensitivity and specificity of each assay will be needed before this approach to classifying variants can be utilized for clinical purposes.

Acknowledgments

Received 8/30/2004; revised 10/22/2004; accepted 11/12/2004.

Grant support: Breast Cancer Research Foundation awards, U.S. Army Medical Research and Materiel Command grant DAMD17-1-00-0328, NIH grant P01 CA82267, and American Cancer Society grant CCE-107497 (F.J. Couch); Ruth L. Kirschstein National Research Training Award (K. Wu); and U.S. Army Medical Research and Materiel Command postdoctoral fellowship grant DAMD17-1-00-0424 (S. Hinson).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

- Tavtigian SV, Simard J, Rommens J, et al. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet* 1996;12:333-7.
- Wooster R, Weber BL. Breast and ovarian cancer. *N Engl J Med* 2003;348:2339-47.
- van der Heijden MS, Yeo CJ, Hruban RH, Kern SE. Fanconi anemia gene mutations in young-onset pancreatic cancer. *Cancer Res* 2003;63:2585-8.
- Spain BH, Larson CJ, Shihabuddin LS, Gage FH, Verma IM. Truncated BRCA2 is cytoplasmic: implications for cancer-linked mutations. *Proc Natl Acad Sci U S A* 1999;96:13920-5.
- Frank TS, Deffenbaugh AM, Reid JE, et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 2002;20:1480-90.
- Goldgar DE, Easton DF, Deffenbaugh AM, et al. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* 2004;75:535-44.
- Miller MP, Kumar S. Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* 2001;10:2319-28.
- Abkevich V, Zharkikh A, Deffenbaugh AM, et al. Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation. *J Med Genet* 2004;41:492-507.
- Scott SP, Bendix R, Chen P, et al. Missense mutations but not allelic variants alter the function of ATM by dominant interference in patients with breast cancer. *Proc Natl Acad Sci U S A* 2002;99:925-30.
- Vallon-Christersson J, Cayan C, Haraldsson K, et al. Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. *Hum Mol Genet* 2001;10:353-60.
- Scully R, Ganesan S, Vlasakova K, et al. Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol Cell* 1999;4:1093-9.
- Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 2002;108:171-82.
- Wong AK, Pero R, Ormonde PA, Tavtigian SV, Bartel PL. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *J Biol Chem* 1997;272:31941-4.
- Mizuta R, LaSalle JM, Cheng HL, et al. RAB22 and RAB163/mouse BRCA2: proteins that specifically interact with the RAD51 protein. *Proc Natl Acad Sci U S A* 1997;94:6927-32.
- Moynahan ME, Pierce AJ, Jasin M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell* 2001;7:263-72.
- Sharan SK, Morimatsu M, Albrecht U, et al. Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* 1997;386:804-10.
- Connor F, Bertwistle D, Mee PJ, et al. Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation. *Nat Genet* 1997;17:423-30.
- Patel KJ, Yu VP, Lee H, et al. Involvement of Brca2 in DNA repair. *Mol Cell* 1998;1:347-57.
- Morimatsu M, Donoho G, Hasty P. Cells deleted for Brca2 COOH terminus exhibit hypersensitivity to γ -radiation and premature senescence. *Cancer Res* 1998;58:3441-7.
- Yu VP, Koehler M, Steinlein C, et al. Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev* 2000;14:1400-6.
- Yuan SF, Lee SY, Chen G, et al. BRCA2 is required for ionizing radiation-induced assembly of rad51 complex *in vivo*. *Cancer Res* 1999;59:3547-51.
- Tutt A, Bertwistle D, Valentine J, et al. Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *EMBO J* 2001;20:4704-16.
- Kraakman-van der Zwet M, Overkamp WJ, van Lange REE, et al. Brca2 (XRCC11) deficiency results in radioresistant DNA synthesis and a higher frequency of spontaneous deletions. *Mol Cell Biol* 2002;22:669-79.
- Howlett NG, Taniguchi T, Olson S, et al. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 2002;297:606-9.
- Marmorstein LY, Kinev AV, Chan GKT, et al. A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. *Cell* 2001;104:247-57.
- Tutt A, Gabriel A, Bertwistle D, et al. Absence of Brca2 causes genome instability by chromosome breakage and loss associated with centrosome amplification. *Curr Biol* 1999;9:1107-10.
- Wu KJ, Jiang SW, Couch FJ. p53 mediates repression of the BRCA2 promoter and down-regulation of BRCA2 mRNA and protein levels in response to DNA damage. *J Biol Chem* 2003;278:15652-60.
- Lutz W, Lingle WL, McCormick D, Greenwood TM, Salisbury JL. Phosphorylation of centrin during the cell cycle and its role in centriole separation preceding centrosome duplication. *J Biol Chem* 2001;276:20774-80.
- Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000;302:205-17.

30. Antoniou A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003;72:1117-30.
31. Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 2003;73:652-5.
32. Struwing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N Engl J Med* 1997;336:1401-8.
33. Mazoyer S, Dunning AM, Serova O, et al. A polymorphic stop codon in BRCA2. *Nat Genet* 1996;14:253-4.
34. Hofmann W, Horn D, Huttner C, Classen E, Scherneck S. The BRCA2 variant 8204G>A is a splicing mutation and results in an in frame deletion of the gene. *J Med Genet* 2003;40:e23.
35. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* 1999;13:2633-8.
36. Ward IM, Reina-San-Martin B, Orlaru A, et al. 53BP1 is required for class switch recombination. *J Cell Biol* 2004;165:459-64.
37. Daniels MJ, Wang Y, Lee M, Venkitaraman AR. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science* 2004;306:876-9.
38. Zuckerkandl E, Pauling L. Molecules as documents of evolutionary history. *J Theor Biol* 1965;8:357-66.
39. Jukes TH, King JL. Deleterious mutations and neutral substitutions. *Nature* 1971;231:114-5.
40. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* 1974;185:862-4.
41. Phelan CM, Dapic V, Tice B, et al. Classification of BRCA1 missense variants of unknown clinical significance. *J Med Genet*. In press 2004.
42. Couch FJ, Weber BL, Borresen AL, et al. Mutations and polymorphisms in the familial early-onset breast cancer (BRCA1) gene. *Hum Mutat* 1996;8:8-18.
43. Chen PL, Chen CF, Chen YM, et al. The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment. *Proc Natl Acad Sci U S A* 1998;95:5287-92.



Genetic, functional and histopathological evaluation of two C-terminal BRCA1 missense variants

Paul K Lovelock, Sue Healey, Wendy Au, Eleanor Sum, Andrea Tesoriero, Ee Ming Wong, Shannon Hinson, Ross Brinkworth, Anna Bekessy, Orland Diez, Louise Izatt, Ellen Solomon, Mark Jenkins, Helene Renard, John Hopper, Paul Waring, kConFab Investigators, Sean V Tavtigian, David Goldgar, Geoffrey J Lindeman, Jane E Visvader, Fergus J Couch, Beric R Henderson, Melissa Southey, Georgia Chenevix-Trench, Amanda B Spurdle and Melissa A Brown

J. Med. Genet. published online 2 Jun 2005;
doi:10.1136/jmg.2005.033258

Updated information and services can be found at:
<http://jmg.bmjournals.com/cgi/content/abstract/jmg.2005.033258v2>

These include:

Rapid responses

You can respond to this article at:
<http://jmg.bmjournals.com/cgi/eletter-submit/jmg.2005.033258v2>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

Online First contains unedited articles in manuscript form that have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Online First articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Online First articles must include the digital object identifier (DOIs) and date of initial publication.

To order reprints of this article go to:
<http://www.bmjournals.com/cgi/reprintform>

To subscribe to *Journal of Medical Genetics* go to:
<http://www.bmjournals.com/subscriptions/>

Genetic, functional and histopathological evaluation of two C-terminal BRCA1 missense variants

Paul K. Lovelock^{1,2}, Sue Healey², Wendy Au³, Eleanor Y.M. Sum⁴, Andrea Tesoriero⁵, Ee Ming Wong⁵, Shannon Hinson⁶, Ross Brinkworth¹, Anna Bekessy², Orland Diez⁷, Louise Izatt⁸, Ellen Solomon⁸, Mark Jenkins⁵, Helene Renard⁹, John Hopper⁵, Paul Waring¹⁰, kConFab¹⁰, Sean V. Tavtigian⁹, David Goldgar⁹, Geoffrey J. Lindeman⁴, Jane E. Visvader⁴, Fergus J. Couch⁶, Beric R. Henderson³, Melissa Southey^{5,9}, Georgia Chenevix-Trench², Amanda B. Spurdle², Melissa A. Brown¹

¹School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Australia; ²Queensland Institute of Medical Research, Brisbane, Australia; ³Westmead Institute for Cancer Research, University of Sydney, Westmead Millennium Institute at Westmead Hospital, Sydney, Australia; ⁴The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia; ⁵University of Melbourne, Melbourne, Australia; ⁶Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.; ⁷Servei de Genetica, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ⁸Division of Genetics and Molecular Medicine, King's College, Guy's Hospital, London, United Kingdom; ⁹International Agency for Research on Cancer, Lyon, France; ¹⁰Peter MacCallum Cancer Centre, Melbourne, Australia.

Corresponding Author:

Georgia Chenevix-Trench, Queensland Institute of Medical Research, Herston, Brisbane Queensland, Australia 4000.
Ph +61 7 3362 0390 Fax +61 7 3362 0105 E-mail Georgia.Trench@qimr.edu.au

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in JMG and any other BMJPG products and sublicenses such use and exploit all subsidiary rights, as set out in our license (<http://jmg.bmjournals.com/misc/ifora/licenceform.shtml>).

Keywords: BRCA1, unclassified variants, functional analysis

Introduction: The vast majority of *BRCA1* missense sequence variants remain uncharacterised for their possible effect on protein expression and function, and therefore are unclassified in terms of their pathogenicity. *BRCA1* plays diverse cellular roles and it is unlikely that any single functional assay will accurately reflect the total cellular implications of missense mutations in this gene. The aim of this study was to elucidate the effect of two *BRCA1* variants, 5236G>C (G1706A) and 5242C>A (A1708E) on *BRCA1* function, and to survey the relative usefulness of several assays to direct the characterisation of other unclassified variants in *BRCA* genes.

Methods: A range of bioinformatic, genetic and histopathological analyses, and *in vitro* functional assays were performed.

Results: These assays indicated that the 1708E variant was associated with the disruption of different cellular functions of *BRCA1*. In transient transfection experiments in T47D and 293T cells, the 1708E product was mislocalised to the cytoplasm and induced centrosome amplification in 293T cells. The 1708E variant also failed to transactivate transcription of reporter constructs in mammalian transcriptional transactivation assays. In contrast the 1706A variant displayed a phenotype comparable to wild-type *BRCA1* in these assays. Consistent with functional data, tumours from 1708E carriers showed typical *BRCA1* pathology while tumour material from 1706A carriers displayed few histopathological features associated with *BRCA1*-related tumours.

Discussion: A comprehensive range of genetic, bioinformatic and functional analyses have been combined for the characterisation of *BRCA1* unclassified sequence variants. Consistent with the functional analyses, the combined odds of causality calculated for the 1706A variant after multifactorial likelihood analysis (1:142) indicates a definitive classification of this variant as 'benign'. In contrast, functional assays of the 1708E variant indicate that it is pathogenic, possibly through subcellular mislocalisation. However, the combined odds of 262:1 in favour of causality of this variant does not meet the minimal ratio of 1000:1 for classification as pathogenic, and A1708E remains formally designated as unclassified. Our findings highlight the importance of comprehensive genetic information, together with detailed functional analysis for the definitive categorisation of unclassified sequence variants. This combination of analyses may have direct application to the characterisation of other unclassified variants in *BRCA1* and *BRCA2*.

Introduction

The pathogenicity of many genetic variants in disease-associated genes can be predicted from the nature of the genetic variation. For example, sequence changes that prevent protein expression or that cause loss of important functional domains can be classified as loss-of-function mutations. However, single exonic nucleotide changes can present a challenge. Such changes have been associated with alterations in transcript stability [1], transcript splicing [2], translation efficiency [1], protein folding [3], protein-protein interactions [4] and the capacity to perform specific cellular functions [4]. Interpretation of the pathogenicity of single nucleotide changes is a challenge for the clinical management of many inherited diseases and predispositions including Duchenne muscular dystrophy [5], cystic fibrosis [6], X-linked mental retardation [7], and inherited cancer syndromes such as Hereditary Non Polyposis Colon Cancer [4] and familial breast cancer [8]. There is a growing interest in developing efficient and reliable ways to classify the pathogenicity of these variants and a variety of approaches have been reported recently. These include analyses of patterns of co-segregation [9], assessment of variant frequency in unaffected controls [6], predictions based on the position and nature of the amino acid change [7] and biochemical and functional assays [17-35].

BRCA1 is a breast cancer susceptibility gene encoding a protein of 1863 amino acids with multiple roles in DNA repair, transcriptional activation, apoptosis and cell cycle regulation (reviewed in [10]). Pathogenic mutations in the *BRCA1* gene have been identified in about 15-20% of families with multiple cases of breast and ovarian cancer [11]. However, the mechanism(s) by which most mutations in *BRCA1* contribute to breast cancer are poorly understood. A further level of complexity is added by the spectrum of unclassified sequence variants described in this gene to date, with over 1000 different unclassified sequence variants in *BRCA1* reported on the Breast Information Core (BIC) database alone (<http://research.nhgri.nih.gov/bic>). The pathogenicity of only a small number of these variants has been inferred genetically [12-16] or tested functionally [17-35]. Underscoring the interest in and need for accurate classification of sequence variants in BRCA1 are recent descriptions of novel approaches to predict the pathogenicity of non-synonymous amino acid substitutions. These approaches include analysing interspecies sequence variation [36, 37] and a combination of genetic and bioinformatic predictive investigations [8]. This approach offers some advantages over more laborious and expensive experimental analyses, particularly in a clinical setting. How well the data from such predictions correlates with the results of biochemical and functional studies on the same variant however, is yet to be established.

We have set out to determine the pathogenicity of two sequence variants, G1706A and A1708E, using a number of genetic, bioinformatic, biochemical and cellular investigations. Although A1708E has been classified as a missense mutation by BIC, the precise molecular defect is not understood, nor has the variant been formally evaluated in the multifactorial model of causality [8]. These variants were initially selected for analysis because of their proximity within the C-terminal region of BRCA1, for which several functional assays have been developed. The C-terminal region is a highly conserved structure containing two *BRCA1* C-terminal (BRCT) tandem repeat domains. Missense changes within this motif cause protein folding defects [24, 31] and inhibit transcriptional transactivation [20]. The G1706A and A1708E variants are located in the first BRCT domain (amino acids 1650-1736) and have been partially characterised by predictive modelling and some functional assays (reviewed in [38]). Here we present data from a wider variety of approaches that further test these predictions and provide a cellular and molecular basis for pathogenic risk assessment.

Materials and methods

Genetic analysis.

Australian study pedigrees were recruited by the Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab) according to eligibility criteria established by the consortium (www.kconfab.org/epidemiology/eligibility). Two Australian pedigrees in which affected index cases were ascertained to be carriers of the *BRCA1* 5236G>C (G1706A) variant, and one Australian pedigree with the affected index case carrying the 5242C>A (A1708E) variant were selected for analysis. Sequencing or Denaturing High Performance Liquid Chromatography (DHPLC) analysis of the coding and flanking intronic regions of *BRCA1* and *BRCA2* in the index cases revealed no other mutations.

A Spanish pedigree carrying the G1706A variant was ascertained by the Genetics Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. In the index case of this pedigree, the 1706A change was detected by mutation analysis of *BRCA1* and *BRCA2* using SSCP methodology and sequencing to confirm the nucleotide alteration, and relatives were screened for the variant by sequencing.

Two English A1708E pedigrees were ascertained by the Clinical Genetics Service at Guy's Hospital in London, United Kingdom. The A1708E variant was identified in the index case of the first pedigree (UK1708E1) by SSCP, followed by sequencing to confirm the mutation. The other member of the family tested was screened for the variant by restriction enzyme analysis. In the second English pedigree (UK1708E2), the A1708E variant was identified in the proband by hydroxylamine fluorescent chemical cleavage of mismatch followed by sequencing confirmation. Subsequently, other carriers within the family were identified by restriction enzyme assay.

In total, 29 individuals from the G1706A pedigrees and nineteen individuals from the A1708E pedigrees were genotyped. We estimated the odds for causality associated with these variants using a Bayes factor analysis by maximising the evidence in favour of causality over the hazard ratio (HR), based on the method described by Thompson et al. [9]. In order to determine the penetrance associated with the variants we used a modified segregation analysis [39], which estimates cumulative risk to age 70 of breast cancer in carriers (although, due to the small sample size, 95% confidence intervals of cumulative risk may be underestimated), where models were fitted under maximum likelihood theory using the MENDEL statistical package [40]. As controls, 180 unaffected females over the age of 45 with no reported family history of breast cancer were recruited through the Australian Breast Cancer Family Study [41], and screened by DHPLC for these variants. Variants absent in a sample of 180 individuals are estimated to have an allele frequency with an upper 95% confidence limit less than 1%, the formal definition of a polymorphism. All pedigrees used in the genetic analysis can be supplied by the corresponding author upon request.

Loss of heterozygosity analysis of tumours.

Loss of heterozygosity (LOH) at *BRCA1* was assessed for index cases by macrodissection of tumour-rich (70%) regions of available paraffin sections. DNA was extracted as described previously [42]. A PCR product including the 1706 and 1708 positions was generated using the primer pair 5236SeqF (5'-GAGGCTCTTTAGCTTCTTAGGAC-3') and 5236SeqR (5'-AAACGTTAGGTGTAATAAATGCAA-3'), and sequenced in both directions using the ABI Big Dye terminator system and the forward (5236SeqF) or reverse (5236SeqR) primers.

LOH was scored by the significant reduction (<50% peak height relative to normal sequence trace) or absence of the heterozygous peak seen in the germline control.

Single Nucleotide Primer Extension (SNUPE) assay.

To verify expression of the wild-type and variant BRCA1 transcripts in lymphoblastoid cell-lines (LCLs) generated from heterozygous carriers of these variants, total RNA was extracted from LCLs using Trizol reagent (Invitrogen), and analysed by Single Nucleotide Primer Extension (SNUPE) [43]. A 383bp fragment of cDNA flanking the 1706 and 1708 positions of BRCA1 was generated by RT-PCR using the Superscript III Reverse Transcriptase kit from Invitrogen and primers Check1F (5'-CCAGAAGAATTTATGCTCGTG) and OR (5'-CAGCTGTACCATCCATTCCA). As positive and negative controls, fragments were also generated with the same primers from the pZeoSV-based expression constructs containing wild-type, G1706A and A1708E BRCA1 cDNA (see below). PCR products were purified using a Qiagen QIAQuick PCR purification kit. SNUPE assays were performed on these fragments with radiolabelled dNTPs and the primer pairs SNUPE1706F and SNUPE1706R (5'-GGACACTGAAATATTTCTAG and 5'-ACCCATTTTCTCCCGCAATT, respectively) and SNUPE1708F and SNUPE1708R (5'-CTGAAATATTTTCTAGGAATTG and 5'-GCTAACTACCCATTTTCTCCC, respectively). Radiolabelled fragments were separated on a 10% denaturing acrylamide gel, and visualised by autoradiography.

Grantham alignments.

Extensive protein multiple sequence alignments were made with the online alignment engines Toffee and 3Dcoffee [44]. The degree of sequence variation present in the alignment was used to calculate the number of positions that are under strong functional constraint or not [37], and the likelihood ratio for whether a substitution at any particular position will be deleterious or not [8]. Grantham scores [45] were calculated both for each position in the alignments and for each missense mutation versus the canonical human sequence. The relationship between the two Grantham scores was used to determine the fit between the missense substitution and the range of variation observed at its position in the alignment as described previously [8, 37]. The analysis included 13 full-length BRCA1 sequences, the most divergent of which was from the tunicate, *Ciona*.

Protein modelling

Molecular modelling was carried out on the simulated crystal structure of the BRCT repeat region of BRCA1 (JNX1.pdb) that incorporates positions 1706 and 1708 [46], using an SGI work-station running the Insight II software package (Accelrys, San Diego).

ESE Analysis

The wild-type BRCA1 cDNA (NCBI Accession No. U14680) was examined on an exon-by-exon basis for sequences that act as binding sites for the serine/arginine-rich family of splicing enhancers using the ESEfinder program (<http://rulai.cshl.edu/tools/ESE/>). The variant cDNA sequence was then screened by ESEfinder and compared to the wild-type sequence to identify any loss or gain of predicted SR binding sites.

Alternate Splicing analysis.

Exon 18 sequence containing the 1706A and 1708E variants was analysed for splice acceptor or donor sites using SpliceSiteFinder (www.genet.sickkids.on.ca/~ali.splicesitefinder.html). To test for possible alterations in BRCA1 splicing, EBV-transformed LCLs generated from variant carriers were treated for 4 hours with cycloheximide (100ug/mL) to stabilize RNA species. RNA was then extracted using the TriPure Isolation Reagent (Roche). cDNA

synthesis was performed and a 345bp RT-PCR product was amplified using the one step Titan™ One Tube RT PCR-system (Roche) with forward primer 5'ATGCTCGTGTACAAGTTTG 3' (BRCA1 exon 17) and reverse primer 5'CTGTGGGCATGTTGGTGAA 3'(BRCA1 exon 21). Products were visualised on a 2% agarose gel.

Expression plasmids and constructs.

A plasmid containing a full-length BRCA1 cDNA with three N-terminal c-myc tags and an in-frame Kozak sequence on a pcDNA3.1 backbone was kindly provided by D. Livingston (Dana Farber Cancer Institute, Harvard Medical School) for subcloning and generation of the mutagenised expression constructs. The BRCA1 cDNA fragment of this parental plasmid was entirely sequenced prior to subcloning, and several SNPs (rs1799949, rs799917, rs16940, rs16941, rs16942, and rs4986849) were identified. These SNPs are published on the NCBI and BIC SNP databases and form a common haplotype [47]. A fragment containing the c-myc tags, Kozak sequence, BRCA1 cDNA and 114bp of 3'UTR was excised from the parental plasmid with XhoI and ligated into the XhoI site of the pZeoSV vector (Invitrogen). The 1706A and 1708E mutations were generated in this plasmid using the Stratagene QuikChange PCR site-directed mutagenesis protocol. Primer pairs for the G1706A mutagenesis protocol were Quik5236F (5'-GACACTGAAATATTTTCTAGCAATTGCGGGAGGAAAATGGG-3' and Quik5236R (5'-CCCATTTTCTCCCGCAATTGCTAGAAAATATTTTCAGTGTC-3'). The primer pairs for the A1708E mutagenesis protocol were RTW5242F (5'-CTGAAATATTTTCTAGGAATTGAGGGAGGAAAATGGGTAGTTAG-3') and RTW5242R (5'-CTAACTACCCATTTTCTCCCTCAATTCCTAGAAAATATTTTCAG-3'). The BRCA1 5382insC deletion mutant expression plasmid was kindly donated by Dr B. Weber, University of Pennsylvania, USA. Large-scale DNA preps were made using the Qiagen Plasmid MaxiKit and the inserts of each plasmid were sequenced entirely to verify their identity.

Trypsin Sensitivity Assay.

A 1571bp product was amplified by PCR from the pZeo expression constructs containing the wild-type, 1706A and 1708E variants with a T7 forward primer 5'GGATCCTAATACGACTCACTATAGGAACAGACCACCATGGGTCTGAGTGACAA GGAATT3' and reverse primer 5'CTGGGGTATCAGGTAGGT3', encompassing BRCA1 exons 12-24. The products were translated *in vitro* using the Promega TNT Coupled Reticulocyte Lysates System incorporating ³⁵S methionine (NEN). Protein products were digested for 12 mins at 37°C using trypsin (Sigma) as described previously [31] at concentrations of 0, 0.6, 6, and 60 µg/ml. Protein products were separated by 14% SDS-PAGE and visualised by autoradiography.

Cytoplasmic localisation of ectopic BRCA1 in MCF7 cells.

MCF7 human breast cancer cells were maintained in Dulbecco's Modified Eagle's Media (DMEM) supplemented with 10% fetal calf serum (FCS), and grown at 37°C in a humidified 5% CO₂ atmosphere. Cells were seeded onto sterile glass coverslips and transfected at 50-60% confluency with 1 to 2 mg of plasmid DNA using Lipofectamine Reagent (Life Technologies) according to the manufacturer's instructions. At 6 hours post-transfection the transfection mix was removed and replaced with DMEM containing 10% FCS. Cells were fixed and processed 30 hours post-transfection for fluorescence microscopy. Transfected cells were fixed in 3.7% formalin/PBS for 15 minutes and processed for immunostaining. Myc-tagged ectopic BRCA1 was detected by immunofluorescence using the anti-Myc rabbit

polyclonal antibody A-14 (Santa Cruz Biotec). Myc antibody was detected with biotin-conjugated secondary antibodies (Santa Cruz) and Texas Red-avidin D (Vector Laboratories). Cell nuclei were counterstained with the chromosome dye Hoechst 33285 (Sigma). The subcellular localisation of each ectopic protein was determined by scoring cells using an Olympus BX40 epifluorescence microscope, and the proportion of cells displaying nuclear, nuclear/cytoplasmic or cytoplasmic staining of BRCA1 was determined as previously described [21]. Digital images were collected using a SPOT camera.

Transcriptional transactivation assays.

The 1706A and 1708E variants were generated in a pGAL4B expression construct containing the BRCA1 activation domain within a region spanning amino acids 1293-1863 [48] using the Stratagene QuikChange-PCR site-directed mutagenesis protocol. Primer pairs for generating the 1706A variant were G1706AF 5'-GGACACTGAAATATTTTCTAGCAATTGCGGGAGGAAAATG-3' and G1706AR 5'-CATTTTCCTCCCGCAATTGCTAGAAAATATTTTCAGTGTCC-3' and primers for the A1708E mutagenesis protocol were RTW5242F/R, described above. Assays were carried out as described previously [48], with two exceptions. MCF7 cells were used instead of 293T cells in the mammalian cell assay and Fugene 6 (Roche) transfection reagent, used according to manufacturer's instructions, replaced the calcium phosphate precipitation method of cell transfection.

Western Blots.

Transiently transfected MCF7 cells were counted and lysate prepared from identical numbers of cells for each transfection. Lysate was electrophoresed, blotted and probed with FLAG or GAL4 antibodies as described previously [48].

Centrosome amplification Immunofluorescence and Confocal Microscopy.

Cells were cultured on glass cover slips and transfected with Myc-BRCA1 wild-type and mutant constructs using Fugene 6 according to manufacturer's instructions. For indirect immunofluorescence, cells were fixed with cold methanol, permeabilized, and stained with primary anti-pericentrin (1:800) polyclonal (Santa Cruz Biotechnologies) and anti-myc (9E10) (1:200) monoclonal (Santa Cruz Biotechnologies) antibodies four days after transfection. Texas-red goat anti-rabbit (1:800) and Oregon-green goat anti-mouse (1:800) secondary antibodies were subsequently added, along with 1µg/ml Hoechst (Molecular Probes). Centrosome numbers were counted in a minimum of 50 Myc expressing cells from each of two independent experiments. For subcellular localisation experiments, 48 hours post-transfection the cells were fixed with cold methanol, stained with 1µg/ml Hoechst (Molecular Probes), anti-myc (9E10) monoclonal (Santa Cruz Biotechnologies) primary antibody, and Oregon-green goat anti-mouse (1:800) secondary antibody. All images were acquired with a Zeiss LSM510 confocal microscope.

Histopathology.

Available tumour sections from all the pedigrees were analysed for BRCA1-like histology by pathologists blind to mutation status. Sections were scored for parameters recognised to be associated with BRCA1 tumours [49-54]. BRCA1-like phenotype was designated as 'medullary' or 'atypical medullary carcinoma' or 'ductal / no special type', with high grade, a high mitotic count (>16 mitotic figures/10 high power fields), and one or more of the following features: >25% pushing margin; confluent necrosis; prominent lymphocytic infiltrate. The estrogen receptor (ER) and progesterone receptor (PR) status of the tumour sample was available from some pathology reports.

Results

Genetic and LOH analysis

In Australian G1706A pedigree #1 (6-99-006, described in [35]), two of the three breast cancer patients (age at diagnosis 66 and 44 years) and the single ovarian cancer patient (age at diagnosis 45yrs) were found to be 1706A carriers whereas one woman diagnosed with breast cancer at age 53 was not a carrier [35]. The Bayes factor calculated for this family was 0.7086 with a likelihood for causality ratio of 1:1.3. In Australian G1706A pedigree #2, the sole surviving breast cancer patient, the single ovarian cancer patient (age at diagnosis 49 and 52yrs respectively), and two other unaffected women (aged 70 and 90yrs) were found to be 1706A carriers. In this pedigree, the variant was carried by the mother of the index case and not the father whose sister was diagnosed with breast cancer at age 40 years. The Bayes factor calculated for this family was 0.0317, with a likelihood for causality of 1:32. In the Spanish G1706A pedigree, the variant was not carried by the proband's mother who had a family history of breast and ovarian cancer and therefore was presumed to have been inherited from the father (not tested) who had no known family history of cancer. The Bayes factor analysis for the Spanish G1706A family yielded a value of 0.0196 and a likelihood for causality of 1:51. The likelihood of causality for the 1706A variant using the Bayes factor analysis method for the combined Australian and Spanish families was 1:2274.

In the Australian A1708E pedigree, both living members with breast cancer (ages at diagnosis 47 and 58 years) were found to carry the 1708E variant. The deceased mother of a carrier was diagnosed with breast cancer at age 47, and a deceased cousin of a carrier was diagnosed with breast cancer at age 30 years. For the Australian 1708E pedigree, the likelihood of causality was 2.8:1.

In the English pedigree UK1708E1 the variant was identified in the proband (breast cancer diagnosed at age 28) and in the mother of the proband (ovarian cancer diagnosed at age 55). No other individuals in this family were genotyped. The Bayes Factor analysis for this pedigree yielded odds of causality of 2.0:1. In the pedigree UK1708E2, two A1708E variant carriers were identified. One carrier was diagnosed with breast cancer at age 39 years. The other carrier has not been diagnosed with cancer to date. Two other individuals from this pedigree without breast cancer were tested and found to be negative for the A1708E variant. The Bayes Factor analysis for this pedigree yielded a ratio of 0.7982:1, equating to an odds of causality of 1:1.3. The total Bayes factor for the combined Australian and English A1708E pedigrees was 4.5:1 in favour of causality, and the estimated penetrance for A1708E was 100% (95% C.I. 98-100%) to age 70 years.

Neither the 1706A nor 1708E variants were detected in any of the 180 control samples. Loss of the wild-type allele was demonstrated in tumour tissue from one case from the Australian G1706A #1/6-99-006 and A1708E pedigrees (data not shown). However, in the tumour sample isolated from the Spanish G1706A pedigree, no LOH was detected. No LOH data were available from the United Kingdom A1708E pedigrees.

SNuPE assay.

The relative expression of wild-type and variant alleles was measured using the SNuPE assay. Wild-type and variant alleles were expressed at approximately equal levels (Figure 1) indicating that the variant sequences do not affect the stability of their respective transcripts.

ESE and splice variant analysis.

The ESE Finder program was used to predict the effects of the 1706A and 1708E variants on serine-arginine rich splicing factor binding to consensus exonic enhancer sequences. While this program may not identify all ESEs, and splicing predictions may not be definitive [34], replacement of the wild-type BRCA1 sequence with the 1706A variant in exon 18 was predicted by ESE Finder to disrupt a consensus binding site for the SC35 splicing enhancer. ESE Finder modelling of the 1708E variant predicted disruption of SC35 and SRp55 splicing enhancer binding sites. Further, SpliceSiteFinder analysis of the 1708E variant predicted the creation of an additional acceptor sequence. However, RT-PCR analysis revealed no evidence of aberrant splicing (data not shown). Thus, the 1706A and 1708E variants did not appear to affect the splicing of their respective primary transcripts, supporting previous findings for 1706A by Campos et al [34].

In silico analysis and protein modelling.

Both variants occur in a highly conserved region of the BRCT domain of BRCA1. Analysis of the degree of cross-species variation at the positions of these two substitutions predicted that both are likely to be deleterious. Position 1708 is invariant, resulting in a likelihood ratio (LR) of 58:1 in favour of causality for 1708E. The Grantham score for the substitution Alanine > E (Glutamic Acid) is 107. This score is above the average Grantham score for all possible missense single nucleotide substitutions in BRCA1 (=78), tending to reinforce the expectation that the substitution is deleterious. Exactly one substitution is observed at position 1706, from which we calculate a LR of 16:1 in favour of causality for 1706A. Interestingly, the cross-species substitution at 1706 is Glycine > Alanine, in the tunicate *Ciona* (SVT, unpublished data). This means that the human substitution 1706A is outside the range of variation observed in vertebrates but within the range of variation observed from human to tunicate. There are no firm data from which to determine whether, or by how much, that observation modifies the likelihood that 1706A is deleterious. But we also note that the Grantham score for Glycine > Alanine is 60, which is below the average Grantham score for all possible missense single nucleotide substitutions in BRCA1. Coupled with the observation of Alanine at the corresponding position in the tunicate sequence, consideration of the Grantham scores would tend to decrease the likelihood that the 1706A substitution is actually high-risk.

Protein structure modelling predicted that the substitution of an Alanine at position 1706 would cause a moderate increase in hydrophobicity and size. The Glycine at position 1706 occurs within an α -helix structure running from Leucine 1701 to Isoleucine 1707. Substitution of Glycine with a larger Alanine at 1706 was predicted to disrupt a bend in the helix, possibly with functional consequences if normal protein function is reliant on the maintenance of this conformation. The substitution was predicted to cause clashes between the side-chains of Valine 1687 and Valine 1713, which were not relieved by changes in the side-chain rotamers of each residue.

The 1708E variant was predicted to cause significant disruption of protein structure, with a shift from a highly hydrophobic residue (Alanine) to a larger and strongly hydrophilic residue (Glutamic acid). Examination of the BRCA1 crystal structure showed that the Alanine, Glycine and Glycine residues at positions 1708, 1709 and 1710 respectively form a tight turn structure likely to permit only small residues as replacements. Substitution of the Alanine with the much larger Glutamic Acid residue at position 1708 was predicted to be incompatible with maintaining the turn structure. Furthermore, this substitution was predicted to produce clashes, mostly between the carboxyl group of Glutamine at 1708 and the Methionine at 1783 in the second BRCT domain. Clashes predicted for the Glutamate residue

were all with the main chain (peptide) atoms and the γ -methylene group of the Methionine, where the clashes involved distances that were too small and unfavourable interactions between hydrophobic and charged hydrophilic atoms. Changes in the side-chain rotamers of each residue were not predicted to relieve the clash.

Proteolytic degradation assays.

Consistent with the conformational changes predicted by protein modelling for 1708E, resistance of this variant protein fragment to trypsin degradation was reduced compared to wild-type and 1706A proteins. Degradation of the 1708E protein fragment by trypsin occurred at a ten-fold lower concentration of the enzyme than required for the 1706A and wild-type BRCA1 fragments (Figure 2), and supports previous findings showing that the 1708E variant was less resistant to proteolytic digestion [31]. The wild-type and 1706A variant BRCA1 proteins showed similar degradation profiles, indicating that the minor structural change predicted for the 1706A variant protein did not have a measurable effect on the stability of the BRCA1 protein in these assays.

Subcellular localisation of ectopically expressed variant and wild-type BRCA1.

The 1706A and 1708E sequence variants were introduced into myc-tagged full-length *BRCA1* cDNA expression constructs, which were then used to transfect various cell-lines for expression and functional analysis. In MCF7 cells transiently transfected with these constructs, normal nuclear localisation of the transiently expressed wild-type and 1706A proteins was indicated by ubiquitous staining in both the nuclear and cytoplasmic compartments after staining with myc antibodies for ectopic BRCA1 (Figure 3). In contrast, cytoplasmic mislocalisation of the BRCA1 1708E protein was observed. This predominantly cytoplasmic expression of the 1708E variant is similar in pattern to that previously described for another cancer-associated *BRCA1* variant, 5382insC [21]. However, while the nuclear localisation of the 5382insC variant was shown to be increased to wild-type levels with the co-transfection of BARD1 [21], in these experiments BARD1 expression did not restore nuclear localisation of the 1708E variant to a wild-type pattern of expression. Nuclear localisation of the wild-type and 1706A constructs was significantly enhanced by the co-transfection of BARD1.

Transcriptional Transactivation assays.

As the 1706A and 1708E variants map to the region of BRCA1 implicated in transcriptional transactivation, the potential effect of these sequence changes on this function of BRCA1 was investigated by introducing these changes into a cDNA construct encoding amino acids 1293 to 1863. In transiently transfected MCF7 cells, regulation of transcriptional transactivation by the 1706A variant resulted in reporter activity indistinguishable from wild-type levels (Figure 4A). In contrast, reporter construct activity in cells transfected with the cDNA construct containing 1708E was similar to those of the empty vector, suggesting minimal capacity for this variant protein to transactivate transcription. Western blot analysis of total protein indicated that levels of the 1708E protein were significantly lower than the wild type and 1706A proteins (Figure 4B) suggesting that the reduced transcriptional transactivation was likely to reflect aberrant expression rather than function of the BRCA1 protein.

Centrosome amplification.

Another important nuclear function of BRCA1 is regulation of centrosomes (reviewed in Starita et al [55]). To address the consequences of the two variants on this function of BRCA1, 293T cells were transfected with wild-type and variant constructs, followed by centrosome analysis and counting. Centrosome amplification was evident in cells transfected

with the BRCA1 1708E variant (Figure 5), but not in those transfected with the 1706A and wt BRCA1 constructs. Staining for c-Myc in the 293T cells also indicated that the 1708E variant is largely mislocalised to the cytoplasm and not transported to the nucleus, supporting data presented earlier (Figure 3) for MCF7 cells.

Tumour pathology.

The pathology of tumours from the 1706A and 1708E carriers in the Australian pedigrees was strikingly different (Table 1). The two tumours examined from 1708E carriers had many *BRCA1*-related features. In contrast, the histopathology of the 1706A tumour from the Australian pedigree #1/6-99-006 index case did not have a *BRCA1*-related tumour histological profile. Similarly, the Spanish G1706A pedigree tumour sample did not display features consistent with a *BRCA1* defect (Table 1).

Table 1. Histopathology of BRCA1 1706A and 1708E breast tumors shows that only the 1708E-associated tumors have a typical 'BRCA1-like' phenotype.

Histopathological feature	G1706A (Australian Pedigree #1)	G1706A (Spanish pedigree)	A1708E (Australian pedigree)	A1708E (U.K. Pedigree #2)
Histological subtype of invasive carcinoma	Ductal/ no special type	Ductal/ lobular	Medullary	Atypical medullary
Mitotic count	Score 1 (actual count 1)	Score 1	Score 3 (actual count 85)	Score 3 (actual count 103)
Overall Histological Grade	Grade 2	Grade 2	Grade 3	Grade 3
% Tumour diameter as continuous pushing margin	Absent	Absent	>75%	>75%
Confluent necrosis	Absent	Absent	Present	Absent
Tumour cells: discernible cell borders	Present	Present	Absent	Absent
ER/PR Status	Data not available	ER+, PR-	ER-, PR-	ER-, PR-
Conclusion	Not BRCA1-like	Not BRCA1-like	BRCA1-like	BRCA1-like

Discussion

Using genetic and bioinformatics analyses, functional assays and histopathology, we have attempted to determine the pathogenicity of the BRCA1 1706A and 1708E variants, and to assess these approaches for their potential use in the characterisation of other *BRCA1* and *BRCA2* unclassified variants.

Several functional assays indicated a severe functional abrogation of BRCA1 proteins carrying the 1708E variant. Protein modelling of the 1708E variant predicted a significant folding defect. Proteolytic digestion assays by us and others [31] supported this prediction by showing that a protein fragment containing the 1708E variant was significantly more susceptible to degradation *in vitro*. Pathogenicity of the 1708E variant was further supported by assays showing cytoplasmic mislocalisation of ectopically expressed BRCA1 protein carrying the 1708E variant in MCF7 and 293T cells. Cytoplasmic mislocalisation has been demonstrated in MCF7 cells for several other BRCA1 BRCT domain mutants of known pathogenicity, including the 5382insC, M1775R and Y1853X mutations [21].

The failure of co-transfected BARD1 to restore nuclear import of 1708E BRCA1 in MCF7 cells suggests that BARD1 is unable to complex with mutant BRCA1. The mechanism by which BARD1 and BRCA1 work together is currently being elucidated [21, 56-58] and although these two proteins interact via the N-terminal RING domain of BRCA1, it is possible that variant sequences in the C-terminus of BRCA1 affect the three dimensional protein structure. The predicted 1708E conformational change may induce an altered tertiary structure masking the BARD1-binding region and preventing the formation of the BRCA1/BARD1 heterodimer. This would impact on both the nuclear import and ubiquitin ligase functions of BRCA1. Alternatively, it is possible that the conformational change is sufficiently severe in terms of charge and size that nuclear import is physically impeded.

Cytoplasmic mislocalisation of the 1708E variant BRCA1 may impede its function as a regulator of transcription. Indeed, transcriptional transactivation by the 1708E variant was significantly lower than the wild-type protein in our assays. These data are consistent with previously reported disruption of BRCA1 transcriptional transactivation by the 1708E variant in mammalian [59] and yeast-based assays [60].

In the centrosome amplification assays, the 1708E variant displayed features consistent with pathogenicity. Centrosome amplification has been associated with over 70% of invasive breast tumours [61] and has been shown to be a major cause of mitotic defects in cancer cells leading to chromosomal instability, cell cycle checkpoint dysregulation and loss of tissue differentiation (reviewed in [62]). The C-terminus of BRCA1 has been implicated as an important mediator of centrosome regulation, through the formation of a ubiquitin ligase complex with BARD1 for the ubiquitination of γ -tubulin, a major centrosome component involved in the maintenance of centrosome number. Defects in the ubiquitination of γ -tubulin induced by overexpression of the N-terminal γ -tubulin binding fragment of BRCA1 in COS 7 cells [63] or by expression of mutant γ -tubulin in mammary Hs578T cells [55] caused centrosome amplification. C-terminal deletion mutants of BRCA1 also disrupted γ -tubulin ubiquitination [55]. Our experiments showed that, in the MCF7 breast cell line the 1708E variant was not localised to the nucleus with endogenous or ectopic BARD1, indicating that formation of the BRCA1/BARD1 complex was significantly diminished. We suggest that the centrosome defect associated with the 1708E variant may in part be a consequence of dysregulated ubiquitination of γ -tubulin and possibly other centrosome-associated proteins.

In support of the functional data, the histopathology review of the two 1708E-carrying tumours identified several features commonly associated with *BRCA1*-mutation carrying tumours. Evolutionary sequence alignment analysis gave odds in favour of causality of 58:1 for 1708E. However, the genetic evidence of pathogenicity from the A1708E pedigrees is weak, giving an odds of causality of 4.5:1. Applying the multifactorial likelihood model [8] to assess pathogenicity, the combined odds from these two approaches was 262:1 which does

not meet the criteria of Goldgar et al [8] of 1000:1 for a pathogenic variant. Nevertheless, the odds are consistent with the functional assays which indicated that this variant has features of a high-risk mutation.

In contrast, the protein stability, transcriptional transactivation and subcellular localisation of the 1706A variant were all comparable to the wild-type in our assays. Protein modelling of the 1706A variant predicted only a mild conformational change, and this prediction was borne out experimentally, with a proteolytic digestion profile similar to the wild-type protein. This variant also showed minimal effects on subcellular localisation, centrosome amplification, and transcriptional transactivation. Our transcriptional assay data contrasts with evidence from another research group [35], who reported that cells expressing 1706A showed a moderately reduced transcriptional transactivation capacity compared to wild-type *BRCA1* in both mammalian and yeast systems. However, these inconsistencies may be due to differences in cell lines and the length of *BRCA1* cDNA used.

The histopathology review of the G1706A-carrying tumour from the Australian and Spanish pedigrees did not have a *BRCA1*-associated tumour phenotype. The loss of the wild type allele in this tumour may simply reflect background LOH, which is estimated to be about 30% in breast tumours [64, 65]. Sequence alignment analysis gave odds in favour of causality of 16:1 for 1706A. The Bayesian odds of causality for this variant in the combined Australian and Spanish pedigrees studied were 1:2274. Using data from sequence alignment and co-segregation analysis, the combined odds of causality after applying the multifactorial likelihood model are 1:142. This meets the criteria of Goldgar et al [8] of 1:100 for a neutral variant, consistent with the functional assays indicating that this variant is not a high-risk mutation.

In conclusion, we have presented genetic, histopathological and functional evidence to suggest that the A1708E variant of *BRCA1* is pathogenic, as classified by BIC, possibly through subcellular mislocalisation. However, this variant remains formally designated as unclassified according to the criteria of Goldgar et al [8] after application of the multifactorial likelihood model using sequence alignment and co-segregation data. In contrast, analysis of the G1706A variant in functional assays suggests that it is unlikely to be associated with significant risk of breast cancer and may be classified as benign according to the criteria of Goldgar et al [8]. While classification of 1708E and other variants may ultimately be achieved by analysis of further co-segregation data, and inclusion of LOH and control frequency data in the multifactorial likelihood model, we suggest the application of an appropriate selection of the functional assays described here and elsewhere will enhance predictions of pathogenicity of unclassified sequence variants in *BRCA1* and *BRCA2*. Assay selection will largely be determined by the location of unclassified sequence variants within known structural or functional motifs in the gene. In our hands, the cytoplasmic mislocalisation, proteolytic sensitivity and centrosome amplification assays were found to be the most informative for the functional analysis of the G1706A and A1708E variants in the BRCT domain of *BRCA1*. These assays may also be useful for variants affecting other regions of the protein. In light of the diversity of roles played by *BRCA1* at the cellular level, it is unlikely that any single assay will conclusively characterise sequence variants in this gene. Where possible, a comprehensive approach utilising genetic, bioinformatic and functional analyses to classify sequence variants is most likely to reliably indicate potential pathogenicity, and may eventually also assist in delineating variants with moderate risk from those with high or low/no risk.

Acknowledgements

The authors would like to thank the kConFab research nurses and staff for data collection, Heather Thorne, Lynda Williams and Dani Surace for DNA preparation, Jan Groves for establishment of the LCLs, Eveline Niedermayr for supplying data, Don Roberts and Anna Marsh for technical assistance, the staff of the Familial Cancer Clinics for their support of kConFab, and the families for their participation. kConFab has been funded by the Kathleen Cuninghame Foundation, National Breast Cancer Foundation, National Health and Medical Research Council (NHMRC), Cancer Council of Victoria, Cancer Council of South Australia, Queensland Cancer Fund, Cancer Council of New South Wales, Cancer Foundation of Western Australia, and Cancer Council of Tasmania. We also thank ABCFS researchers Margaret McCredie and Graham Giles for their roles in the establishment of the ABCFS, Gillian Dite for database management, and Sarah Steinborner and Deon Venter for preparation of DNA. Collection of ABCFS control DNA was funded by the NHMRC. This research was supported by a grant from the Susan G. Komen Breast Cancer Foundation, the NHMRC, and the INHERIT BRCA programme from the Canadian Institute for Health Research. ABS is funded by an NHMRC Career Development Award. JH, GC-T and BRH are NHMRC Senior Principal, Principal and Senior Research Fellows, and WA is funded by a NHMRC Dora Lush postgraduate scholarship.

Competing Interests: The authors declare they have no competing interests

Figure Legends

Figure 1. Stability of the transcripts derived from the genes encoding BRCA1 variants G1706A and A1708E in carriers from Australian pedigrees. SNUPE assays were carried out on homozygous wild-type (wt) and variant control plasmid PCR products, and RT-PCR products from RNA isolated from LCLs generated from BRCA1 variant carriers. (A) G1706A SNUPE assays in forward and reverse directions show approximately equivalent stability of transcripts containing the wt (G) and variant (C) nucleotides at nt5236 (B) A1708E SNUPE assays in forward and reverse directions show approximately equivalent stability of transcripts expressing the wild-type (C) and variant (A) nucleotides at nt5242 of the *BRCA1* mRNA.

Figure 2. The BRCA1 variant 1708E is significantly more susceptible to proteolytic digestion than the wild-type or 1706A isoforms of BRCA1. *BRCA1* cDNA fragments (wild-type, 1706A, 1708E) were translated *in vitro* and digested with increasing concentration of trypsin.

Figure 3. The BRCA1 1708E variant is mislocalised in MCF7 cells. MCF7 cells were transiently co-transfected with myc-tagged *BRCA1* cDNA (wt = wild-type, 1706v = 1706A, 1708v = 1708E) and YFP or YFP-BARD1 expression constructs. Ectopic BRCA1 expression was analysed by c-myc immunofluorescence.

Figure 4. Transcriptional transactivation of reporter plasmids *in vitro* by 1708E BRCA1 is defective in a mammalian system. (A) Mammalian MCF7 cells co-transfected with *BRCA1* expression constructs shows no significant difference in reporter activity between the cells transfected with the wt*BRCA1* or the cDNA encoding the 1706A variant BRCA1, but a significant decrease with the 1708E variant. Values are the mean and standard deviations of three independent experiments (B) Western Blot expression analysis of MCF7 cells transiently transfected with *BRCA1* cDNA constructs and probed with FLAG antibody. Arrow denotes BRCA1 (wt or mutants)-Gal4DBD fusion protein. Protein levels are normalised for even loading. Markers shown are in kD.

Figure 5. Increased centrosome amplification is observed in 293T cells transiently transfected with the cDNA encoding the BRCA1 1708E variant. (A) 293T cells transfected with *BRCA1* expression constructs containing the 1706A and 1708E variants and the 5382insC deletion mutant were stained with c-myc to identify cellular localisation of the ectopic constructs and pericentrin to identify centrioles. Scale bar represents 10nm. Results show exclusion of 1708E from the nucleus. (B) The number of 293T cells containing more than 2 centrosomes indicating centrosome amplification was counted for each of the transient transfections of the various expression plasmids. Negative control is vector and wild-type *BRCA1*. Positive control is *BRCA1* 5382insC. Results show a significant increase in the number of cells showing centrosome amplification in the 1708E transfection.

References

1. Duan J, Wainwright MS, Comeron JM, *et al.* Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum Mol Genet* 2003;**12**(3):205-216.
2. Vega A, Campos B, Bressac-De-Paillerets B, *et al.* The R71G BRCA1 is a founder Spanish mutation and leads to aberrant splicing of the transcript. *Hum Mutat* 2001;**17**(6):520-1.
3. Jung CH, Na YR, Im H. Retarded protein folding of deficient human alpha 1-antitrypsin D256V and L41P variants. *Protein Sci* 2004;**13**(3):694-702.
4. Kariola R, Hampel H, Frankel WL, *et al.* MSH6 missense mutations are often associated with no or low cancer susceptibility. *Br J Cancer* 2004;**91**(7):1287-92.
5. Hofstra RM, Mulder IM, Vossen R, *et al.* DGGE-based whole-gene mutation scanning of the dystrophin gene in Duchenne and Becker muscular dystrophy patients. *Hum Mutat* 2004;**23**(1):57-66.
6. Tzetis M, Efthymiadou A, Strofalis S, *et al.* CFTR gene mutations--including three novel nucleotide substitutions--and haplotype background in patients with asthma, disseminated bronchiectasis and chronic obstructive pulmonary disease. *Hum Genet* 2001;**108**(3):216-21.
7. Rosenberg EH, Almeida LS, Kleefstra T, *et al.* High prevalence of SLC6A8 deficiency in X-linked mental retardation. *Am J Hum Genet* 2004;**75**(1):97-105.
8. Goldgar DE, Easton DF, Deffenbaugh AM, *et al.* Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to BRCA1 and BRCA2. *Am J Hum Genet* 2004;**75**(4):535-44.
9. Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 2003;**73**(3):652-5.
10. Rosen EM, Fan S, Pestell RG, *et al.* BRCA1 gene in breast cancer. *J Cell Physiol* 2003;**196**(1):19-41.
11. Peto J, Collins N, Barfoot R, *et al.* Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst* 1999;**91**(11):943-9.
12. Kawahara M, Sakayori M, Shiraishi K, *et al.* Identification and evaluation of 55 genetic variations in the BRCA1 and the BRCA2 genes of patients from 50 Japanese breast cancer families. *J Hum Genet* 2004.
13. Valarmathi MT, A A, Deo SS, *et al.* BRCA1 germline mutations in Indian familial breast cancer. *Hum Mutat* 2003;**21**(1):98-9.
14. Valarmathi MT, Sawhney M, Deo SS, *et al.* Novel germline mutations in the BRCA1 and BRCA2 genes in Indian breast and breast-ovarian cancer families. *Hum Mutat* 2004;**23**(2):205.
15. Diez O, Osorio A, Duran M, *et al.* Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: A high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat* 2003;**22**(4):301-12.
16. Deffenbaugh AM, Frank TS, Hoffman M, *et al.* Characterization of common BRCA1 and BRCA2 variants. *Genet Test* 2002;**6**(2):119-21.
17. Coyne RS, McDonald HB, Edgemon K, *et al.* Functional Characterization of BRCA1 Sequence Variants Using a Yeast Small Colony Phenotype Assay. *Cancer Biol Ther* 2004;**3**(5).
18. Trenz K, Landgraf J, Speit G. Mutagen sensitivity of human lymphoblastoid cells with a BRCA1 mutation. *Breast Cancer Res Treat* 2003;**78**(1):69-79.

19. Sakayori M, Kawahara M, Shiraishi K, *et al.* Evaluation of the diagnostic accuracy of the stop codon (SC) assay for identifying protein-truncating mutations in the BRCA1 and BRCA2 genes in familial breast cancer. *J Hum Genet* 2003;**48**(3):130-7.
20. Vallon-Christersson J, Cayan C, Haraldsson K, *et al.* Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. *Hum Mol Genet* 2001;**10**(4):353-60.
21. Rodriguez JA, Au WW, Henderson BR. Cytoplasmic mislocalization of BRCA1 caused by cancer-associated mutations in the BRCT domain. *Exp Cell Res* 2004;**293**(1):14-21.
22. Kawai H, Li H, Chun P, *et al.* Direct interaction between BRCA1 and the estrogen receptor regulates vascular endothelial growth factor (VEGF) transcription and secretion in breast cancer cells. *Oncogene* 2002;**21**(50):7730-9.
23. Li H, Lee TH, Avraham H. A novel tricomplex of BRCA1, Nmi, and c-Myc inhibits c-Myc-induced human telomerase reverse transcriptase gene (hTERT) promoter activity in breast cancer. *J Biol Chem* 2002;**277**(23):20965-73.
24. Williams RS, Glover JN. Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. *J Biol Chem* 2003;**278**(4):2630-5.
25. Perrin-Vidoz L, Sinilnikova OM, Stoppa-Lyonnet D, *et al.* The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. *Hum Mol Genet* 2002;**11**(23):2805-14.
26. Mazoyer S, Puget N, Perrin-Vidoz L, *et al.* A BRCA1 nonsense mutation causes exon skipping. *Am J Hum Genet* 1998;**62**(3):713-5.
27. Keaton JC, Nielsen DR, Hendrickson BC, *et al.* A Biochemical analysis demonstrates that the BRCA1 intronic variant IVS10-2A-->C is a mutation. *J Hum Genet* 2003.
28. Pyne MT, Pruss D, Ward BE, *et al.* A characterization of genetic variants in BRCA1 intron 8 identifies a mutation and a polymorphism. *Mutat Res* 1999;**406**(2-4):101-7.
29. Worley T, Vallon-Christersson J, Billack B, *et al.* A naturally occurring allele of BRCA1 coding for a temperature-sensitive mutant protein. *Cancer Biol Ther* 2002;**1**(5):497-501.
30. Buchholz TA, Wu X, Hussain A, *et al.* Evidence of haplotype insufficiency in human cells containing a germline mutation in BRCA1 or BRCA2. *Int J Cancer* 2002;**97**(5):557-61.
31. Williams RS, Chasman DI, Hau DD, *et al.* Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations. *J Biol Chem* 2003;**278**(52):53007-16.
32. Williams RS, Lee MS, Hau DD, *et al.* Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. *Nat Struct Mol Biol* 2004;**11**(6):519-25.
33. Clapperton JA, Manke IA, Lowery DM, *et al.* Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. *Nat Struct Mol Biol* 2004;**11**(6):512-8.
34. Campos B, Diez O, Domenech M, *et al.* RNA analysis of eight BRCA1 and BRCA2 unclassified variants identified in breast/ovarian cancer families from Spain. *Hum Mutat* 2003;**22**(4):337.
35. Phelan CM, Dapic V, Tice B, *et al.* Classification of BRCA1 missense variants of unknown clinical significance. *J Med Genet* 2005;**42**(2):138-46.
36. Figge MA, Blankenship L. Missense mutations in the BRCT domain of BRCA-1 from high-risk women frequently perturb strongly hydrophobic amino acids conserved among mammals. *Cancer Epidemiol Biomarkers Prev* 2004;**13**(6):1037-41.

37. Abkevich V, Zharkikh A, Deffenbaugh AM, *et al.* Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation. *J Med Genet* 2004;**41**(7):492-507.
38. Mirkovic N, Marti-Renom MA, Weber BL, *et al.* Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposition. *Cancer Res* 2004;**64**(11):3790-7.
39. Scott CL, Jenkins MA, Southey MC, *et al.* Average age-specific cumulative risk of breast cancer according to type and site of germline mutations in BRCA1 and BRCA2 estimated from multiple-case breast cancer families attending Australian family cancer clinics. *Hum Genet* 2003;**112**(5-6):542-51.
40. Lange K, Weeks D, Boehnke M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol* 1988;**5**(6):471-2.
41. Hopper JL, Chenevix-Trench G, Jolley DJ, *et al.* Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS). *J Natl Cancer Inst Monogr* 1999;**26**:95-100.
42. Miller J, Jenny A, Rhyan J, *et al.* Detection of Mycobacterium bovis in formalin-fixed, paraffin-embedded tissues of cattle and elk by PCR amplification of an IS6110 sequence specific for Mycobacterium tuberculosis complex organisms. *J Vet Diagn Invest* 1997;**9**(3):244-9.
43. Singer-Sam J, LeBon JM, Dai A, *et al.* A sensitive, quantitative assay for measurement of allele-specific transcripts differing by a single nucleotide. *PCR Methods Appl* 1992;**1**(3):160-3.
44. Poirot O, Suhre K, Abergel C, *et al.* 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Res* 2004;**32**(Web Server issue):W37-40.
45. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* 1974;**185**(4154):862-4.
46. Williams RS, Green R, Glover JN. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nat Struct Biol* 2001;**8**(10):838-42.
47. Osorio A, De La Hoya M, Rodriguez-Lopez R, *et al.* Over-representation of two specific haplotypes among chromosomes harbouring BRCA1 mutations. *Eur J Hum Genet* 2003;**11**(6):489-92.
48. Sum EY, Peng B, Yu X, *et al.* The LIM domain protein LMO4 interacts with the cofactor CtIP and the tumor suppressor BRCA1 and inhibits BRCA1 activity. *J Biol Chem* 2002;**277**(10):7849-56.
49. Lakhani SR, Jacquemier J, Sloane JP, *et al.* Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst* 1998;**90**(15):1138-45.
50. Armes JE, Egan AJ, Southey MC, *et al.* The histologic phenotypes of breast carcinoma occurring before age 40 years in women with and without BRCA1 or BRCA2 germline mutations: a population-based study. *Cancer* 1998;**83**(11):2335-45.
51. Eisinger F, Jacquemier J, Charpin C, *et al.* Mutations at BRCA1: the medullary breast carcinoma revisited. *Cancer Res* 1998;**58**(8):1588-92.
52. Marcus JN, Watson P, Page DL, *et al.* Hereditary breast cancer: pathobiology, prognosis, and BRCA1 and BRCA2 gene linkage. *Cancer* 1996;**77**(4):697-709.
53. Lakhani SR, Gusterson BA, Jacquemier J, *et al.* The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. *Clin Cancer Res* 2000;**6**(3):782-9.

54. Chappuis PO, Nethercot V, Foulkes WD. Clinico-pathological characteristics of BRCA1- and BRCA2-related breast cancer. *Semin Surg Oncol* 2000;**18**(4):287-95.
55. Starita LM, Machida Y, Sankaran S, *et al.* BRCA1-dependent ubiquitination of gamma-tubulin regulates centrosome number. *Mol Cell Biol* 2004;**24**(19):8457-66.
56. Rodriguez JA, Schuchner S, Au WW, *et al.* Nuclear-cytoplasmic shuttling of BARD1 contributes to its proapoptotic activity and is regulated by dimerization with BRCA1. *Oncogene* 2004;**23**(10):1809-20.
57. Fabbro M, Rodriguez JA, Baer R, *et al.* BARD1 induces BRCA1 intranuclear foci formation by increasing RING-dependent BRCA1 nuclear import and inhibiting BRCA1 nuclear export. *J Biol Chem* 2002;**277**(24):21315-24.
58. Au WW, Henderson BR. The BRCA1 RING and BRCT domains cooperate in targeting BRCA1 to ionizing radiation-induced nuclear foci. *J Biol Chem* 2004.
59. Ostrow KL, McGuire V, Whittemore AS, *et al.* The effects of BRCA1 missense variants V1804D and M1628T on transcriptional activity. *Cancer Genet Cytogenet* 2004;**153**(2):177-80.
60. Monteiro AN, August A, Hanafusa H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc Natl Acad Sci U S A* 1996;**93**(24):13595-9.
61. Lingle WL, Barrett SL, Negron VC, *et al.* Centrosome amplification drives chromosomal instability in breast tumor development. *Proc Natl Acad Sci U S A* 2002;**99**(4):1978-83.
62. D'Assoro AB, Barrett SL, Folk C, *et al.* Amplified centrosomes in breast cancer: a potential indicator of tumor aggressiveness. *Breast Cancer Res Treat* 2002;**75**(1):25-34.
63. Hsu LC, Doan TP, White RL. Identification of a gamma-tubulin-binding domain in BRCA1. *Cancer Res* 2001;**61**(21):7713-8.
64. Courjal F, Theillet C. Comparative genomic hybridization analysis of breast tumors with predetermined profiles of DNA amplification. *Cancer Res* 1997;**57**(19):4368-77.
65. Forozan F, Mahlamaki EH, Monni O, *et al.* Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data. *Cancer Res* 2000;**60**(16):4519-25.

Figures for J.Med genet M/S

Figure 1 SNUPE Assay TIFF

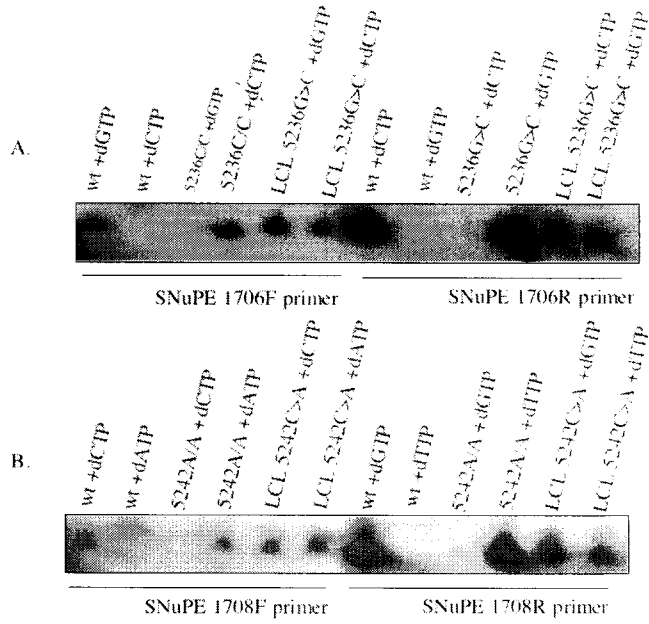


Figure 2 Proteolytic digestion TIFF

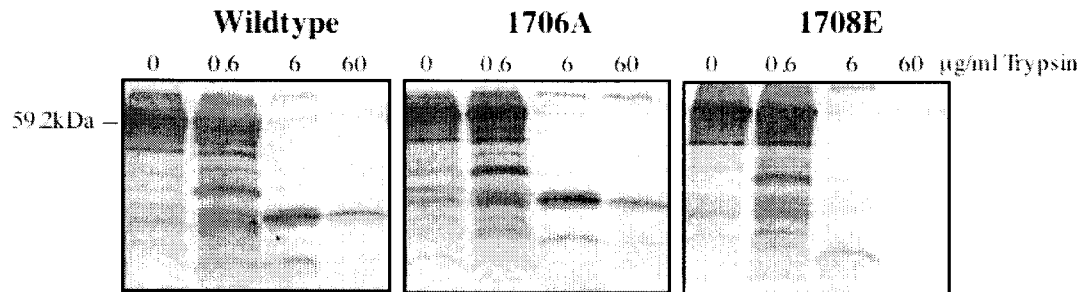


Figure 3 Cytoplasmic localisation in MCF7 cells TIFF

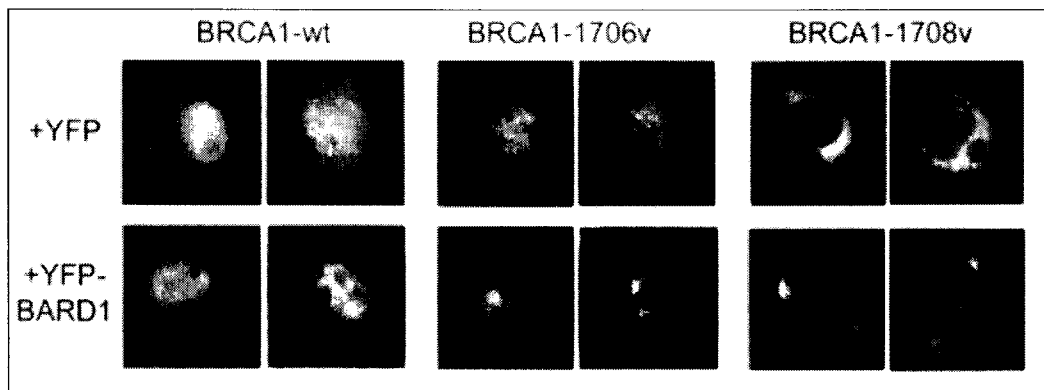


Figure 4. Mammalian transcriptional Activation assay TIFF

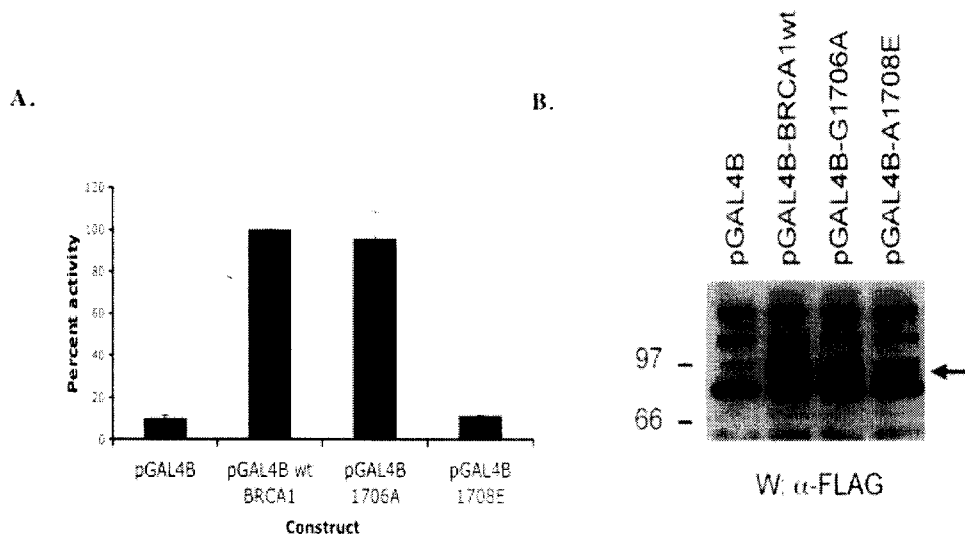


Figure 5 Centrosome amplification TIFF

