

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

**Analysis of *BRCA1* Genomic Structure:
Novel germline mutations and somatic alterations in breast cancer.**

by

Shannon Renée Payne

**A dissertation submitted in partial fulfillment of the
requirements for the degree of**

Doctor of Philosophy

University of Washington

2000

Program Authorized to Offer Degree: Department of Genetics

UMI Number: 9964282

Copyright 2000 by
Payne, Shannon Renee

All rights reserved.

UMI[®]

UMI Microform 9964282

Copyright 2000 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright 2000

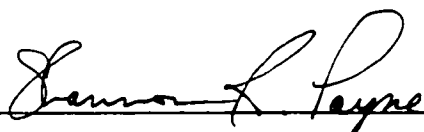
Shannon Renée Payne

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington,

I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Bell and Howell Information and Learning,

300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microfilm form and/or (b) printed copies of the manuscript made from microform."

Signature 

Date January 28, 2000

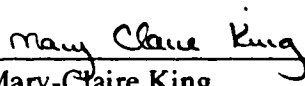
University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Shannon Renée Payne

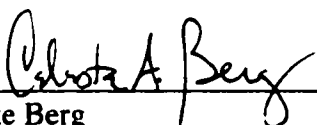
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

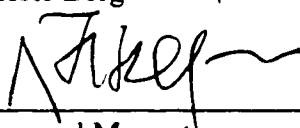


Mary-Claire King

Reading Committee:



Celeste Berg



Raymond Monnat

Date: 3/4/00

University of Washington

Abstract

Analysis of *BRCA1* Genomic Structure:
Novel germline mutations and somatic alterations in breast cancer

by Shannon Renée Payne

Chairperson of the Supervisory Committee

Professor Mary-Claire King

Department of Genetics and Division of Medical Genetics

Germline mutations in *BRCA1* lead to an increased risk of breast and ovarian cancer, with loss of the second, normal allele critical to tumorigenesis. The relative lack of somatic mutations in *BRCA1*, however, has argued against its involvement in non-inherited (sporadic) breast cancer. One explanation for this contradiction is that *BRCA1* mutations exist in sporadic breast carcinomas, but are not identified by current mutation screening procedures. My research specifically addressed the types and frequencies of large genomic rearrangements responsible for inactivation of *BRCA1*. I characterized the types of large rearrangements that occur within the *BRCA1* region and investigated the contribution of two large germline rearrangements to breast cancer in a population-based series of breast cancer patients. Although the structure of the *BRCA1* genomic region was generally well-conserved, one variant allele containing multiple large alterations of the *BRCA1* genomic region was identified. The existence of a variant with multiple large rearrangements in *cis* indicates that an investigation of the

types and frequencies of noncoding variation in the *BRCA1* genomic region may yield a broader understanding of noncoding variation within the human genome.

One argument for the role of BRCA1 in sporadic breast cancer has been the high rate of allelic loss observed for the *BRCA1* locus. In order to determine whether BRCA1 is inactivated somatically by large rearrangements of *BRCA1*, I analyzed 92 breast carcinomas for genomic loss in the *BRCA1* region of chromosome 17q. I investigated genomic loss using a combination of loss of heterozygosity (LOH), Long PCR, and Southern analysis. Although two large germline rearrangements were detected in our series, no large somatic rearrangements were identified. LOH results were correlated with BRCA1 protein immunohistochemistry data generated by Rachel Gonzalez-Hernandez in order to test whether LOH is a mechanism for inactivating BRCA1 in sporadic breast cancer. Reduced BRCA1 protein in sporadic breast carcinomas was associated significantly with loss of the most 5' *BRCA1* intragenic marker, D17S1323. LOH at the more 3' *BRCA1* intragenic markers was not associated with reduced BRCA1 protein. Interestingly, 8 of 14 breast carcinomas retaining all three *BRCA1* intragenic markers showed reduced BRCA1 protein. Thus, there are likely other mechanisms for inactivation of BRCA1 in sporadic breast cancer.

TABLE OF CONTENTS

List of Figures.....	iii
List of Tables.....	iv
Chapter One: BRCA1: The Early Years.....	1
Cloning and sequence characterization of <i>BRCA1</i>	2
Functional characterization of <i>BRCA1</i>	4
Tumor suppression and <i>BRCA1</i>	6
Germline and somatic alterations in <i>BRCA1</i>	7
Project goals.....	12
Chapter Two: Both Homologous and Nonhomologous Mechanisms	
Generate Large Germline Rearrangements of <i>BRCA1</i>	15
Experimental logic and chapter contributions.....	16
Materials and methods.....	17
Results and discussion.....	22
Chapter Three: Somatic Alterations of <i>BRCA1</i> in Breast Cancer: Loss of	
Heterozygosity, but No Large Rearrangements.....	27
Experimental logic and chapter contributions.....	28
Materials and methods.....	29
Results.....	31
Discussion.....	41
Chapter Four: Characterization of a Variant <i>BRCA1</i> Allele.....	50
Experimental logic and chapter contributions.....	51
Materials and methods.....	51
Results.....	53
Discussion.....	64

TABLE OF CONTENTS (cont.)

Chapter Five: Summary and Future Directions.....69

 Large germline rearrangements of *BRCA1*.....70

 Somatic inactivation of *BRCA1*.....71

 Noncoding variation within the *BRCA1* genomic region.....73

List of References.....76

Appendix One: “Complex germline rearrangement of *BRCA1* associated
with breast and ovarian cancer.”.....87

LIST OF FIGURES

Number

1.1	BRCA1 protein domains.....	3
1.2	Genomic region between the <i>BRCA1</i> and <i>NBR1</i> genes.....	13
2.1	Detection of the g.43368ins6081 mutation.....	25
3.1	Microsatellite markers within the <i>BRCA1</i> region.....	32
3.2	Southern analysis of <i>BRCA1</i> in primary breast carcinomas.....	39
3.3	<i>BRCA1</i> region of chromosome 17q21.....	40
4.1	977 bp germline deletion (g.56024del977).....	55
4.2	Analysis of RT-PCR products from Case 72.....	56
4.3	Southern analysis of the <i>BRCA1</i> exon 8 genomic region.....	57
4.4	PCR analysis of a minimal genomic region for the Case 72 insertion.....	58
4.5	BRCA1 intron 7 compound microsatellite.....	60
4.6	Single nucleotide polymorphisms identified in Case 72 variant allele.....	61
4.7	24 bp insertion in <i>BRCA1</i> intron 15.....	63
4.8	Loss of transcript analysis in cDNA from Case 72.....	65

LIST OF TABLES

Number

1.1	Somatic mutations in <i>BRCA1</i>	9
1.2	Characterized large germline rearrangements of <i>BRCA1</i>	10
2.1	Methods for detection of large rearrangements in <i>BRCA1</i>	18
2.2	Long PCR primers used for analysis of <i>BRCA1</i>	20
2.3	Restriction site variants identified in <i>BRCA1</i> from families.....	24
3.1	Genomic loss in primary breast carcinomas.....	33
3.2	Restriction site variants identified in <i>BRCA1</i> from breast cancer cases.....	38
3.3	Marker D17S1323 LOH and BRCA1 protein expression.....	42
3.4	Genomic loss and BRCA1 protein expression.....	43

ACKNOWLEDGEMENTS

My whole-hearted debt of gratitude goes to the many people who were involved in the work that I describe in this dissertation. The immunohistochemistry work carried out by Rachel Gonzalez-Hernandez was beautiful and crucial in its relationship to my research. I would also like to thank the many breast cancer patients who, in their own moments of deepest uncertainty, were generous enough to share themselves with researchers.

The work of my committee members (Mary-Claire King, Celeste Berg, Clem Furlong, and Raymond Monnat) on my behalf cannot be understated. Mary-Claire introduced me to some of my best friends and I am eternally grateful to her ability to gather such wonderful people together in a single lab. Celeste happily mentored me throughout graduate school, two different departments, and two different advisors. Our coffee breaks together kept me sane. Clem taught me that it is possible to lead a balanced life and shared with me many wonderful stories that helped me maintain perspective. Ray gave me the joy of sharing my work with someone who reveled in the same aspects of my research as I did. This group would not be complete without mention of my Botany department advisor, Luca Comai. The further I go in scientific research, the more I realize that I owe to him the kinds of questions that I ask and the ways in which I ask them.

I am deeply indebted to all the members of the King Lab, both past and present. Elizabeth Schubert was my “baymate” for many years and, try as she might, she cannot seem to escape me as I will soon be working just the down the hall from her again. I am grateful for all the sage advice that she has given me in the past and continues to give me. Piri Welsh has helped me to “keep my cool” and to attend some of the best parties of my life. Michele Harvey-Blankenship truly wants to make the world a better place. I say she has already done so. Heather Christy Mefford has the gift of instinctively knowing when I need to go for a very long bike ride and making sure that I

take it. She's been a good friend and a good riding partner. Rachel Gonzalez-Hernandez taught me to always look for the "good pearl" in each day. Thank you all.

My acknowledgements would not be complete without thanking my family: my mother, my father, my sister Jenny, and her two children. No matter what fool thing grabbed my attention, my parents always supported me 100%. They supported me not because they understood what I was doing, but because they trusted me without question to make the best judgement. My sister has shown more courage in the past year than most people must show in an entire lifetime. She is an inspiration to me and, I am sure, to her children (Chad and Reyna). If I ever manage to accomplish anything in my life, it is because my family taught me how.

Finally, I want to thank Roddy Grant, my best friend and my husband. He introduced me to the world of cycling, scotch, and Freehand. Most importantly, however, he made me smile when I needed to smile. Graduate school was more bearable because I could go home to him.

DEDICATION

The work described herein is dedicated to Archibald Cox, who followed the path less traveled, and to all those who follow that path with him. Remember to live your life in such a way as to make the comments of your detractors seem as false as they are.

CHAPTER ONE

BRCA1: The Early Years

Cloning and sequence characterization of *BRCA1*

Breast cancer is the most prevalent malignancy among Western European women and is second only to lung cancer in its mortality rate (ACS, 1997). Approximately 5-10% of breast cancer is attributable to inherited mutations, although the exact number is population-dependent (Newman et al., 1988; Szabo & King, 1997; Newman et al., 1998). For other human diseases such as colorectal cancer and melanoma, relatively rare cancer susceptibility syndromes have yielded a critical foothold into the understanding of these diseases. Thus, it was widely presumed that the cloning and characterization of genes involved in hereditary breast and ovarian cancer would lead to a better understanding of the more common non-inherited (sporadic) forms of breast cancer.

The breast cancer susceptibility gene, *BRCA1*, was mapped to chromosome 17q21 in 1990 and in 1994, *BRCA1* was cloned (Hall et al., 1990; Miki et al., 1994). The major *BRCA1* message encodes a protein of 1863 amino acids, although numerous alternatively spliced products that maintain the *BRCA1* reading frame have been identified (Xu, 1997; Lu et al., 1996). In most cases, the function(s) of the alternatively spliced product is unclear. It has been suggested that one alternatively spliced product that is missing the majority of exon 11 is localized exclusively to the cytoplasm and may have a role distinct from that of full-length *BRCA1* in cell growth and tumorigenesis (Wilson et al., 1997).

The amino acid sequence of *BRCA1* revealed few clues as to its function. No homology with any known protein was identified with the exception of a RING finger domain (Figure 1.1). The NH₂ terminus of *BRCA1* contains a pattern of cysteine and histidine residues (aa 24-64) found in members of the RING finger family, a sub-family of the zinc finger proteins (Miki et al., 1994; Lovering et al., 1993). The RING finger is a Zn²⁺-binding motif that is found in a diverse group of proteins and often mediates either protein-DNA or protein-protein interactions (Freemont, 1993; Saurin et al., 1996). The RING finger domain of *BRCA1* contains the only consensus missense mutations leading to breast and ovarian cancer predisposition, C61G and C64G (Castilla et al.,

Figure 1.1

BRCA1
1863 aa

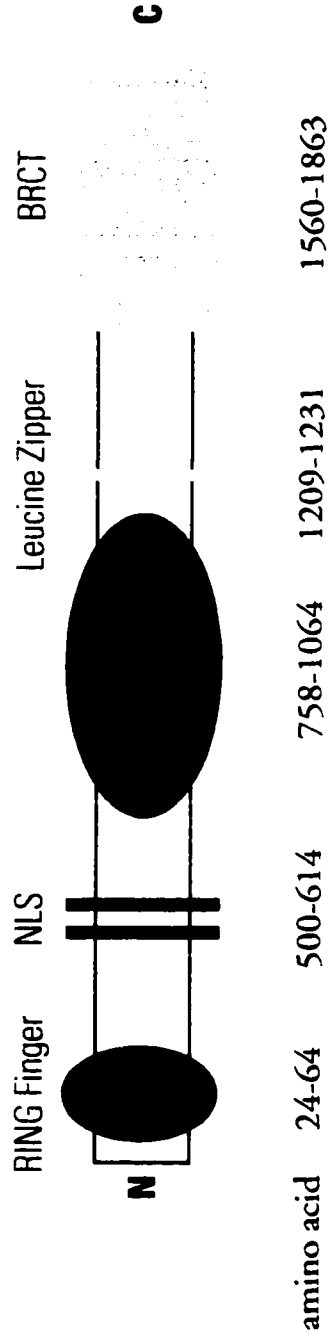


Figure 1.1 Schematic diagram of domains in the BRCA1 protein. The amino acids corresponding to each domain are indicated below as are the functions of the domain when these are known.

1994; Friedman et al., 1994). The naturally occurring missense mutations in this region disrupt BRCA1 homodimerization and heterodimerization with BARD1 (BRCA1-Associated RING Domain protein 1) (Brzovic et al., 1998; Meza et al., 1999).

Although several RING finger proteins interact directly with DNA through the RING finger domain, there is no evidence for binding of DNA by either the BRCA1 homodimer or heterodimer (Meza et al., 1999).

- Other recognizable features of the BRCA1 amino acid sequence include two putative nuclear localization signals (aa 500-508 and aa 609-615), a leucine zipper (aa 1209-1231), and an excess of negatively charged residues in the COOH-terminus. The presence of an acidic region in the COOH-terminus correlates with a transactivation domain in many eukaryotic transcription activators. Accordingly, the BRCA1 COOH-terminus is able to transactivate reporter constructs when fused to a heterologous DNA-binding domain (Chapman & Verma, 1996; Monteiro et al., 1996).

Finally, a novel protein motif, designated the BRCT (BRCA1 COOH-Terminal) domain, was identified through analysis of a repeated motif in the BRCA1 protein and comparison to other proteins, including p53 binding protein (*53BP1*) and the yeast RAD9 protein (Koonin et al., 1996). The BRCT domain is present in many proteins known to be involved in DNA repair including XRCC1, RAD4, REV1, Crb2, RAP1 and several eukaryotic DNA ligases (Callebaut & Morion, 1997).

Functional characterization of BRCA1

Potential roles for BRCA1 in DNA damage response and in transcription activation have been proposed. Several lines of evidence argue BRCA1 is involved in the cellular response to DNA damage. The presence of the BRCT motif in the COOH-terminus of BRCA1 suggests a role for BRCA1 in DNA damage repair pathways. BRCA1 and human RAD51 colocalize in discrete nuclear foci during S phase of the cell cycle and relocate to sites of nonduplex DNA structure in response to treatment with hydroxyurea or ultraviolet light (Scully et al., 1997a ; Scully et al., 1997b). Deletion mapping indicated that amino acids 758-1064, encoded by exon 11 of *BRCA1*, mediated

the interaction with RAD51. Additionally, BRCA1 undergoes hyperphosphorylation in response to DNA damaging agents (Scully et al., 1997b).

The most convincing evidence that BRCA1 is involved in the cellular response to DNA damage was provided by mouse embryonic stem cells deficient in BRCA1. These cells are defective in the ability to carry out transcription-coupled repair of oxidative DNA damage and are hypersensitive to ionizing radiation and hydrogen peroxide (Gowen et al., 1998). Transcription-coupled repair is a process in which DNA damage is repaired more rapidly in transcriptionally active DNA than in the genome as a whole (Mellon et al., 1987; Hanawalt, 1994). The accelerated rate of repair is due to faster repair of lesions in the transcribed strand than in the non-transcribed strand and requires an active RNA polymerase II complex (Leadon & Lawrence, 1992; Christians & Hanawalt, 1992).

BRCA1 is linked to the RNA polymerase II holoenzyme via RNA helicase A (Scully et al., 1997c; Anderson et al., 1998). The presence of BRCA1 in the RNA polymerase II holoenzyme complex integrates the two primary functions proposed for BRCA1: response to DNA damage via transcription-coupled repair and transcription activation. *In vitro* evidence of a transactivating role for BRCA1 was first provided by experiments in which the BRCA1 COOH-terminus was fused to the GAL4 DNA-binding domain and was able to transactivate a variety of reporter constructs (Chapman & Verma, 1996; Monteiro et al., 1996). *In vivo*, full-length BRCA1 is able to induce transcription from the promoter of the cyclin-dependent kinase inhibitor p21^{WAF1/CIP1} when transiently transfected into human cancer cells (SW480 cells) (Somasundaram et al., 1997). It is not clear, however, whether the induction of p21 in this model represents a direct or indirect response to the presence of BRCA1 protein.

Other evidence suggests that BRCA1 may have a global role in DNA and RNA metabolism through chromatin remodeling. BRCA1 interacts with the HDAC1 and HDAC2 complexes through the BRCT domain both *in vitro* and *in vivo* (Yarden & Brody, 1999). The HDAC1 and HDAC2 complexes are involved in the establishment of transcriptionally silenced chromatin by deacetylating the nucleosomal histones. Deacetylation appears to facilitate tighter interaction between DNA and nucleosomes by

unmasking the positively charged lysine residues at the histone NH₂-termini. The tighter nucleosome interaction hinders access of transcription factors to DNA regulatory elements.

Additionally, the acidic trans-activation domain of BRCA1, when tethered to a GAL4 DNA binding domain, alters local chromatin structure and stimulates chromosomal DNA replication *in vivo* in *Saccharomyces cerevisiae* (Hu et al., 1999). Stimulation of eukaryotic DNA replication by transcription factors that bind DNA near an origin of replication is well-documented (Van der Vliet, 1996). It has been suggested that the chromatin remodeling complex responsible for mediating transcription activation may also be involved in activation of DNA replication (Hu et al., 1999). Thus, BRCA1 may represent an important link for understanding the relationship between chromatin remodeling, cell cycle progression, and tumorigenesis.

Tumor suppression and *BRCA1*

Germline mutations in *BRCA1* lead to an increased risk of breast and ovarian cancer, with loss of the second, wild-type allele critical to tumorigenesis (Merajver et al., 1995). Even before *BRCA1* was cloned, it was generally accepted that BRCA1 functioned as a tumor suppressor because in tumors from *BRCA1*-mutation carriers, loss of heterozygosity (LOH) was always observed for the unlinked, “normal” chromosome (Smith et al., 1992).

LOH including the *BRCA1* region occurs frequently in sporadic breast and ovarian cancer (Bieche & Lidereau, 1995; Devilee & Cornelisse, 1994). Anywhere from 30-70% of breast tumors show LOH of the *BRCA1* region. Thus, it was thought that BRCA1 would function as a tumor suppressor in sporadic cancer as well. The role of BRCA1 in sporadic breast cancer has been controversial due to the apparent absence of somatic mutations (Futreal et al., 1994; Matsushima et al., 1995; Takahashi et al., 1995; Hosking et al., 1995; Merajver et al., 1995; Berchuck et al., 1998).

In several ways, *BRCA1* fits the classic tumor suppressor model as defined by *TP53*, *APC*, *RBI*, *CDKN2A*, and others. Antisense inhibition of *BRCA1* accelerates

mammary epithelial cell growth in culture (Thompson et al., 1995). Wild-type *BRCA1* message is frequently decreased in sporadic breast tumors as compared to normal breast tissue (Thompson et al., 1995; Rice et al., 1998; Magdinier et al., 1998). Finally, evaluation of sporadic breast tumors by immunohistochemistry reveals that expression of BRCA1 protein is reduced or lost in most breast tumors of high histologic grade (Wilson et al., 1999; Jarvis et al., 1998; Taylor et al., 1998; Lee et al., 1999).

Germline and somatic alterations in *BRCA1*

The *BRCA1* gene encodes a 7.8 kilobase pair (kb) mature message that is divided among 24 exons located over a more than 80 kb genomic region (Smith et al., 1996). More than 400 distinct *BRCA1* germline mutations have been identified to date (BIC, 1999). Although ~30% of distinct *BRCA1* mutations have been observed more than once, the majority of mutations are unique (Szabo & King, 1995). Mutation detection in *BRCA1* is complicated further by the ubiquitous distribution of these mutations throughout the *BRCA1* coding region (BIC, 1999). As sequencing of the entire *BRCA1* genomic region in each sample potentially harboring a *BRCA1* mutation is logistically prohibitive, most investigators have chosen to use a combination of techniques that rely on amplification of small, ~150-700 bp fragments from the *BRCA1* genomic region. However, mutation detection based on amplification of small fragments will not identify large-scale alterations of the *BRCA1* genomic region such as deletions, duplications, or inversions of hundreds of base pairs (bp).

This approach has worked relatively well for germline mutation screening of *BRCA1*. Most germline *BRCA1* mutations are small insertions, deletions, or single bp substitutions that lead to premature protein truncation (BIC, 1999). Based on the identification of these "small" mutations, it is thought that mutations in *BRCA1* and, to a lesser extent *BRCA2*, account for the majority of hereditary breast and ovarian cancer (Rebbeck et al., 1997; Narod et al., 1995). In a study of 48 breast and/or ovarian cancer families, Schubert et al. (1997) encountered 9 families with no detectable mutations in either *BRCA1* or *BRCA2*. Of these 9 families, 2 demonstrated positive LOD scores with

BRCA1. In a similar study of 23 families, Rebbeck et al. (1997) found no detectable mutations in 13 families. In at least 2 of these families, linkage of cancer to *BRCA1* yielded positive LOD scores. It is uncertain how many of the unexplained families in these studies are due to as yet undetected mutations in *BRCA1* or *BRCA2*, and how many are due to mutations in susceptibility genes yet to be identified.

Of the breast and ovarian tumors screened by several independent groups, only a few ovarian tumors with somatic mutation in *BRCA1* have been reported (Table 1.1) (Hosking et al., 1995; Merajver et al., 1995; Berchuck et al., 1998). Thirteen of the fourteen *BRCA1* somatic mutations detected in ovarian tumors were nonsense and frameshift mutations caused by either point mutation or the insertion or deletion of a single bp. The remaining somatic mutation was a single bp substitution in a conserved residue of the *BRCA1* RING finger domain, a missense mutation that has been seen previously as a germline mutation. All fourteen of the tumors harboring these somatic mutations exhibited LOH for the *BRCA1* region of chromosome 17.

Thus, *BRCA1* somatic mutations detectable by conventional PCR-based techniques are extremely rare in breast and ovarian tumors. These conventional screening procedures, however, neglect a class of mutations: large genomic rearrangements. Several large germline deletions and one large germline duplication in *BRCA1* have been reported (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b; Montagna et al., 1999; Carson et al., 1999; Rohlf's et al., 1999). All of the large germline rearrangements were undetectable by conventional genomic DNA (gDNA) screening procedures (Table 1.2). Large germline deletions appear to be a relatively common feature of inherited breast cancer in the Dutch population, due in large part to founder mutations (Petrij-Bosch et al., 1997).

Several features of the *BRCA1* locus provide clues as to the nature of large genomic rearrangements that might occur. *BRCA1* has one of the highest densities of Alu elements of genes deposited in GenBank (41.5% of 81 kb) (Smith et al., 1996). Large genomic rearrangements in human genes frequently are associated with Alu repetitive elements (Purandare & Patel, 1997; Mazzarella & Schlessinger, 1998).

Table 1.1 Somatic Mutations in *BRCA1*

<u>Nt Change^a</u>	<u>Exon</u>	<u>Coding Effect</u>	<u>Effect on Protein</u>	<u>Site^b</u>	<u>Observed^c</u>	<u>Reference</u>
151 ins C	2	frameshift	N16X	ovary	1	Merajver et al., 1995
G ivs2+1 T	ivs 2	splicing	unknown	ovary	1	Berchuck et al., 1998
T 300 G	5	missense	C61G	ovary	1	Merajver et al., 1995
2080 ins A	11	frameshift	K672X	ovary	1	Berchuck et al., 1998
NA ^d	11	nonsense	E730X	ovary	1	Berchuck et al., 1998
2388 del G	11	frameshift	V764X	ovary	2	Berchuck et al., 1998
2569 del G	11	frameshift	I845X	ovary	1	Berchuck et al., 1998
2893 ins A	11	frameshift	K937X	breast	1	Khoo et al., 1999
C 3145 A	11	nonsense	S1009X	ovary	1	Khoo et al., 1999
G 5293 T	19	nonsense	E1725X	ovary	1	Merajver et al., 1995
5502 ins T	22	frameshift	M1827X	ovary	1	Merajver et al., 1995
5559 del G	23	frameshift	V1833X	ovary	1	Hosking et al., 1995
NA	24	nonsense	R1835X	ovary	1	Berchuck et al., 1998

a. Nt Change: The observed nucleotide change with respect to the *BRCA1* cDNA sequence, HSU 14680.

b. Site: The site of the primary tumor in which the indicated somatic mutation was observed.

c. Observed: The number of independent tumors in which the indicated somatic mutation was observed.

d. NA: The exact nucleotide change was not published.

Table 1.2 Characterized Large Germline Rearrangements of *BRCA1*

<u>Characterization</u>	<u>Location</u>	<u>Size</u>	<u>Breakpoints</u>	<u>Effect</u>	<u>Reference</u>
**	<i>NBR2</i> to ivs 2	14 kb	both Alu	no transcript	Swensen et al., 1997
g.12977ins10del1039	exon 3 to ivs 3	1039 bp	exon 3 ivs 3	exon 3 skipped codon 27 stop	Family 5 (Appendix One)
g.26966del23763	ivs 7 to ivs 13	23,763 bp	both Alu	exons 8-13 deleted codon 50 stop	Puget et al., 1999a
g.43368ins6081	ivs 12 to ivs 13	6081 bp	both Alu	exon 13 duplicated codon 1460 stop	Puget et al., 1999b founder mutation
g.44519del3835	ivs 12 to ivs 13	3835 bp	Alu ivs 13	exon 13 deleted codon 1398 stop	Petrij-Bosch et al., 1997 Dutch founder mutation
g.53101del2998	ivs 14 to ivs 15	2998 bp	both Alu	exon 15 deleted codon 1574 stop	Puget et al., 1999a

** breakpoint amplified using primers designed from PAC P1 1141 and *BRCA1* intron 2

*** not fully characterized

Table 1.2 Large Germline Rearrangements of *BRCA1* (cont.)

<u>Characterization</u>	<u>Location</u>	<u>Size</u>	<u>Breakpoints</u>	<u>Effect</u>	<u>Reference</u>
***	ivs 16 to ivs 19 ivs 19 to ivs 20	8216 bp ~3.5 kb	all four Alu	exon 17-19 deleted exon 20 deleted in frame deletion of 97 a.a.	Rohlf's et al., 1999
g.58758del3094	ivs 16 to ivs 17	3094 bp	both Alu	exon 17 deleted codon 1672 stop	Montagna et al., 1999 founder mutation
g.60453del 1008	ivs 16 to ivs 17	1008 bp	both Alu	exon 17 deleted codon 1672 stop	Puget et al., 1997
g.71348del3984	ivs 19 to ivs 20	3984 bp	both Alu	exon 20 deleted in-frame deletion of 28 a.a.	Carson et al., 1999
g.79504del510	ivs 21 to ivs 22	510 bp	ivs 21 Alu	exon 22 deleted codon 1833 stop	Petrij-Bosch et al., 1997 Dutch founder mutation

** breakpoint amplified using primers designed from PAC P1 1141 and *BRCA1* intron 2

*** not fully characterized

Another predisposing feature of the *BRCA1* locus is a tandem duplication involving the 5' region of *BRCA1* (Brown et al., 1996; Barker et al., 1996). The duplication includes the *BRCA1* promoter region and a neighboring gene, *NBR1* (Figure 1.2). The result of the duplication was the creation of a new gene, *NBR2*, with which human *BRCA1* shares a bi-directional promoter and the creation of a pseudo-copy of *BRCA1* situated in a similar head to head orientation with the *NBR1* gene approximately 30 kb upstream of the functional *BRCA1* locus (Xu, 1997). The pseudo-copy of *BRCA1* contains a nonsense mutation in exon 2. Based on Southern analysis, the pseudo-copy of *BRCA1* includes exons 1a, 1b, and 2, but not exons 3, 4, 13, and 21. Other exons were not analyzed. The presence of a pseudo-copy of *BRCA1* with >90% identity located so near on chromosome 17 further predisposes the functional *BRCA1* locus to inactivation by misaligned homologous recombination or gene conversion.

Project goals

Several possible interpretations could explain the paradox presented by the lack of somatic mutations in *BRCA1*. One formal explanation is that *BRCA1* is not involved in sporadic breast and ovarian cancer. It is possible that the high rate of LOH for the *BRCA1* region in sporadic tumors actually targets a nearby gene, and not *BRCA1*.

A second explanation is an alternative mechanism for loss of the second *BRCA1* allele in sporadic tumors. Aberrant patterns of DNA methylation are among the most common genomic alterations seen in cancer (Jones & Laird, 1999, 1996; Baylin et al., 1998). Abnormal methylation in the promoters of well-characterized tumor suppressor genes such as *RBI* and *CDKN2A* can contribute to their functional inactivation (reviewed in Jones & Laird, 1999). Several groups have reported a decrease in the amount of *BRCA1* message detected in breast tumor tissue as compared to normal breast tissue (Thompson et al., 1995; Rice et al., 1998; Magdinier et al., 1998).

To date, analyses of methylation within the *BRCA1* CpG island have produced conflicting results. In one study, the decrease in *BRCA1* expression was associated with aberrant methylation within the *BRCA1* δ promoter (Rice et al., 1998). Mancini et al.

Figure 1.2

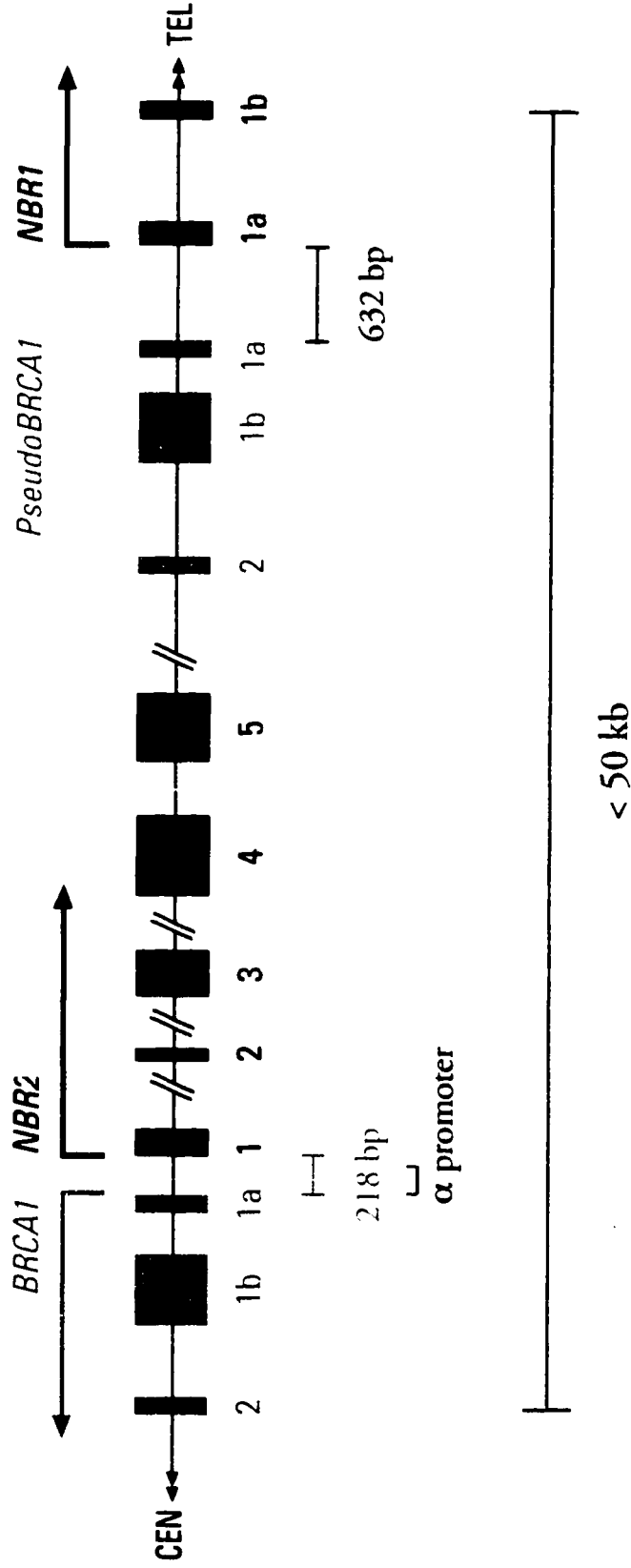


Figure 1.2 Schematic diagram of the genomic region between the *BRCA1* and *NBR1* genes. Arrows represent the direction of transcription and boxes represent exons. Exons and introns without hash bars are drawn to scale. The distance between the first exons of *BRCA1* and *NBR2* is shown in red. The distance between the first exons of pseudo-*BRCA1* and *NBR1* is shown in black. The duplication encompasses the 5' end of both *BRCA1* and *NBR1* and spans a maximum of 50 kb (Brown et al., 1996).

(1998) also detected aberrant methylation within the *BRCA1* δ promoter in four of eleven breast and ovarian tumors. Their study did not analyze *BRCA1* expression. In the largest study of methylation within the *BRCA1* δ promoter published to date, Catteau et al. (1999) detected aberrant methylation in 11 of 96 breast tumors. Magdinier et al. (1998), however, did not detect aberrant methylation in any of 37 breast tumors analyzed.

A third possibility is that the lack of *BRCA1* mutations in sporadic tumors actually reflects limitations in current mutation screening procedures. The spectrum of somatic mutation in *BRCA1* may differ dramatically from the germline mutation spectrum. Given the genomic organization of the *BRCA1* locus and the methods commonly used to screen for *BRCA1* mutations, it is conceivable that *BRCA1* somatic mutations in breast and ovarian cancer exist but have remained undetected by conventional mutation detection methods.

The research detailed in the following chapters further elucidates the role of *BRCA1* in breast cancer by analyzing the genetic alterations that lead to inactivation of *BRCA1*. Chapter Two describes the characterization of large germline rearrangements identified in *BRCA1* and their contribution to breast cancer in a population-based series of breast cancer patients. Chapter Three details the nature of somatic alterations identified from sporadic breast tumors and their relationship to *BRCA1* protein expression. Chapter Four describes a rare variant allele of *BRCA1* and the potential significance of noncoding variation in the *BRCA1* genomic region. Finally, Chapter Five reviews the significant advances in *BRCA1* biology that have occurred over the time that this research was conducted and how this research has contributed to these advances.

CHAPTER TWO

Both Homologous and Nonhomologous Mechanisms Generate Large Germline Rearrangements of *BRCA1*

Experimental Logic and Chapter Contributions

My primary interest has been to understand what role (if any) *BRCA1* plays in the genesis of sporadic breast cancer. Because of my interest in whether large rearrangements of *BRCA1* lead to somatic inactivation of *BRCA1*, I first wanted to survey the types of large rearrangements that can occur in the *BRCA1* genomic region.

I addressed this question by screening for large germline rearrangements in families with inherited breast and/or ovarian cancer in which no mutations in either *BRCA1* or *BRCA2* had been detected previously. These included four families with positive LOD scores to the *BRCA1* region of chromosome 17q21 (Families 41, 48, 58 and 94). Families with multiple cases of breast cancer and with negative LOD scores for linkage to *BRCA1* might nonetheless harbor mutations in *BRCA1* if some cases are sporadic (ie. phenocopies). Seven such families were screened (Families 43, 46, 52, 57, 61, 66, and 110). The characterization of one large germline rearrangement predisposing to breast and ovarian cancer is discussed in detail in Appendix One. Appendix One represents work that has been published previously.

Additionally, I screened for large rearrangements of *BRCA1* in several cell culture lines. If *BRCA1* is inactivated somatically by large rearrangements of *BRCA1*, this would be more easily detected in a clonal population (such a cell culture line) than in a mixed population such as found in tumor tissue. The cell lines analyzed were standard lines used by a number of investigators and available through the American Type Culture Collection. The lines represented a range of phenotypes such as estrogen receptor negative and estrogen receptor positive. They included one normal human mammary epithelial cell line (HMEC2595), one abnormal breast cell line (HBL100), ten breast cancer cell lines (BT20, BT483, HTB24, MDA-MB-231, MDA-MB-468, Hs578T, T47D), three ovarian cancer cell lines (CaOv3, ES2, PA1), and three prostate cancer cell lines (DU145, PC3, PPC1).

I screened for large genomic rearrangements of *BRCA1* in two ways: Long PCR analysis and Southern analysis. Long PCR allows rapid screening of samples using minimal amounts of DNA. Due to the nature of PCR, however, Long PCR will be

biased strongly toward detection of genomic deletions as opposed to genomic duplications. Additionally, very large rearrangements (larger than the size of Long PCR products used) will not be detected by Long PCR. The majority of the published large rearrangements within the *BRCA1* region are in the size range detectable by Long PCR. Most of the characterized large rearrangements in this size range would be detectable by the Long PCR screen designed (Table 2.1) (Puget et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Montagna et al., 1999; Carson et al., 1999; Rohlf's et al., 1999).

Two very large rearrangements of *BRCA1*, deletions of ~14 kb and ~23.8 kb, have been described (Swensen et al., 1997; Puget et al., 1999a). These rearrangements were detected by Southern hybridization. Therefore, I also examined family and cell line DNA by Southern blot analysis. One caveat to Southern analysis is that unambiguous detection of rearrangement requires the probe used for detection to be present in the rearranged allele. Thus, when using a cDNA probe, genomic deletion of a single exon will be detected most readily if that exon resides on a restriction fragment containing other exons. This presents a technical problem for analysis of the *BRCA1* genomic region in which many small exons (14 exons are <100 bp and 4 are <60 bp) are spread over a large genomic region. As a result, some mutations are most easily detected by Long PCR analysis (for example, the deletion involving the 54 bp *BRCA1* exon 3 described in Appendix One).

Finally, in order to begin to ascertain the contribution of large germline rearrangements to breast cancer in the American population, I investigated two large germline rearrangements (one large deletion and one large duplication) in a population-based series of 242 white and 164 African-American breast cancer patients unselected for family history. Patients were ascertained previously as part of the Carolina Breast Cancer Study (Newman et al., 1995; Newman et al., 1998).

Materials and Methods

Long PCR Analysis

Table 2.1 Methods for Detection of Large Rearrangements in *BRCA1*

<u>Characterization</u>	<u>Location</u>	<u>Size</u>	<u>Long PCR?^a</u>	<u>Restriction Digest?^b</u>	<u>Southern?^c</u>
Promoter deletion	<i>NBR2</i> to ivs 2	14 kb del	no	no	yes (PstI)
g.12977ins10del1039	exon 3 to ivs 3	1039 bp del	yes	yes	no junction fragment
g.26966del123763	ivs 7 to ivs 13	23,763 bp del	no	no	yes (Eco RI, Hind III)
g.43368ins6081	ivs 12 to ivs 13	6081 bp dup	no	no	yes (Hind III, Pst I)
g.44519del3835	ivs 12 to ivs 13	3835 bp del	no	no	yes? (Hind III)
g.53101del2998	ivs 14 to ivs 15	2998 bp	no	no	yes (EcoRI, HindIII, Pst I)
***	ivs 16 to ivs 19 ivs 19 to ivs 20	8216 bp del ~3.3 kb del	no yes	no ?	yes (Eco RI, Hind III) ?
g.58758del3094	ivs 16 to ivs 17	3094 bp del	yes	yes	yes (Eco RI, Hind III)
g.60453del 1008	ivs 16 to ivs 17	1008 bp del	yes	yes	yes (Eco RI, Hind III)
g.71348del3984	ivs19 to ivs 20	3984 bp del	yes	yes	yes (Eco RI, Hind III)
g.79504del510	ivs 21 to ivs 22	510 bp del	yes	yes	yes (PstI)

a. Is the indicated large rearrangement detectable by Long PCR analysis alone?

b. Is the indicated large rearrangement detectable by Long PCR followed by restriction digest?

c. Is the indicated large rearrangement detectable by Southern analysis of Eco RI-, Hind III-, or Pst I-digested gDNA?

*** not fully characterized

PCR template was lymphocyte gDNA from members of high-risk breast and/or ovarian cancer families with no identified mutations in either *BRCA1* or *BRCA2* and representative cell lines. The families were screened using lymphocyte gDNA from one linked and one unlinked family member. In the event that linkage was unclear, two high risk individuals (when available) were screened.

Fifteen unique Long PCR primer pairs were designed across the *BRCA1* genomic region (Table 2.2). The pairs amplify overlapping products ranging in size from 2.4 to 10.9 kb with an average size of 7.7 kb and a median size of 8.5 kb. The primer pairs provided full coverage of the *BRCA1* region, from the promoter to 5 kb downstream of the final exon with one exception. The region (GenBank L78833 bp~24,500-28,700) encompassing exon 6, exon 7 and the Alu-dense intron 7 was refractory to Long PCR amplification. A primer pair amplifying a smaller 3.8 kb product was eventually designed to screen exons 6 and 7. However, a 1187 bp gap in intron 7 remained (GenBank L78833 bp 27,666-28,853). This region was screened solely by Southern analysis.

Long PCR was carried out using the Boehringer Mannheim Expand™ Long Template PCR System. PCR products were amplified in 25 µl volumes containing 250 ng gDNA, 1x buffer 3, 500 µM dNTP's, 300 nM each primer, 0.25 mM MgCl₂ (in addition to 2.25 mM MgCl₂ from buffer 3), and 1.25 units enzyme mix. Amplification conditions were 10 cycles of denaturation at 94°C for 10 s, annealing at optimal temperature of each primer pair for 30 s, and extension at 68°C for 8 min. This routine was followed by 20 cycles in which the extension time was increased by 20 s each cycle and a final extension at 68°C for 7 min. The Long PCR products were analyzed further by restriction enzyme digestion and examined for variant banding patterns.

Southern Analysis of the BRCA1 Genomic Region

Southern blotting and hybridizations were performed as described previously using 10 µg of gDNA (Monnat et al., 1992). Southern hybridizations were performed by probing a single blot three times with different regions of the *BRCA1* cDNA (exons

Table 2.2 Long PCR Primers Used for Analysis of *BRCA1*

Primer Name	Exons^a	Temp.^b	Size (bp)	GenBank L78833^c	Sequence
n2.1.1bR BA2	NBR2 1b to 2	65	2364	2309-4672	5'-GCA ATG CAA AGA CCG TCC GCT G-3' 5'-GTA CTT CTT CAA CGC GAA GAG CAG ATA AAT C-3'
1aF 3R	1a to 3	65	9936	3028-12963	5'-TTT GGA CAA TAG GTA GCG ATT CTG ACC TTC-3' 5'-AAC TCC AGA CTA GCA GGG TAG GGG GGG-3'
3F 4R	3	61	5531	12792-18322	5'-TCC TGA CAC AGC AGA CAT TTA-3' 5'-CCC GTC TCT ACA GAA AAC AC-3'
ivs3F 7R	5 to 7	65	10857	13715-24571	5'-TGT GAA GAC AGG AAA GGA CCT GAT ACC AGT TTC-3' 5'-CAC GGT TTC TGT AGC CCA TAC TTT GGA TGA TAG-3'
6F-2 ivs7R	6 to 7	65	3828	23839-27666	5'-GCT TTT CAG CTT GAC ACA GGT TTG G-3' 5'-CCC CAG CAC TCC TAA GAA CAT TTA GTA TAG G-3'
8F 11R	8 to 11	65	8361	28853-37213	5'-CAG GAA ACC AGT CTC AGT GTC CAA CTC TCT AAC CTT G-3' 5'-TGT CAC TCA GAC CAA CTC CCT GGC TTT CAG AC-3'
11F 13R	11 to 13	65	9518	36783-46300	5'-CCA TAC ACA TTT GGC TCA GGG TTA CCG AAG AGG G-3' 5'-TTC GCA GGT CCT CAA GGG CAG AAG AGT CAC-3'

a. *BRCA1* exons unless otherwise indicated.

b. Temp.: Long PCR annealing temperature.

c. Region of GenBank L78833 amplified by the primer pair.

Table 2.2 Large Germline Rearrangements of *BRCA1* (cont.)

Primer Name	Exons^a	Temp.^b	Size (bp)	GenBank	
				L78833^c	Sequence
12F ivs13R	12 to 13	65	8927	37715-46641	5'-GGT GTG AGA GTG AAA CAA GCG TCT CTG AAG ACT GC-3' 5'-GCC TGT CAC CAA TTT CTC CCA TTC CAC TTA GCT TC-3'
13F 15R	13 to 15	65	8071	46235-54275	5'-GGA GCC AGC CTT CTA ACA GCT ACC CTT CCA TC-3' 5'-GAC TCC CAG AGC AAC TGT GCA TGT ACC ACC TAT C-3'
15F-2 18R-2	15 to 18	65	10613	54241-64852	5'-TGA TAG GTG GTA CAT GCA CAG TTG CTC TGG-3' 5'-GCT AAC TAC CCA TTT TCC TCC CGC AAT TCC-3'
17F 19R	17 to 19	61	4581	60922-65502	5'-GTG TAG AAC GTG CAG GAT TG-3' 5'-CAT TGT TAA GGA AAG TGG TGC-3'
18F 20R	18 to 20	65	6897	64777-71673	5'-TGC AGA TGC TGA GTT TGT GTG TGA ACG GAC-3' 5'-CCT GGG ATT CTC TTG CTC GCT TTG GAC C-3'
ivs19F 21R	20 to 21	65	7495	70154-77648	5'-TCC CAG TGA GGT GAA AAG CCG ATT GTT AAG TTC-3' 5'-CCC ATA GCA ACA GAT TTC TAG CCC CCT GAG G-3'
20F 23R	20 to 23	61	9662	71518-81179	5'-ATA TGA CGT GTC TGC TCC AC-3' 5'-ACT GTG CTA CTC AAG CAC CA-3'
22F ivs24R	22 to 24	63	8536	79560-88095	5'-AGC TGT GTG GTG CTT CTG TGG TGA AGG-3' 5'-AGA GCC AGC AAG ATC AGA TGG TCT ACA GGA C-3'

a. *BRCA1* exons unless otherwise indicated.

b. Temp.: The Long PCR annealing temperature.

c. Region of GenBank L78833 amplified by the primer pair.

1a-7, 3-11, and 11-24). Family and cell line gDNA was analyzed using at least two different restriction enzymes. Due to limiting quantities of tumor DNA, breast tumor gDNA was analyzed using a single restriction enzyme (HindIII), unless variant bands were observed. Case 72, the only breast tumor in which an unexpected restriction fragment was observed, was analyzed further using four different restriction enzymes (See Chapter Four). Regions involved in variant banding patterns were determined using the complete *BRCA1* genomic sequence (GenBank L78833) (Smith et al., 1996).

Population-based Screening for Two Large Germline Rearrangements

Patients were ascertained previously as part of the Carolina Breast Cancer Study (Newman et al., 1995; Newman et al., 1998). A PCR primer pair was designed to detect a 514 bp breakpoint junction fragment for the g.12977 ins10 del1039 mutation. Primers used to amplify the breakpoint junction fragment were: (forward) 5'-TTT-TTC-TCC-CCC-CCT-ACC-CTG-3'; (reverse) 5'-GCT-CAG-CAT-TTG-TTA-CTC-AAG-CTG-3'. PCR was performed in 25µl volumes containing 100 ng gDNA, 1x reaction buffer (Boehringer Mannheim), 250 µM dNTP's, nM each primer, and 1.25 units Taq enzyme. Primers used to amplify a 980 bp breakpoint junction fragment for the g.43368ins6081 mutation were: (forward) 5'-ATT-ATT-TCC-CCC-CAG-GCT-ACC-CAG-3'; (reverse) 5'-GGT-CCA-TTT-CAA-AGA-AGA-GTG-TGC-3'.

Results and Discussion

Characterization of a large germline deletion in the BRCA1 region leading to breast and ovarian cancer predisposition

I identified one large germline rearrangement from Family 58 (Family 5 in Hall et al., 1990). The rearrangement consisted of a 1039 bp deletion with a 10 bp inverted duplication inserted at the breakpoint junction (g.12977 ins10 del1039). The research characterizing this deletion is described in Appendix One. Due to the loss of 32 bp

from exon 3 in the g.12977 ins10 del1039 mutation, it was detected by Long PCR but was not obvious when examined by Southern analysis.

No other large germline rearrangements were detected in the families or cell lines using either Long PCR or Southern analysis. Several polymorphic restriction sites were observed. The variant restriction sites identified and their observed frequencies are listed in Table 2.3.

Population-based investigation of two large germline rearrangements predisposing to breast and ovarian cancer

Next I investigated the contribution of two different large germline rearrangements to breast cancer in the American population. I analyzed one large deletion (the g.12977 ins10 del1039 mutation) and one large duplication (the g.43368ins6081 mutation) in a population-based series of American breast cancer patients unselected for family history (Appendix One; Puget et al., 1999b). The g.12977 ins10 del1039 mutation was investigated in 242 white and 164 African-American breast cancer patients. No mutation carriers were detected in 406 cases. The g.43368ins6081 mutation was investigated in 220 white and 153 African-American breast cancer patients. One mutation carrier (CBCS 950749) was detected in 373 cases (Figure 2.1).

The g.43368ins6081 mutation was previously identified in four American families of mixed European descent with multiple cases of breast and/or ovarian cancer and one Portuguese family with three cases of breast cancer (Puget et al., 1999b; Heather Mefford, pers. comm.). A founder effect for the g.43368ins6081 mutation was postulated based on the observation that four of the families shared the same haplotype at nine polymorphic markers within or flanking the *BRCA1* locus (Puget et al., 1999b).

It was determined previously that young age at diagnosis alone did not predict *BRCA1* carrier status in the CBCS population, but that a family history of breast and ovarian cancer or at least four cases of breast cancer was predictive of *BRCA1* mutation carrier status (Newman et al., 1998). The g.43368ins6081 mutation carrier (CBCS

Table 2.3 Restriction Site Variants Identified in *BRCA1* from High-risk Breast and Ovarian Cancer Families without Known *BRCA1* or *BRCA2* Mutations

<u>Location</u>	<u>Primer Pair</u>	<u>Restriction Site</u>	<u>GenBank L78833^a</u>	<u>Families^b</u>	<u>Frequency</u>
ivs 3	ivs3F/7R	loss of PstI	16,569	1	1/33 chromosomes
ivs 19	18F/20R	loss of Aval	66,041	1	1/33 chromosomes
ivs 20	ivs19F/21R	loss of PstI	76,381	6	8/33 chromosomes
ivs 21	20F/23R	creation of Sty I	78,056 Civs21+386T	6	7/33 chromosomes

- a. Location of the restriction site affected unless the specific nucleotide alteration was characterized.
- b. The number of independent families in which the restriction site variant was identified.

Figure 2.1

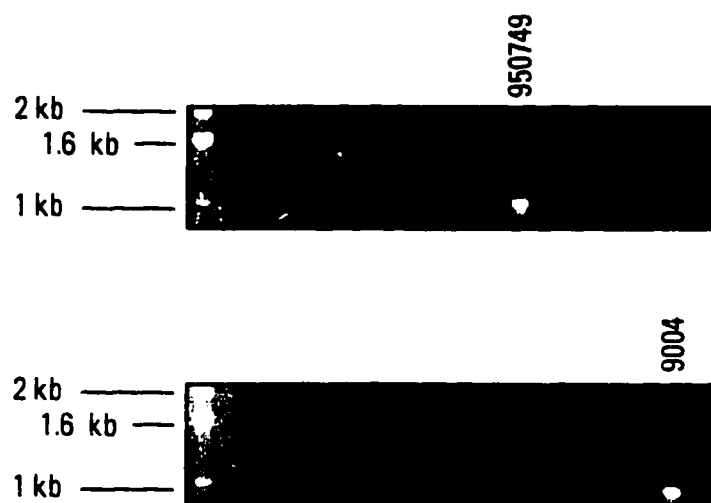


Figure 2.1 Detection of the g.43368ins6081 mutation in a population-based series of American breast cancer patients. The PCR product amplified across the breakpoint junction fragment is 980 bp. No product is amplified from wild-type *BRCA1*. Mutation carrier 950749 was identified among 373 breast cancer cases in the Carolina Breast Cancer Study. Individual 9004 from Family 90 serves as a positive control.

950749) was 53 years at diagnosis and did have a family history of breast cancer. CBCS 950749 was diagnosed with breast cancer at age 53 and with stomach cancer at age 60. The daughter of CBCS 950749 was diagnosed with breast cancer at age 29.

It is not known what contribution large germline rearrangements make to breast cancer in the American population. A growing number of such rearrangements are being described (Table 2.1). Detection of these large germline rearrangements, once fully characterized, will be a matter of conventional PCR over a breakpoint junction fragment using diagnostic primers that can be designed from the known genomic sequence of *BRCA1*. Small insertions, deletions, and point mutations account for approximately 3% of breast cancer in the American population when adjusted for sampling probabilities (Newman et al., 1998). It will be important to know whether large germline rearrangements make a similar contribution to breast cancer. This possibility is of particular interest because large rearrangements are the mutation types most likely to be missed by conventional mutation detection methods.

CHAPTER THREE

Somatic Alterations of *BRCA1* in Breast Cancer: Loss of Heterozygosity, but No Large Rearrangements

Experimental Logic and Chapter Contributions

The primary goal of my research was to test whether *BRCA1* is inactivated somatically by large genomic rearrangements that have not been identified by conventional mutation detection methods. I addressed this question by screening for large somatic rearrangements using 92 paired normal and tumor breast tissues. I examined the tumor gDNA for large somatic rearrangement of *BRCA1* in three different ways: loss of heterozygosity (LOH) analysis, Long PCR analysis, and Southern analysis. The tumors were not analyzed routinely for microsatellite instability as it is rare in breast cancer (Krajinovic et al., 1998; Anbazhagan et al., 1999). However, tumors that showed multiple alleles for microsatellite markers within the *BRCA1* genomic region were analyzed further with markers diagnostic of microsatellite instability (Dietmaier et al., 1997).

The tumors were not analyzed for small insertion, deletion and single bp somatic mutations. Several independent groups have screened breast and ovarian tumors for *BRCA1* mutations and only a few tumors with somatic mutation have been identified (Hosking et al., 1995; Merajver et al., 1995; Berchuck et al., 1998; Khoo et al., 1999). Thus, I reasoned that small somatic mutations of *BRCA1* were not likely to be present in our samples.

The paired normal and tumor samples were obtained through collaboration with two separate groups: the Cooperative Human Tissue Network (CHTN) and the Louisiana Women's Hospital in Baton Rouge. CHTN contributed 44 of the 92 samples. These samples were diagnosed by at least two independent pathologists prior to delivery. The Louisiana Women's Hospital agreed to send samples specifically for *BRCA1* research. These 48 samples included a pathology report and detailed family history for each patient. Samples from both groups were obtained on a prospective basis. I was responsible for organization and maintenance of the large collection of paired normal and tumor breast tissues. This effort involved coordination with the various divisions of CHTN, isolation of DNA from the 92 breast tissue pairs used for the study, and the creation of a computer database for the paired breast tissue collection.

While I was involved with the Louisiana Women's Hospital as well, Piri Welsh was largely responsible for the coordination of the collaboration.

The work described in this chapter was part of a collaborative effort. In this chapter, I will discuss those aspects of the research in which I was involved directly and the major observations from that research. The purpose of our study was to delineate the putative role of *BRCA1* in sporadic breast cancer. Using a single large set of 92 breast carcinomas of various histologies, we investigated several potential mechanisms for the decrease in *BRCA1* expression observed previously by other investigators. I screened the tumor set for large deletions and rearrangements as well as for conventional LOH. In addition, Tom Walsh analyzed the same tumor set for aberrant methylation within the *BRCA1* promoter. Finally, Rachel Gonzalez-Hernandez used immunohistochemistry to determine whether *BRCA1* protein was present within the tumors. By using an integrated approach involving analysis of large somatic rearrangements, LOH, methylation, and protein in a large group of breast carcinomas unselected for family history, we could determine whether any of these genetic alterations correlate with the previously observed reduction of *BRCA1* protein (Wilson et al., 1999; Jarvis et al., 1998; Taylor et al., 1998; Lee et al., 1999).

Materials and Methods

Tissue Samples

Ninety-two paired normal and tumor breast samples were obtained through the Cooperative Human Tissue Network (CHTN), and the Louisiana Women's Hospital in Baton Rouge. They comprised 74 invasive ductal carcinomas, 12 invasive lobular carcinomas, 4 invasive ductal with invasive lobular carcinoma, 1 mucinous carcinoma, and 1 apocrine carcinoma. Patients were not selected for family history or age at diagnosis. The samples were diagnosed by at least two independent pathologists, and contained at least 60% tumor cells. All tissues were snap frozen in liquid nitrogen following surgery. DNA from breast tumor tissue and normal breast tissue was isolated

using the Stratagene DNA Extraction kit. Histologic grade on 68 tumors was confirmed by an independent pathologist (Allan Gown).

LOH Analysis

LOH in the *BRCA1* genomic region was analyzed in paired normal and tumor gDNA using six microsatellite markers. D17S1323, D17S1322, D17S855 lie in *BRCA1* introns 12, 19, and 20 respectively (Smith et al., 1996). The *BRCA1* transcription start site is 39kb from D17S1323, 66kb from D17S1322, and 73kb from D17S855 (Smith et al., 1996). D17S1326 and D17S1325 are approximately 150 and 300kb from the 5' end (distal) of *BRCA1*, and D17S1320 is approximately 500kb from the 3' end (proximal) of *BRCA1* (Neuhausen et al., 1996). Markers were typed using PCR and electrophoresis conditions as described previously (Anderson et al., 1993). All *BRCA1* markers were evaluated for all normal and tumor breast tissues. An allele was scored as lost if intensity was reduced by >50%. Markers used for typing microsatellite instability were BAT26, BAT40, D2S123, D5S346, and Mfd15 (Dietmaier et al., 1997).

Long PCR and Southern Analyses

Long PCR, Southern blotting, and Southern hybridization were as described in Chapter Two.

Statistical analysis

Proportions of tissues in various categories were compared by χ^2 and Fisher's exact test, as appropriate. Dependent variables with ordinal categories were evaluated using χ^2 test for trend. Stratification for confounders was addressed as suggested for unmatched comparisons by Breslow and Day (1993). All P values represent 2-tailed tests.

Results

LOH of the BRCA1 region in sporadic breast carcinoma

The *BRCA1* genomic region investigated is shown in Figure 3.1. Of the paired breast normal and tumor samples informative for LOH, 50% (39/78) of invasive ductal tumors lost an allele at one or more markers within the *BRCA1* region (Table 3.1). None of the 12 lobular carcinomas showed any allelic loss. Of the 56 tumors informative for the three *BRCA1* intragenic markers (D17S1323, D17S1322, D17S855), 36 tumors (64.3%) had lost at least one intragenic marker. Allelic loss at the three *BRCA1* intragenic markers as well as at the three flanking markers (D17S1320, D17S1326, D17S1325) was associated significantly with tumor grade. LOH was detected in 44% (4/9) of grade I tumors, 28% (11/40) of grade II tumors, and 63% (26/41) grade III tumors ($p < 0.005$). This association is consistent with previous reports of association between somatic LOH at the *BRCA1* locus and tumor grade (Beckmann et al., 1996; Niederacher et al., 1997; Rio et al., 1998; Silva et al., 1999). Estrogen receptor status was independently associated with LOH ($p = 0.046$) as reported previously (Rio et al., 1998; Silva et al., 1999). There was no independent association between LOH and tumor stage or age at diagnosis.

Microsatellite instability in an infiltrating ductal carcinoma

Case 12 was unique in that additional alleles appeared at four markers coincident with LOH at the remaining two markers. In order to test whether the additional alleles were due to microsatellite instability (MSI), I analyzed markers indicative of MSI for Case 12. These markers were shown previously to be diagnostic of a mismatch repair defect in tumor cells (Dietmaier et al., 1997). Five of ten diagnostic markers were analyzed for Case 12: BAT26, BAT40, D2S123, D5S346, and Mfd15. Case 12 showed instability at four of the five markers tested. Only the BAT26

Figure 3.1

Microsatellite markers in the *BRCA1* region

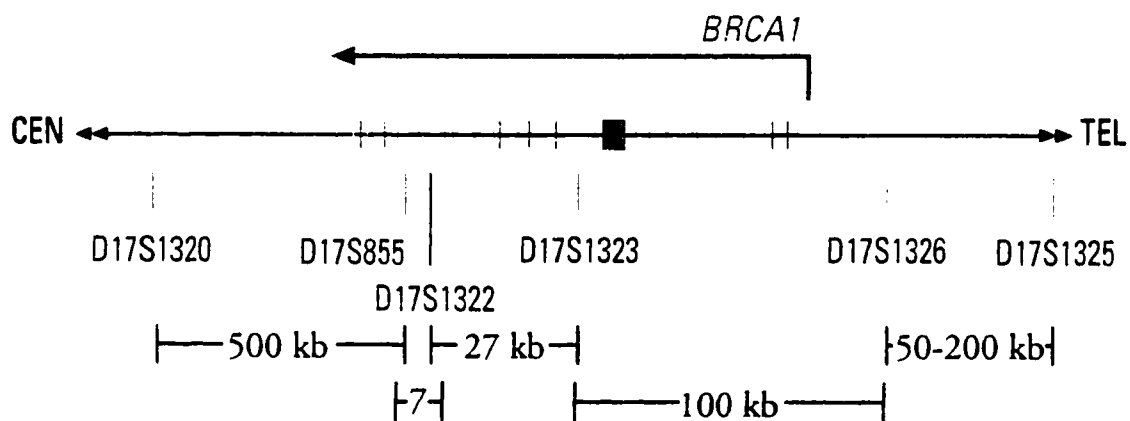


Figure 3.1 Microsatellite markers used to test loss of heterozygosity within the *BRCA1* genomic region. The distances between markers are diagrammed below. Arrows indicate the direction of transcription and boxes indicate exons.

Table 3.1 Genomic Loss in Primary Breast Carcinomas

Case ^a	Dx ^b	Grade	D17S1320	D17S855	D17S1322	D17S1323	D17S1326	D17S1325
72	66	II	●	●	●	●	●	NI
47	64	II	○	●	●	●	●	●
4	35	III	●	●	●	●	●	●
27 ^{mct??}	41	III	●	●	●	●	●	●
24	47	III	NI	●	●	●	●	●
41	57	III	NI	●	●	●	●	○
94	40	III	NI	●	●	●	●	NI
13		III	●	●	●	●	●	●
29	65	III	●	●	●	●	NI	NI
88	37	I	●	●	●	NI	NI	●
42	68	I	NI	●	●	NI	●	●
7	35	III	NI	●	●	NI	●	●
59	36	III	●	●	●	NI	●	NI
81	38	III	NI	●	●	NI	NI	NI
99	69	III	○	●	○	○	NI	NI
67	44	II	NI	●	NI	●	NI	NI
32	72	III	○	●	NI	●	NI	NI
79	69	I	●	●	NI	NI	●	NI
75	38	II	NI	●	NI	NI	●	NI
65	60	II	○	●	NI	NI	NI	NI
71	49	II	●	●	NI	NI	●	NI
62	38	III	NI	●	NI	NI	NI	●
18	49	III	●	●	NI	NI	NI	○
38	62	III	NI	●	NI	NI	NI	●
77	46	III	NI	●	NI	NI	NI	NI
1	66	II	○	○	●	●	●	●
39 ^{mixed}	45	III	NI	○	●	●	●	●
76	54	III	●	○	●	NI	●	●
98	63	II	NI	○	○	●	○	NI

a. Histologic type is indicated if other than ductal.

b. Dx indicates age at diagnosis.

c. Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.

Table 3.1 Genomic Loss in Primary Breast Carcinomas (cont.)

Case ^a	Dx ^b	Grade	D17S1320	D17S855	D17S1322	D17S1323	D17S1326	D17S1325
203	46	III	●	○	○	●	NI	●
54 ^{lob}	82	I	NI	○	○	○	○	NI
19	57	I	○	○	○	○	○	○
48 ^{muc}	71	I	●	○	○	○	○	○
46 ^{lob}	83	II	NI	○	○	○	NI	○
69	59	II	○	○	○	○	○	NI
85	45	II	○	○	○	○	○	NI
34 ^{mixed}	76	II	○	○	○	○	●	●
3	40	II	NI	○	○	○	○	NI
82		II	○	○	○	○	○	NI
60	42	II	○	○	○	○	○	○
56	52	II	○	○	○	○	○	NI
61	72	II	○	○	○	○	○	NI
44 ^{lob}	52	II	○	○	○	○	○	NI
201 ^{lob}	82	II	○	○	○	○	NI	○
95	44	III	○	○	○	○	NI	○
17	35	III	○	○	○	○	○	NI
43	46	III	NI	○	○	○	○	NI
96 ^{lob}	56	III	NI	○	○	○	○	○
78	42	III	○	○	○	○	NI	○
9		III	○	○	○	○	○	○
49 ^{lob}	59	I	NI	○	○	NI	NI	○
73	82	I	○	○	○	NI	○	NI
25	39	II	○	○	○	NI	○	○
58	42	II	NI	○	○	NI	○	NI
51	47	II	○	○	○	NI	○	○
50	31	II	○	○	○	NI	○	○
52	74	II	○	○	○	NI	○	○
202	62	II	NI	○	○	NI	NI	NI

a. Histologic type is indicated if other than ductal.

b. Dx indicates age at diagnosis.

c. Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.

Table 3.1 Genomic Loss in Primary Breast Carcinomas (cont.)

Case ^a	Dx ^b	Grade	D17S1320	D17S855	D17S1322	D17S1323	D17S1326	D17S1325
86	62	III	○	○	○	NI	NI	●
93	48	III	○	○	○	NI	NI	NI
23	61	III	●	○	NI	●	●	●
91 ^{lob}	53	I	NI	○	NI	○	NI	NI
20	65	II	NI	○	NI	○	NI	○
15	74	III	NI	○	NI	○	○	○
92	59	II	○	○	NI	NI	NI	○
22	40	II	NI	○	NI	NI	○	NI
33 ^{lob}	73	II	○	○	NI	NI	NI	NI
100	51	III	○	○	NI	NI	NI	●
28 ^{lob}	42	III	NI	○	NI	NI	○	NI
55	56	III	NI	○	NI	NI	NI	○
16	49	III	○	○	NI	NI	NI	NI
89 ^{lob}	62	III	○	○	NI	NI	○	NI
12		III	msi ^c	msi	msi	●	msi	●
21 ^{mixed}	50	III	○	NI	●	●	●	NI
30	45	III	●	NI	●	●	NI	NI
36	69	III	●	NI	●	NI	●	NI
64 ^{apn}	41	II	○	NI	○	●	●	NI
83	38	II	○	NI	○	○	○	○
2 ^{lob}	55	II	○	NI	○	○	○	NI
84 ^{mixed}	57	II	NI	NI	○	NI	NI	NI
204		II	NI	NI	○	NI	○	NI
40	89	III	○	NI	○	NI	○	NI
97	44	III	○	NI	○	NI	NI	○
90 ^{lob}	47		○	NI	○	NI	NI	○
63	52	II	NI	NI	NI	○	NI	○
11	52	II	●	NI	NI	NI	NI	●

a. Histologic type is indicated if other than ductal.

b. Dx indicates age at diagnosis.

c. Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.

d. "msi" indicates that multiple alleles were observed with this marker.

Table 3.1 Genomic Loss in Primary Breast Carcinomas (cont.)

Case^a	Dx^b	Grade	D17S1320	D17S855	D17S1322	D17S1323	D17S1326	D17S1325
53	62	II	○	NI	NI	NI	○	○
68	80	II	○	NI	NI	NI	○	NI
6	38	II	○	NI	NI	NI	○	NI
37	72	III	○	NI	NI	NI	○	NI
14	50		○	NI	NI	NI	○	○
35	71	II	○	NI	○	NI	NI	NI

a. Histologic type is indicated if other than ductal.

b. Dx indicates age at diagnosis.

c. Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.

marker was stable in tumor gDNA. Thus, Case 12 showed 80% unstable loci among the first five markers.

Long PCR and Southern analysis of sporadic breast carcinomas

Of the 92 samples, 84 were amplified successfully for all 15 Long PCR primer pairs. However, no large somatic rearrangements were detected. Several variant restriction sites were observed in both normal and tumor gDNA from the cases. The polymorphic restriction sites are listed in Table 3.2. One large germline deletion in intron 15 of *BRCA1* was detected in Case 72. This rearrangement will be described in detail in Chapter Four.

In order to detect very large deletions, duplications, and rearrangements that would be missed by PCR, I also performed Southern analysis. As tumor DNA was limiting and because no large rearrangements were identified using Long PCR, Southern analysis was performed with only the 40 tumors that demonstrated LOH within the *BRCA1* region (Figure 3.2). The eight samples that were not screened fully by Long PCR were analyzed by Southern as well. One ~250 bp germline insertion in intron 7 of *BRCA1* was observed in Case 72. This insertion will be described further in Chapter Four. No large somatic rearrangements were detected.

BRCA1 protein expression and loss of heterozygosity

Rachel Gonzalez-Hernandez analyzed BRCA1 protein expression in breast normal and tumor tissue sections by immunohistochemistry (IHC) using monoclonal antibody MS110. Her data are summarized in this paragraph in order to provide a context for my data in the following paragraphs. The relationship between the coding regions for the epitope recognized by MS110 and the microsatellite markers used for LOH analysis is shown in Figure 3.3. Rachel scored the normal and tumor breast tissues on a scale of 0 to 4. Normal breast epithelium exhibited staining consistent with a score of 3. Rachel observed reduced BRCA1 protein (scores of 0 to 2) as compared to

Table 3.2 Restriction Site Variants Identified in *BRCA1* from Breast Cancer Cases

<u>Location</u>	<u>Primer Pair</u>	<u>Restriction Site</u>	<u>GenBank L78833*</u>	<u>Frequency^b</u>
ivs 3	3F/4R	loss of Ava I	17,320	5/184 chromosomes
ivs 3	3F/4R	loss of Sty I	15,679	1/184 chromosomes
ivs 8	8F/11R	loss of Eco RI	30,718	7/184 chromosomes

- a. Location of the restriction site affected unless the specific nucleotide alteration was characterized.
- b. The number of chromosomes in which the restriction site variant was identified.

Figure 3.2

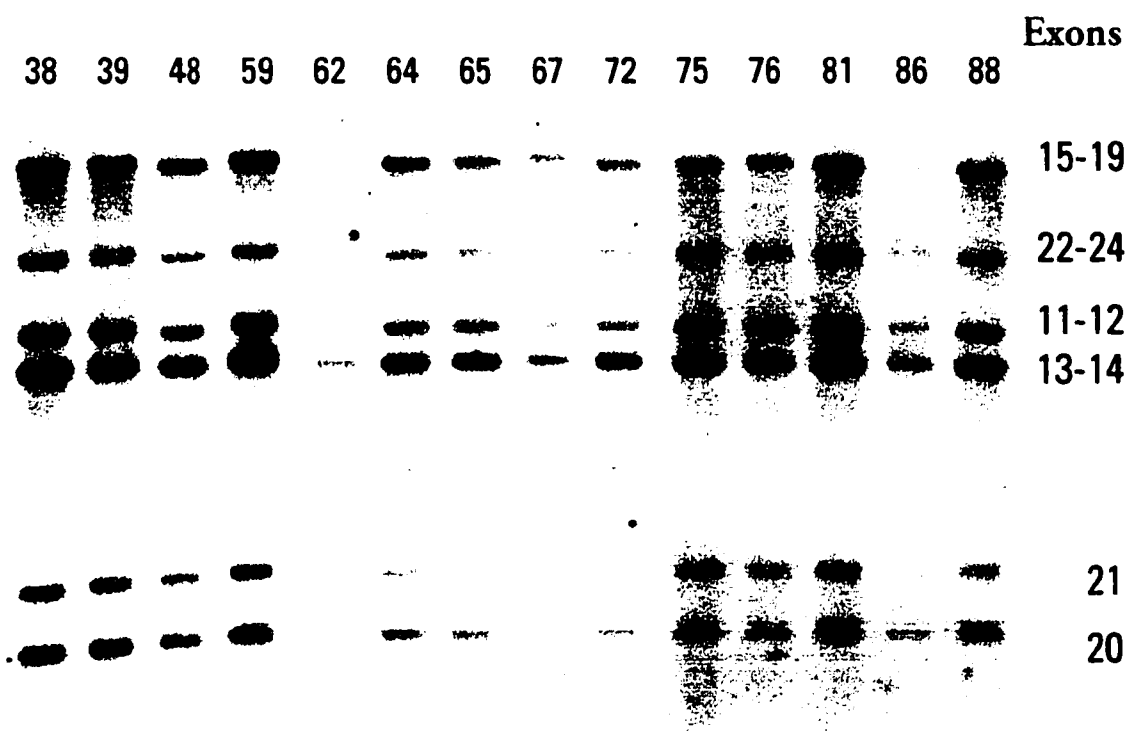
BREAST CANCER CASES SHOWING *BRCA1* REGION LOH

Figure 3.2 Southern analysis of *BRCA1* in primary breast carcinomas. 10 μ g of Hind III-digested genomic DNA was electrophoresed and Southern blotted as described in Materials and Methods. The blot was hybridized to a radioactively labelled probe made from exons 11 to 24 of the *BRCA1* cDNA. Of the 46 tumors analyzed by Southern, 14 are shown here. All 14 tumors appear wild-type.

Figure 3.3

BRCA1 region of Chromosome 17q21

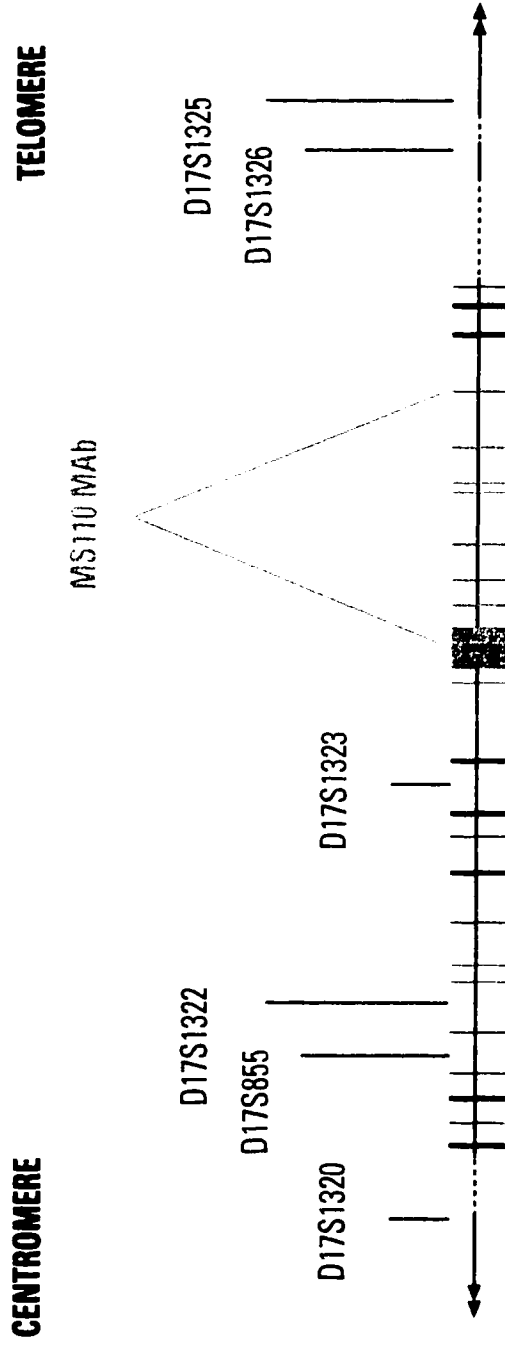


Figure 3.3 *BRCA1* region of chromosome 17q21. 92 breast carcinomas were analyzed for LOH using the six microsatellite markers indicated in blue. 84 of 92 breast carcinomas were analyzed by Long PCR and 46 of 92 breast carcinomas were analyzed by Southern analysis for large rearrangement of *BRCA1*. 68 of 92 breast carcinomas were analyzed by immunohistochemistry using MS110 MAb. The *BRCA1* exons that encode the epitope recognized by MS110 are indicated in orange.

normal breast epithelial tissue in 72% (49/68) of the breast tumors and in the majority of tumors (82%, 28/34) with high histologic grade ($p < 0.001$).

Reduced BRCA1 protein expression was associated with genomic loss at D17S1323, the intragenic marker closest to the *BRCA1* transcription start site (Table 3.3). Reduced BRCA1 protein staining (scores of 0 to 2 on a scale of 0 to 4) occurred in 92% (11/12) of tumors with genomic loss at D17S1323 but in 55% (12/22) without genomic loss at this marker ($p < 0.02$). Genomic loss at marker D17S1325 was associated marginally with reduced BRCA1 protein expression ($p = 0.05$), although the significance of this association is unclear. Marker D17S1325 is located ~50 to 300 kb 5' of *BRCA1*. Although marker D17S1326 lies between D17S1323 and D17S1325, LOH at D17S1326 is not associated with reduced BRCA1 protein staining (Table 3.4). The marker D17S1326 is only slightly less informative than the marker D17S1325. The maximum heterozygosity for marker D17S1326 is 0.8300 whereas the maximum heterozygosity for D17S1325 is 0.8858.

55% (12/22) of tumors in the series had reduced BRCA1 protein staining (scores of 0 to 2) despite no detectable genomic loss of the *BRCA1* transcription start as indicated by marker D17S1323. Reduced BRCA1 protein was observed in a high proportion (8/14, 57%) of the tumors that were informative at all three intragenic markers and fully retained both alleles of *BRCA1*. In this subset of tumors, other mechanisms are likely responsible for BRCA1 inactivation.

Discussion

Microsatellite instability and breast cancer

According to the criteria of Dietmaier et al. (1997), MSI is defined by at least 40% unstable loci among the ten diagnostic markers. These authors found that 14 of 15 tumors with >20% unstable loci had lost either hMSH2 or hMLH1 protein expression in tumor cells. Of the five MSI markers tested in Case 12, 80% showed instability. Therefore, Case 12 exhibits instability for at least 40% of the diagnostic microsatellite

**Table 3.3 Marker D17S1323 LOH
and BRCA1 Protein Expression**

	BRCA1 Protein Score	
	<u>Reduced BRCA1 (0-2)</u>	<u>Normal BRCA1 (3-4)</u>
<u>D17S1323</u>		
Loss	11	1
No loss	12	10

Table 3.4 Genomic Loss and BRCA1 Protein Expression

<u>Case^a</u>	<u>Grade</u>	<u>D17S1323^b</u>	<u>D17S1326</u>	<u>D17S1325</u>	<u>BRCA1^d</u>
4	III	●	●	●	0
27 ^{met?}	III	●	●	●	0
7	III	NI	●	●	0
12	III	●	msi ^c	●	1
23	III	●	●	●	1
24	III	●	●	●	1
41	III	●	●	○	1
15	III	○	○	○	1
95	III	○	NI	○	1
17	III	○	○	NI	1
43	III	○	○	NI	1
62	III	NI	NI	●	1
76	III	NI	●	●	1
86	III	NI	NI	●	1
88	I	NI	NI	●	1
100	III	NI	NI	●	1
18	III	NI	NI	○	1
25	II	NI	○	○	1
28 ^{lob}	III	NI	○	NI	1
37	III	NI	○	NI	1
40	III	NI	○	NI	1
58	II	NI	○	NI	1
59	III	NI	●	NI	1
75	II	NI	●	NI	1
84 ^{mixed}	II	NI	NI	NI	1

- Histologic type is indicated if other than ductal.
- Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.
- "msi" indicates multiple alleles were observed with this marker.
- Scores for MS110 detection of BRCA1 protein. The data in this column were generated by Rachel Gonzalez-Hernandez and are presented only to illustrate the correlation between low protein scores and LOH at the marker D17S1323.

Table 3.4 Genomic Loss and BRCA1 Protein Expression (cont.)

<u>Case^a</u>	<u>Grade</u>	<u>D17S1323</u> ^b	<u>D17S1326</u>	<u>D17S1325</u>	<u>BRCA1^d</u>
93	III	NI	NI	NI	1
39 ^{mixed}	III	●	●	●	2
64 ^{apo}	II	●	●	NI	2
67	II	●	NI	NI	2
72	II	●	●	NI	2
94	III	●	●	NI	2
46 ^{lob}	II	○	NI	○	2
78	III	○	NI	○	2
83	II	○	○	○	2
96 ^{lob}	III	○	○	○	2
54 ^{lob}	I	○	○	NI	2
69	II	○	○	NI	2
85	II	○	○	NI	2
91 ^{lob}	I	○	NI	NI	2
11	II	NI	NI	●	2
38	III	NI	NI	●	2
49 ^{lob}	I	NI	NI	○	2
51	II	NI	○	○	2
53	II	NI	○	○	2
55 [?]	III	NI	NI	○	2
92	II	NI	NI	○	2
97	III	NI	NI	○	2
65	II	NI	NI	NI	2

- Histologic type is indicated if other than ductal.
- Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.
- "msi" indicates multiple alleles were observed with this marker.
- Scores for MS110 detection of BRCA1 protein. The data in this column were generated by Rachel Gonzalez-Hernandez and are presented only to illustrate the correlation between low protein scores and LOH at the marker D17S1323.

Table 3.4 Genomic Loss and BRCA1 Protein Expression (cont.)

<u>Case^a</u>	<u>Grade</u>	<u>D17S1323</u> ^b	<u>D17S1326</u>	<u>D17S1325</u>	<u>BRCA1^d</u>
73	I	NI	○	NI	2
35	II	NI	NI	NI	2
9	III	○	○	○	3
19	I	○	○	○	3
48 ^{muc}	I	○	○	○	3
99	III	○	NI	NI	3
42	I	NI	●	●	3
16	III	NI	NI	NI	3
36	III	NI	●	NI	3
77	III	NI	NI	NI	3
47	II	●	●	●	4
34 ^{mixed}	II	○	●	●	4
20	II	○	NI	○	4
63	II	○	NI	○	4
2 ^{lob}		○	○	NI	4
3	II	○	○	NI	4
82	II	○	○	NI	4
50	II	NI	○	○	4
71	II	NI	●	NI	4
81	III	NI	NI	NI	4

- Histologic type is indicated if other than ductal.
- Open circles denote retention. Closed circles denote allelic loss. "NI" indicates the marker was uninformative.
- "msi" indicates multiple alleles were observed with this marker.
- Scores for MS110 detection of BRCA1 protein. The data in this column were generated by Rachel Gonzalez-Hernandez and are presented only to illustrate the correlation between low protein scores and LOH at the marker D17S1323.

markers. Given these criteria, the infiltrating ductal carcinoma from Case 12 likely represents a mismatch repair defective breast tumor. MSI is indicative of widespread genomic instability and is well-documented in hereditary non-polyposis colorectal cancer (HNPCC). The MSI phenotype is observed in many sporadic cancers as well, including colorectal, gastric, pancreatic and endometrial cancers (Peltomaki, 1997). There have been reports of breast tumors with MSI from affected members of HNPCC kindreds (Bergthorsson et al., 1995; Risinger et al., 1996; Boyd et al., 1999). However, MSI is rare in sporadic breast cancer (Krajinovic et al., 1998; Anbazhagan et al., 1999). In a recent study, 0 of 267 breast tumors exhibited MSI (Anbazhagan et al., 1999).

Although MSI is rare in sporadic breast cancer, LOH of some DNA mismatch repair loci is frequent (Benachenhou et al., 1999). Chromosomal region 3p21, which harbors hMLH1, shows LOH in 46% of breast tumors (Benachenhou et al., 1999). Benachenhou et al. (1999) proposed that hMLH1 could be involved in breast tumorigenesis through cellular functions other than replication error correction. It is possible that hMLH1 plays a tumor suppressing role in sporadic breast cancer through functions other than replication error correction and that Case 12 represents a rare example in which loss of the breast tumor suppressing function also affects replication error correction.

A more likely possibility is that Case 12 represents an affected member of a HNPCC kindred. Case 12 was 42 years at diagnosis (early-onset), but no family history is available. It is, therefore difficult to assess the probability that Case 12 is a member of a HNPCC kindred.

Absence of large somatic rearrangements in sporadic breast cancer

Despite extensive analysis by our group and others, very few somatic mutations in *BRCA1* have been reported in breast cancer (Hosking et al., 1995; Merajver et al., 1995; Berchuck et al., 1998; Khoo et al., 1999). Nuclear staining BRCA1 protein, nevertheless, is reduced or absent as compared to normal mammary ductal epithelium in the great majority of the breast carcinomas analyzed by IHC. The reduction of BRCA1

nuclear protein observed in our samples confirms previous observations by numerous groups using a variety of *BRCA1* antibodies directed against diverse regions of *BRCA1* (Wilson et al., 1999; Jarvis et al., 1998; Taylor et al., 1998; Lee et al., 1999). Thus, while somatic mutation of the *BRCA1* gene is rare in sporadic breast cancer, reduction of *BRCA1* protein staining is extremely common.

Paradoxically, analysis of *TP53*, one of the first tumor suppressors identified, indicated that characterization of genes involved in hereditary breast cancer would identify factors involved in the genesis sporadic forms of breast cancer as well. Germline mutations in *TP53* predispose to Li-Fraumeni syndrome, a cancer susceptibility syndrome of which breast cancer is an integral component. In contrast to *BRCA1*, *TP53*, is mutated somatically in 20–40% of sporadic breast carcinomas (Osborne et al., 1991; Coles et al., 1992; Grenblatt et al., 1994).

BRCA1, however, is not the only gene in which germline mutations predispose to breast cancer, but in which somatic mutations in sporadic breast cancer are rare. Inherited mutations in *BRCA2* also predispose to breast and ovarian cancer with a slightly lowered risk of ovarian cancer as compared to *BRCA1* (Wooster et al., 1995; Stratton et al., 1996). As seen for *BRCA1*, somatic mutations of *BRCA2* in sporadic breast cancer are very rare (Lancaster et al., 1996). Mutations in the *PTEN* and *LKB1* genes lead to inherited predisposition syndromes in which breast (*PTEN*) or breast and ovarian cancer (*LKB1*) are integral cancers (Li et al., 1997; Steck et al., 1997; Hemminki et al., 1998). Although germline mutations in *PTEN* and *LKB1* predispose to breast and/or ovarian cancer, somatic mutations of these genes in breast and ovarian cancer are rare (Rhei et al., 1997; Bignell et al., 1998). Interestingly, *PTEN*, like *BRCA1*, shows reduced protein expression in many sporadic breast carcinomas (33%) despite the absence of detectable somatic mutations (Perren et al., 1999).

Investigators have interpreted the lack of somatic mutations in these genes (*BRCA1*, *BRCA2*, *PTEN*, and *LKB1*) as a general indication that these genes are not involved in sporadic breast cancer. It is possible, however, that pathways involving these tumor suppressor genes are targets of inactivation in sporadic breast cancer by somatic mutation of another pathway component. The *RB1* gene product, a classic

tumor suppressor identified by analysis of families with inherited predisposition to retinoblastoma, shares a common pathway with p16 (*CDKN2A*) and cyclin D1 (*D1*). The pRb pathway is targeted for inactivation in a number of sporadic cancer types. Inactivation of any one component of the pRb pathway in a tumor greatly decreases the probability of identifiable damage to other components (reviewed in Sherr, 1996).

The absence of DNA mutations affecting the coding and/or splicing of a gene does not necessarily imply that the gene itself is not targeted for inactivation (Jones & Laird, 1999). As described in the introduction to this chapter, Tom Walsh has tested whether *BRCA1* is subject to epigenetic regulation via methylation in sporadic breast cancer using methylation-specific PCR (MSP). MSP of the *BRCA1* δ promoter indicated that the δ promoter was aberrantly methylated in a number of breast tumors. Magdinier et al. (1998), however, did not observe methylation of the *BRCA1* δ promoter in any of 37 sporadic breast carcinomas while Catteau et al. (1999) detected aberrant methylation in only 11 of 96 sporadic breast carcinomas. The discrepancy is most likely explained by differential sensitivity of the techniques used. Whereas Magdinier et al. and Catteau et al. employed Southern hybridization using methylation sensitive restriction enzymes, our study and those of Mancini et al. (1998) and Rice et al. (1998) employed sodium bisulfite treatment followed by PCR-based assays.

The loss of a necessary transcription factor upstream of *BRCA1* might also be responsible for loss of BRCA1 protein in tumors without either a heavily methylated *BRCA1* δ promoter or genomic loss of the critical region near the transcription start site (as indicated by LOH at D17S1323). Very little is known about the *trans*-acting factors required for *BRCA1* expression. A 31 bp minimal *cis*-acting region within the *BRCA1* promoter has been defined (Thakur & Croce, 1999). Alterations in the regulatory factors that bind this region may lead to suppression of *BRCA1* expression.

Correlation of LOH at marker D17S1323 with loss of BRCA1 protein

The association of reduced BRCA1 protein with *BRCA1* LOH in tumors is specific to LOH at marker D17S1323, the intragenic marker closest to the 5' end of

BRCA1. It was unexpected that LOH at the more 3' (proximal) markers was not associated with reduced BRCA1 protein staining. Since disease-associated protein truncating mutations can occur throughout the gene, loss of markers at the 3' end of *BRCA1* would be expected to be as deleterious as loss of markers at the 5' end of *BRCA1*. The specificity of the association may explain the previous observation that BRCA1 protein expression was not correlated with allelic loss as defined by loss at two out of the three intragenic markers (Bernard-Gallon et al., 1999). It is worthwhile to note that the one tumor (Case 12) exhibiting MSI, shows classic LOH at the two markers most closely associated with reduced BRCA1 protein (D17S1323 and D17S1325).

LOH at D17S1323 in *BRCA1* intron 12 might reflect specific targeting of either the *BRCA1* promoter or *BRCA1* exon 11 for loss. BRCA1 has a role in maintenance of genomic stability and is required for the G2-M checkpoint in response to certain types of DNA damage. The function of BRCA1 that is involved in maintenance of genetic stability through the G2-M checkpoint has been shown to require exon 11 (Xu et al., 1999).

Alternatively, LOH at D17S1323 may target the promoter region and transcription start site of *BRCA1*. Other markers more 5' of D17S1323 could enable one to distinguish between these two possibilities. If loss of the *BRCA1* promoter region and transcription start site is targeted, then the association between LOH and reduced BRCA1 protein should become stronger with more 5' markers. If, however, exon 11 is the target of LOH, then the association between LOH and reduced BRCA1 protein will weaken with increasing distance from exon 11.

CHAPTER FOUR

Characterization of a Variant *BRCA1* Allele

Experimental Logic and Chapter Contributions

In the course of analyzing *BRCA1* for large rearrangement, two large germline rearrangements of unknown significance were identified from a single allele in normal and tumor gDNA from Case 72. In addition, several small alterations were identified in the same allele from Case 72 containing the two large germline rearrangements.

One large rearrangement identified from Case 72 involves germline deletion of 977 bp from *BRCA1* intron 15. The second large rearrangement involves germline expansion of a compound microsatellite sequence in *BRCA1* intron 7. While neither of the rearrangements alter coding regions of *BRCA1*, they are of interest because the allele that harbors both of these rearrangements is retained in an infiltrating ductal carcinoma from Case 72 while the more “common” allele of *BRCA1* is lost.

There is a 50% probability that the Case 72 variant allele of *BRCA1* is retained in the tumor randomly. In order to test whether the large rearrangements might affect processing of the *BRCA1* transcript from the rare allele, we analyzed splicing of exons flanking the intronic rearrangements. Additionally, I tested the variant allele for loss of transcript using a single nucleotide polymorphism (SNP) that was identified in exon 16.

The difficulty in classifying the status of the large germline rearrangements identified from Case 72 highlights an area of increasing interest: noncoding variation in the human genome. The determination of the amount and the degree to which individual genomes vary will be critical to determining both the relevance to human disease for variant alleles such as the rare allele described from Case 72 and the use of genetic variation to infer the evolutionary history of human populations.

Materials and Methods

Long PCR and Southern Analyses

Long PCR, Southern blotting, and Southern hybridization were as described in Chapter Two.

Conventional PCR, sequencing, and nested PCR analysis

PCR products were cloned using the Original TA Cloning® kit by Invitrogen. Following the manufacturer's instructions, cloned PCR products were sequenced using the Big Dye Terminator Cycle Sequencing Ready Reaction DNA Sequencing Kit (Perkin Elmer) and reactions were read on an ABI Model 377 DNA Sequencer (Applied Biosystems). Primers used to sequence cloned PCR products were exon16 gDNA SSCP reverse: 5'-AAT-TCT-TAA-CAG-AGA-CCA-GAA-C-3' and a primer designated L78833:57322-57303: 5'-GGC-ATT-ATG-TAG-CAA-ACA-CC-3'.

The ivs 7b primers were: (forward) 5'-GAG-TGC-TGG-GGT-TTT-ATT-GTC-ATC-3' and (reverse) 5'-ATG-ATG-CCT-GGA-AAA-AAT-GCC-3'. Amplification conditions were initial denaturation at 94° for 2 minutes, 35 cycles of denaturation at 94° for 45 sec, annealing at 56° for 45 sec, and extension at 72° for 45 sec, followed by a 7 minute final extension at 72°. PCR products amplified using the ivs 7b primer pair were separated in 0.8% agarose and gel-extracted using the Qiagen Qiaex Extraction kit. 50 ng of gel-extracted DNA was used in a sequencing reaction with the primer ivs7bR3: (forward) 5'-ACA-AGC-GTG-TGC-AAC-TAT-G-3' and Big Dye terminator chemistry.

In order to estimate the size of the intron 7 compound microsatellite more accurately, nested PCR primers were designed: (forward) 5'-GGA-GAA-TCA-ACT-GAA-CCA-GGG-AG-3' and (reverse) 5'-ACA-AGC-GTG-TGC-AAC-TAT-GCC-3'. Nested PCR was performed on 1:100 dilutions of the ivs 7b product using $\delta^{32}\text{P}$ -dCTP. The radioactively labeled products were separated on a 5% polyacrylamide gel for 4 hrs at 80 W.

RT-PCR Analysis

RNA was obtained from 75 mg of tissue taken from an infiltrating ductal carcinoma occurring in Case 72 using the Qiagen RNeasy® Midi kit. Reverse

transcription (RT) was performed with 1 μ g of RNA using random decamers and the Ambion Retroscript™ kit. PCR on cDNA was performed using 1-2 μ l of the RT reaction in a 50 μ l volume containing 1x reaction buffer (Boehringer Mannheim), 250 μ M dNTP's, 500 nM each primer, and 1.25 units Taq enzyme.

Amplification conditions and primer pairs were as previously described (Friedman et al., 1994). Briefly, amplification conditions were initial denaturation at 94° for 2 minutes, 35 cycles of denaturation at 94° for 1 minute, annealing at 58° for 1 minute, and extension at 72° for 1 minute, followed by a 7 minute final extension at 72°. The C3 forward primer (exon 6) was paired with either C3 reverse (exon 8) or C4 reverse (exon 11) to analyze splicing of exons 6 and 7. Splicing of exons 15 and 16 was analyzed using either the C7 forward primer (exon 13) or the C8 forward primer (exon 15) paired with either the C9 reverse primer (exon 17) or the C10 reverse primer (exon 20). The C8 forward primer was also paired with the C8 reverse primer (exon 16) to test whether portions of intron 15 might be spliced into the *BRCA1* message inappropriately as a result of the g.56024 del977 deletion.

Results

Characterization of a large germline deletion of unknown significance from Case 72.

A large germline deletion of *BRCA1* was identified from a sporadic breast cancer patient (Case 72). Using primer pair 10, designed to amplify a 10.6 kb product encompassing exons 15-18, variant restriction fragments were detected using four different restriction enzymes (data not shown). The variant restriction fragments were present in equal intensity in gDNA from normal breast epithelial tissue of Case 72, but the variant bands were preferentially retained in gDNA prepared from an infiltrating ductal carcinoma of Case 72.

Based on the restriction digests, a minimal region of ~3800 bp (GenBank L78833 bp 54885-58706) containing a deletion of ~1000 bp was identified near *BRCA1* exon 16. Wild-type and variant PCR products encompassing the deletion were cloned

and sequenced. All clones from the variant PCR products contained an intact exon 16, indicating that the breakpoint occurred in intron 15. Sequence of all variant clones revealed the simple deletion shown in Figure 4.1. Based on *BRCA1* genomic sequence L78833, the deletion is a 977 bp deletion of GenBank L78833 bp 56025-57001 from *BRCA1* intron 15. The notation for the deletion (Antonarakis et al., 1998) is g.56024 del977 based on GenBank L78833. The 5' breakpoint for g.56024 del977 occurs in the 3' tail of an Alu-Sp element and the 3' breakpoint occurs 9 bp into the 5' head of an oppositely oriented Alu-Y element. Southern analysis using a probe to *BRCA1* exons 11-24 confirmed the deletion and indicated that the deleted allele was retained in Case 72 tumor DNA.

It is unclear what effect, if any, the g.56024 del977 deletion has on the transcription and translation of *BRCA1*. The deletion does not affect splicing in the region as determined by RT-PCR using a variety of primer pairs located in exons 13, 15, 16, 17, and 20 (Figure 4.2). The 977 bp deletion in intron 15 of Case 72 was not the only large germline rearrangement detected in the allele preferentially retained in the infiltrating ductal carcinoma of Case 72.

Characterization of a large expansion of a compound microsatellite sequence in BRCA1 intron 7

Southern analysis of Hind III-digested Case 72 tumor gDNA revealed an extra band of ~4.3 kb. Further Southern analysis of Case 72 normal and tumor gDNA using multiple restriction enzymes indicated that a minimal region of ~3500 bp (GB L78833 bp 26,911-30394) including *BRCA1* exon 8 contained an insertion of ~250 bp that was preferentially retained in tumor DNA (Figure 4.3 and data not shown). Overlapping PCR primer pairs were designed to span the region and tumor DNA from Case 72 was amplified. The exon 8 containing PCR product was of expected size, but a primer pair (ivs7b) amplifying a region just upstream of exon 8 (GenBank L78833 bp 27,656-28,733) yielded a product ~250 bp larger than expected from Case 72 tumor gDNA (Figure 4.4).

Figure 4.1

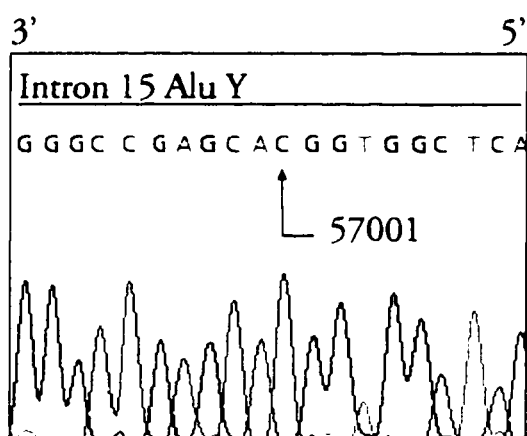
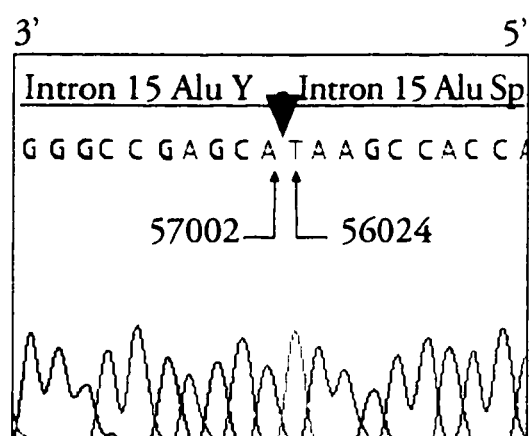
Wild-type BRCA1 Sequence:**977 bp Germline Deletion:**

Figure 4.1 977 bp germline deletion (g.56024del977) On the left is sequence from the unretained Case 72 *BRCA1* allele. The sequence from the unretained allele matches GenBank L78833 bp 57,010-56,991, near the 3' breakpoint of the g.56024del977 allele. On the right is sequence from the 977 bp germline deletion breakpoint junction of Case 72. The red arrow indicates the breakpoint junction. The 3' and 5' symbols represent sequence orientation with respect to *BRCA1* transcription.

Figure 4.2

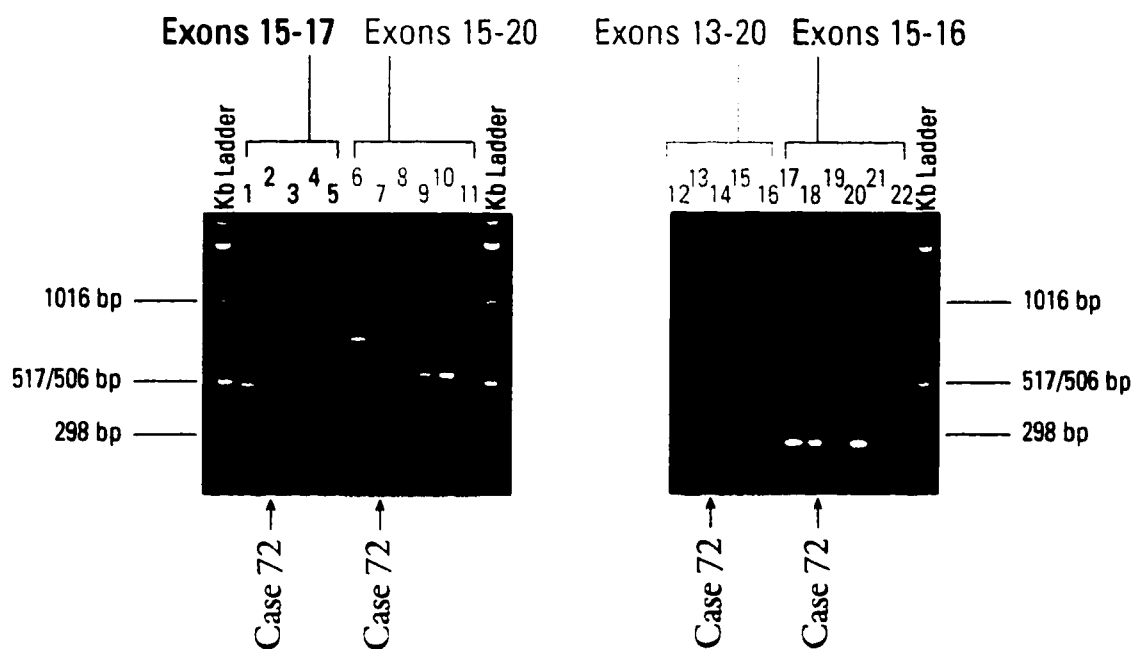


Figure 4.2 Analysis of RT-PCR products from Case 72. RT-PCR primer pairs were as follows: C8/C9 (lanes 1-5) amplifies a 499 bp product from *BRCA1* exons 15-17, C8/C10 (lanes 6-11) amplifies a 737 bp product from exons 15-20, C7/C10 amplifies a 1059 bp product from exons 13-20, and C8/C8 amplifies a 266 bp products from exons 15-16. Lanes 1, 6, 12, 17: control cDNA from an individual without *BRCA1* mutation (9104). Lanes 2, 7, 13, 18: cDNA from the Case 72 breast tumor. Lanes 4, 9, 15, 20: cDNA from another breast tumor (Case 265). Lanes 3, 8, 14, 19 (Case 72) and lanes 5, 10, 16, 21 (Case 265): control reverse transcription reactions without reverse transcriptase. Lanes 11 and 22: negative water controls.

Figure 4.3

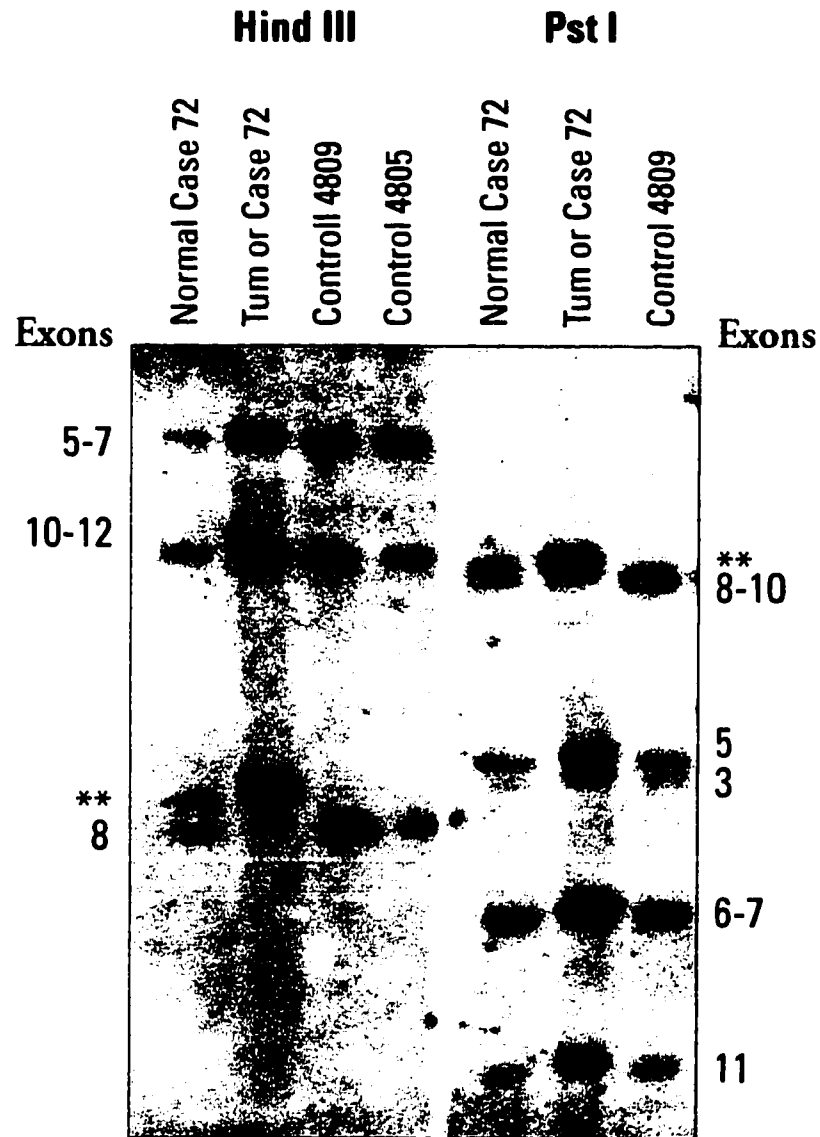


Figure 4.3 Southern analysis of a large germline rearrangement involving the *BRCA1* exon 8 genomic region. 10 μ g of restriction enzyme-digested genomic DNA was electrophoresed and Southern blotted as described in Materials and Methods. The blot was hybridized to a radioactively labelled probe made from exons 3 to 11 of the *BRCA1* cDNA. ** indicates the rearrangement junction fragment. Control individuals are from Family 48.

Figure 4.4

PCR Analysis of Minimal Genomic Region for Case 72 Insertion

L78833 bp 26,911-30,394

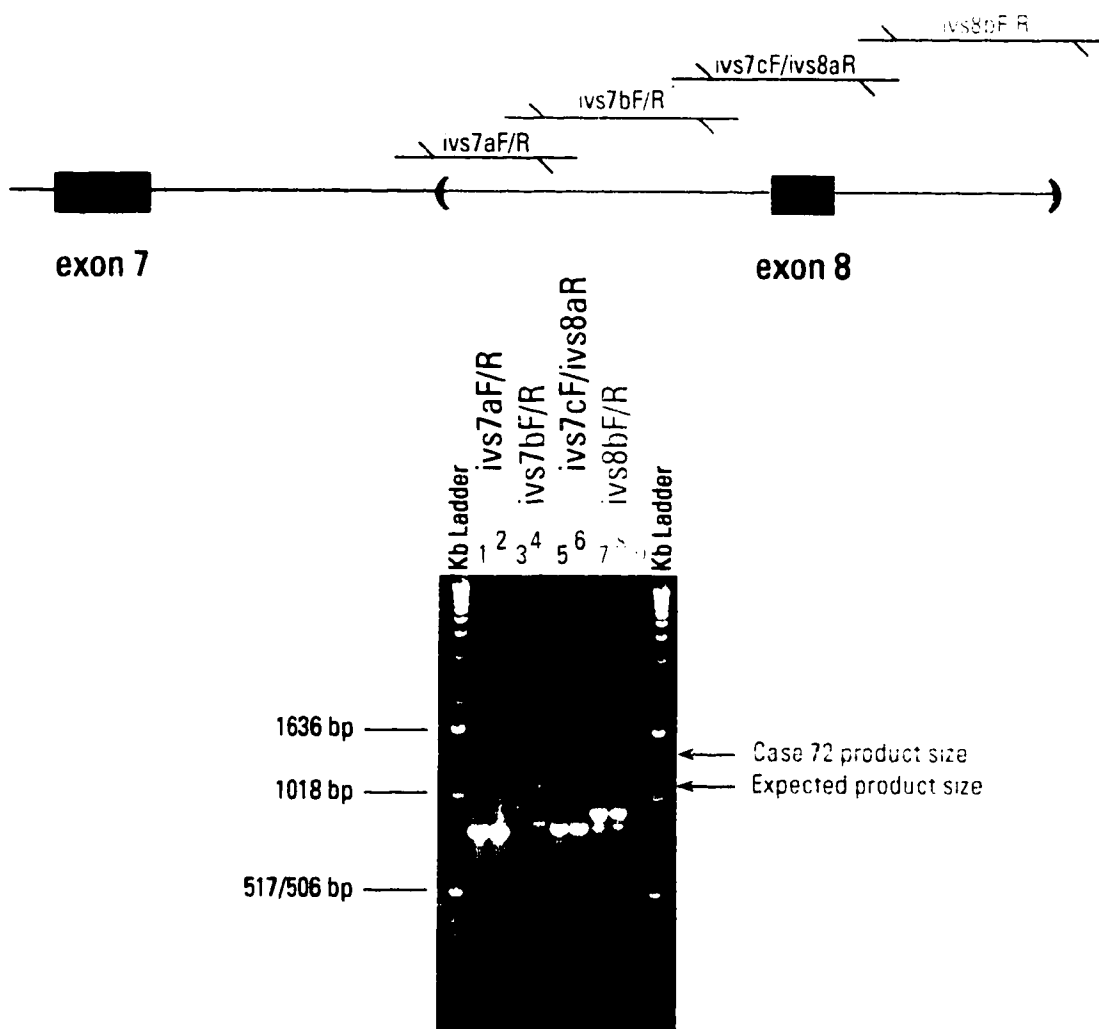


Figure 4.4 PCR analysis of the minimal genomic region for the 250 bp insertion in the retained allele from Case 72. Overlapping PCR primer pairs were designed to cover the region. The minimal genomic region containing *BRCA1* exon 8 is indicated as are the overlapping PCR products used to analyze the region. All amplified products were of the expected size except the product amplified with the primer pair ivs7b (indicated by red arrows). Lanes 1, 3, 5, 7: control individual (4310). Lanes 2, 4, 6, 8: Case 72 tumor gDNA. Lane 9: water.

The ivs 7b primer pair amplifies a number of minor bands in addition to the intended product. These minor bands represent internal products as gel extraction of the largest band and subsequent amplification of the specific product yields the same minor bands. The variant ivs 7b product for Case 72 was gel-extracted and sequenced using internal primers. Multiple independent sequencing reactions revealed an expansion of a compound microsatellite located at GenBank L78833 bp ~28,087 to 28,174 (Figure 4.5). In the published genomic sequence the imperfect compound microsatellite GTN (AT)₄₋₈ is ~87 bp in size. In the variant allele from Case 72, the compound microsatellite is ~333 bp in size.

The compound microsatellite appears to be polymorphic in the population. Nested PCR analysis of 40 affected and 55 unaffected individuals from multigeneration breast cancer families indicated that allele size ranges from ~300 bp to ~460 bp, with the ~350bp allele (the size predicted from the published genomic sequence) being the most common allele among those surveyed (data not shown). Nested PCR of the allele from Case 72 yielded a 600 bp product. The ~600 bp allele from Case 72 is the largest allele observed to date.

As with the g.56024 del977 deletion, the effect, if any, of the expanded compound microsatellite in *BRCA1* intron 7 on the transcription and translation of *BRCA1* is unclear. The expanded repeat does not affect splicing in the region as determined by RT-PCR using a variety of primer pairs located in exons 6, 8, 9, and 11 (data not shown).

Other sequence alterations observed in the variant allele from Case 72

Several other alterations were observed in the variant allele from Case 72 (Figure 4.6). Two single bp substitutions in *cis* to the g.56024 del977 deletion were observed in exon 16. One is a C to T transition at nucleotide (nt) 4801 of HSU14680. The substitution does not affect translation of the *BRCA1* message. This SNP has been observed before, although at a very low frequency. It was identified in one African-American patient in the Carolina Breast Cancer Study and is present in about 1% of the

Figure 4.6

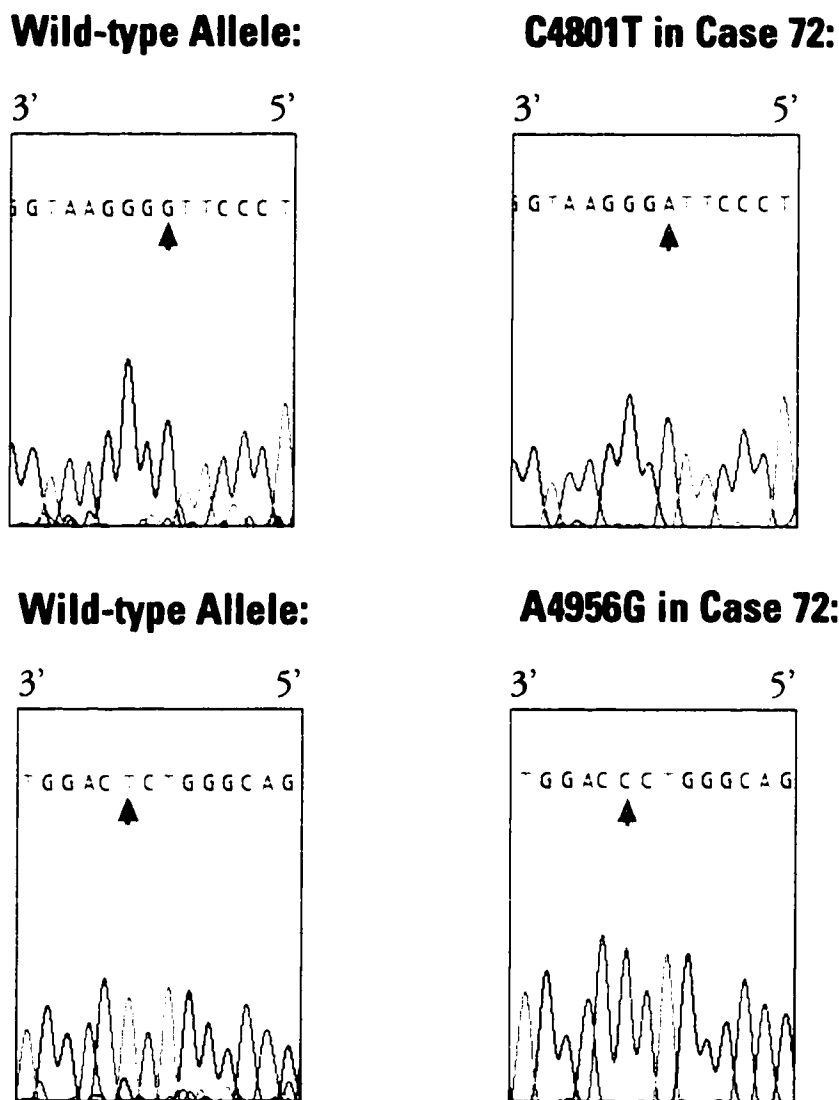


Figure 4.6 Single nucleotide polymorphisms identified in the variant allele from Case 72. The panels on the left represent wild-type sequence in the wild-type allele from Case 72. The panels on the right represent sequence of single nucleotide polymorphisms in *BRCA1* exon 16 of the variant allele from Case 72. In the top panels, sequence flanking a C to T transition at nt 4801 of HSU14680 is shown. In the bottom panels, sequence flanking an A to G transition at nt 4956 of HSU 14680 is shown. The red arrows indicate the location of the transition substitutions.

population (BIC, 1999). The second substitution is an A to G transition at nt 4956. The substitution results in a missense polymorphism at codon 1613 (S1613G). This missense polymorphism has been observed before as well and is present in 32% of the population (Dunning et al., 1997). Finally, a 24 bp insertion at GenBank L78833 bp 57,211 was identified in the variant allele (Figure 4.7). The origin of the 24 bp insertion is unknown, but it contains a region of 13-20 bp identity with sequence from human chromosome 17 clone hRPK.268_F_2, human chromosome 6 PAC 271G9, and an unassigned human BAC RG191D16.

Given the high percentage of rare variant sites within the retained allele from Case 72, tumor gDNA was tested for microsatellite instability using diagnostic markers as described previously (Dietmaier et al., 1998). No microsatellite instability was observed in the tumor DNA from Case 72.

Analysis of transcript loss in Case 72

As splicing of the variant allele was not affected around either of the large rearrangements, we tested whether the variant allele might produce an unstable transcript. Two SNP's in *BRCA1* were identified from cloned PCR products containing the g.56024 del977 breakpoint junction fragment. One of these, the A4956G missense substitution in exon 16 creates an Ava II site. The AvaII site was used to test the stability of the transcript from the variant allele. I tested this hypothesis by determining whether *BRCA1* message obtained from the Case 72 breast tumor contained transcripts from the variant allele (Ava II site at position 4956). In Case 72 breast tumor gDNA, only the variant allele is present. If only the wild-type allele (no Ava II site at position 4956) is present in cDNA from the tumor, then the message from the variant allele is likely to be unstable. The amplified cDNA with no Ava II site at position 4956 (A4956) would represent transcription from the wild-type allele in contaminating normal cells within the tumor.

RT-PCR was performed on cDNA prepared from Case 72 tumor tissue using a forward primer in exon 15 and reverse primers in either exon 17 or exon 20. Results are

Figure 4.7

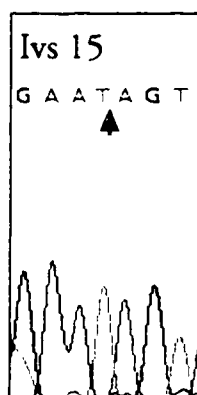
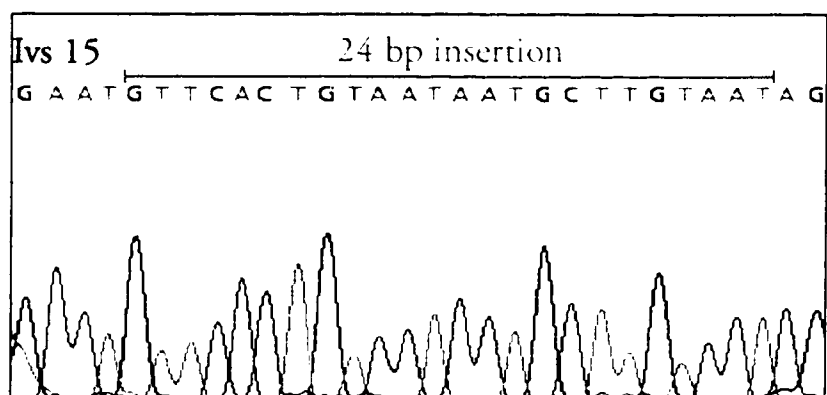
Wild-type:**24 bp Germline Insertion in Case 72:**

Figure 4.7 24 bp insertion in *BRCA1* intron 15. Sequence from the unretained Case 72 *BRCA1* allele is shown in the left panel. The sequence from the unretained allele matches the GenBank L78833 sequence. Sequence from the junction fragment of the 24 bp insertion identified from the variant allele of Case 72 is shown in the right panel. The red bar indicates the inserted nucleotides and the red arrow indicates the site of insertion.

shown in Figure 4.8. The majority of the cDNA amplified from the Case 72 tumor contained an *Ava* II site at position 4956 (G49556) and represented the variant, retained allele. Thus, the transcript from the variant allele is stable.

Discussion

Do any of the germline alterations from Case 72 represent disease-associated mutation?

It is possible that the large germline rearrangements identified in Case 72 merely represent a rare variant allele of *BRCA1*. Although the coding regions of *BRCA1* have been studied extensively by many groups and several coding polymorphisms have been identified, the extent of polymorphism within the noncoding regions of *BRCA1* has not yet been studied systematically. It is interesting that the allele containing g.56024 del977 also contains several other sequence variants, including an expanded repeat in intron 7, two SNP's in exon 16, and a 24 bp insertion of unknown origin in intron 15. The novelty of the allele combined with the observation that it is retained in an infiltrating ductal carcinoma from Case 72 is intriguing. Further, Case 72 was diagnosed with a primary colon carcinoma concurrent with the primary infiltrating ductal carcinoma. The rarity of two concurrent primary tumors hints at a potential genetic component that may reside in the *BRCA1* genomic region or elsewhere in the genome.

Most disease-associated *BRCA1* mutations identified from high-risk breast and ovarian cancer families result in premature protein truncation. Only two consensus missense mutations have been identified. Because the tumor-suppressing function of *BRCA1* is not yet understood, the nature of what constitutes a disease-predisposing mutation in *BRCA1* is not easily defined. If any one (or combination) of the alterations in the variant allele from Case 72 led to obvious effects on transcript processing or stability, then we would be able to detect these effects in RNA from the Case 72 breast tumor. Despite analysis of several RT-PCR products amplified from exons flanking

both large rearrangements, no RT-PCR product length or stability differences were observed.

The alterations in the variant allele from Case 72 might also lead to subtle regulatory effects on transcription or alternative splicing of the *BRCA1* message, a process that is not yet fully understood. These possibilities are difficult to test. A more directed approach is to ascertain whether the variant allele from Case 72 is more common in breast cancer cases than in controls. Large sample sizes for both cases and controls are required, however, in order to be informative for rare alleles.

It is also possible that the variants described from the rare allele are in *cis* to a previously uncharacterized small insertion/deletion or point mutation within the *BRCA1* coding sequence. If so, the retention of the variant allele would indicate the presence of the germline mutation and the numerous noncoding region variations would indicate the evolutionary history of the mutated allele. It will be important to screen the coding region of the variant allele to rule out the presence of mutations in *cis* to the noncoding region variants.

The variant allele from Case 72 (as defined by the large expansion of the intron 7 compound microsatellite and the g.56024 del977 deletion) was not observed in 216 other chromosomes from either breast cancer families or breast cancer patients unselected for family history. It is possible that in some populations the variant allele identified in Case 72 is common. At the least, individuals carrying the Case 72 variant allele are represented only once in our series. Samples from the DNA Polymorphism Discovery Resource, which consists of samples from many populations around the world, could help to determine whether the Case 72 variant allele is represented in other populations (Collins et al., 1998). Although this resource would not identify the population(s) in which the variant allele is present, it would help to determine the frequency of the variant allele worldwide.

Characterization of noncoding variation in the BRCA1 genomic region

The variant allele from Case 72 may have implications for the types and frequencies of genomic diversity within the *BRCA1* region. *BRCA1* lies within a large region (~300-500 kb) of recombination suppression on chromosome 17q. A minimal region of nearly complete linkage disequilibrium extends from microsatellite marker D17S1328 (~225 kb proximal to *BRCA1*) to the microsatellite marker D17S1325 (~50-300 kb distal to *BRCA1*) (Liu & Barker, 1999). Multiple studies have shown that two major haplotypes within the *BRCA1* region account for $\geq 90\%$ of chromosomes (Dunning et al, 1997; Liu & Barker, 1999). There are at least two separate reports of multiple large rearrangements within a single allele of *BRCA1* (Rohlf's et al., 1999 and this study). It is possible that in the absence of recombination, other mechanisms exist for generating diversity within the *BRCA1* region.

Most of the variation in the *BRCA1* region is expected to be located in noncoding regions. In a study of the human lipoprotein lipase gene (*LPL*) in which a 9.7 kb region was sequenced from 71 individuals from three different populations, 88 polymorphic sites were identified (Nickerson et al., 1998). Of these sites, 79 were SNP's and only 9 involved insertion/deletion variants. Of the 9 insertion/deletion variants, 8 involved copy number changes in regions known to be polymorphic such as mononucleotide, dinucleotide, and tetranucleotide repeat sequences.

By contrast, the variant allele from Case 72 alone contains two large insertion/deletion alterations involving 100's of bp. Additionally, an insertion of 24 bp of sequence of unknown origin was observed in intron 15. Only exons 8, 15, and 16 and introns 7, 8, and 15 were examined by sequence analysis. Given the high degree of variability in these regions, other small alterations in unsequenced regions of *BRCA1* are likely to exist in the variant allele from Case 72. Either the variant allele from Case 72 represents an evolutionary exception or variability within the *BRCA1* region is much higher than that observed for *LPL*.

The degree of variability within the *BRCA1* region has important implications for assessing the age of founder mutations and the history of human populations. An elevated mutation rate within the *BRCA1* region could implicate *BRCA1* as a target of genomic instability. It will be important to assess the degree and nature of noncoding

variation within the *BRCA1* region in order to determine the effect of rare alleles such as the variant allele retained in the infiltrating ductal carcinoma of Case 72.

CHAPTER FIVE

Summary and Future Directions

My work has contributed to *BRCA1* biology by expanding the germline mutation spectrum in *BRCA1* and by helping to characterize the role of *BRCA1* in sporadic breast cancer. Finally, I have contributed a knowledge base for approaching the study of noncoding variation in the *BRCA1* genomic region. Each of these contributions will be discussed in detail below.

Large germline rearrangements of *BRCA1*

At the inception of the research described in this dissertation, the *BRCA1* complete genomic sequence had been published only recently (Smith et al., 1996). The *BRCA1* genomic region contains a high density of Alu sequence, one of the highest Alu densities observed for genes deposited in GenBank (41.5% of 81 kb). At the same time, it was becoming clear that not all high-risk breast and/or ovarian cancer families, even among those with breast cancer linked to the *BRCA1* region of chromosome 17q21, were explained by small insertion, deletion, or point mutations within *BRCA1*. Large genomic rearrangements in human genes frequently are associated with Alu repetitive elements (Purandare & Patel, 1997; Mazzarella & Schlessinger, 1998). We therefore designed a Long PCR screen to identify the large rearrangements of *BRCA1* that conventional PCR does not detect.

Eleven large germline deletions and one large germline duplication of *BRCA1* have been characterized since that time (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b; Montagna et al., 1999, Carson et al., 1999; Rohlf's et al., 1999; and this study). Of the 24 breakpoints involved, 20 occurred in Alu sequence. Even given that Alu sequence comprises almost half of the *BRCA1* genomic region, there is an excess of breakpoints in Alu sequence. The complex germline rearrangement we identified using Long PCR is the only characterized large germline deletion or duplication in which a nonhomologous mechanism is implicated in the generation of the rearrangement.

The minor contribution of small germline mutations in *BRCA1* to breast cancer in a population-based series of women has been established (Newman et al., 1998). An

important goal for the near future will be to determine the contribution of large germline rearrangements of *BRCA1* to breast cancer in a similar population-based series of women. This is particularly important because large rearrangements are the mutation types most likely to be missed by conventional mutation detection methods. Based on my data from g.12977 ins10 del1039 and g.43368ins6081 frequency in the CBCS population, it is unlikely that these two large germline rearrangements make a large contribution to breast cancer in the American population.

Somatic inactivation of BRCA1

At the start of our research, the prevailing view was that somatic inactivation of *BRCA1* was not common in sporadic tumorigenesis (Futreal et al., 1995; Hosking et al., 1995; Merajver et al., 1995). We reasoned that *BRCA1* might be inactivated by somatic mutations that were undetectable by conventional mutation screening procedures.

I attempted to determine the nature and frequency of large somatic rearrangement of *BRCA1* by analyzing breast tumor gDNA using Long PCR and Southern analysis. Despite Long PCR analysis in 84 tumors and Southern analysis in 46 tumors (including all tumors with genomic loss at any *BRCA1* region marker), no large somatic rearrangement of *BRCA1* was detected. Our research was a critical link in establishing that BRCA1 is not inactivated somatically by either small or large mutation in the majority of sporadic breast tumors.

Over the course of our research an initial observation by a single group documenting reduced *BRCA1* message in sporadic breast tumors was substantiated by several independent researchers (Thompson et al., 1995; Rice et al., 1998; Magdinier et al., 1998). Recently, evaluation of sporadic breast tumors by immunohistochemistry has revealed that expression of BRCA1 protein is reduced or lost in most breast tumors of high histologic grade (Wilson et al., 1999; Jarvis et al., 1998; Taylor et al., 1998; Lee et al., 1999). Using a single large set of 92 breast carcinomas of various histologies, I investigated potential mechanisms for the observed decrease in *BRCA1* expression.

My results indicated that genomic loss at the intragenic marker D17S1323 were associated independently with reduced BRCA1 protein in sporadic breast tumors. This is, to our knowledge, the first description of a mechanism for reduced BRCA1 protein in sporadic breast tumors.

Several potential explanations for the lack of *BRCA1* somatic mutations exist. A trivial explanation is supplied by the difficulty of obtaining pure tumor samples from breast carcinomas. Unlike ovarian or colon carcinomas, breast tumors are more heterogeneous. Breast tumor samples tend to contain a certain percentage of normal stromal cells in addition to tumor cells. In our samples, the percentage of contaminating normal cells ranged from 10 to 40%. If *BRCA1* somatic mutation were a late event in breast tumorigenesis, then any somatic mutations would occur in a small percentage of cells in a given tumor sample and would be more difficult to detect. Two observations argue against this explanation. First, small *BRCA1* somatic mutations are rare in ovarian tumors, a tumor type in which pure tumor samples are more easily obtained. Second, Long PCR detection of large rearrangement favors detection of deletions. From mixing experiments using cloned Long PCR products (primer pair no. 3), deletions of even 1000 bp were significantly favored. Using the g.12977 ins10 del1039 PCR product (4552 bp) and the wild-type PCR product (5581 bp), the g.12977 ins10 del1039 product could be detected in as few as 10% of the cells in a population (data not shown).

A second possibility is that BRCA1 is haploinsufficient. Although *BRCA1* somatic mutations in breast cancer are rare, genomic loss within the *BRCA1* region, as defined by LOH, is common. Further, genomic loss at an intragenic marker nearest the *BRCA1* transcription start site is associated significantly with reduced BRCA1 protein in breast tumors. Under the hypothesis of haploinsufficiency, BRCA1 has multiple roles in breast tumorigenesis, including roles in initiation and progression. Individuals carrying germline mutation of *BRCA1* are apparently wild-type with the exception of early-onset breast and/or ovarian cancer. In germline mutation-carriers, loss of the second, wild-type allele leads to tumor initiation. Thus, BRCA1, by definition, is not haploinsufficient for tumor initiation. However, BRCA1 may be haploinsufficient for

tumor progression. This would explain the high rate of genomic loss in sporadic breast cancer without accompanying somatic mutation. *KIP1*, another tumor suppressor gene in which somatic mutations are rare, is haploinsufficient for the suppression of radiation-induced tumors in mice (Fero et al., 1998). In favor of this idea is the observation that *BRCA1*-mutation carriers have the same risk of developing colon cancer as the general population, but the age of onset in *BRCA1*-mutation carriers is earlier than in the general population (Lin et al., 1999).

A third possibility is that *BRCA1* protein stability is modified in sporadic breast cancer. Because our research used protein staining as a measure of *BRCA1* expression, we cannot exclude the possibility that *BRCA1* message is translated, but the protein is targeted for degradation. It will be important in the future to determine whether *BRCA1* message levels correlate with *BRCA1* protein staining in the breast tumor samples from our collection.

Noncoding variation within the *BRCA1* genomic region

The research described in this dissertation was begun less than 3 years from the initial cloning of *BRCA1*. It was increasingly evident that coding region variation in *BRCA1*, particularly disease-associated mutations, were capable of reflecting the histories of the people in which these variants arose. At the time, the complete genomic sequence of *BRCA1* had been published, but it was not yet clear how variable that sequence was in the general population. Numerous coding and splice site region variants have been identified. There are at least 48 polymorphisms, 43 unclassified variants, and more than 400 distinct mutations (BIC, 1999). The extent of noncoding region variation has not yet been investigated systematically.

The Long PCR screen for large rearrangements was designed using the published genomic sequence (GenBank L78833). In most of the 217 alleles investigated, the general organization of the *BRCA1* region was consistent with the published sequence. More than 48 restriction sites over a region of more than 85 kb were analyzed in high-risk breast and/or ovarian cancer family members and sporadic

breast cancer patients unselected for family history. Only seven variant restriction sites were observed, most of these were seen more than once (Tables 2.3 and 3.2). Another useful measure of noncoding variation was Southern analysis using restriction enzyme-digested gDNA. Using three different enzymes for Southern analysis, no further restriction site variants were revealed.

Despite the general conservation of the *BRCA1* region, at least at the level investigated, a single allele of *BRCA1* contained two large germline rearrangements of unknown significance. Additionally, one rare SNP and a novel 24 bp insertion in *BRCA1* intron 15 were identified in *cis* to the two large germline rearrangements. Interestingly, this allele was retained in an infiltrating ductal carcinoma from Case 72. While no obviously deleterious effects of the allele on *BRCA1* expression were identified, the novelty of the allele, combined with its retention in a ductal carcinoma, is intriguing as the presence of two concurrent primary carcinomas in a single individual is rare.

The difficulty in classifying the status of the Case 72 large germline rearrangements highlights an area of increasing interest: noncoding variation in the human genome. The determination of the amount and the degree to which individual genomes vary will be critical to determining the relevance to human disease for variant alleles such as the rare allele described from Case 72.

Sequence comparisons of large contiguous genomic regions (>30 kb) between humans and rodents have revealed that coding regions are relatively well-conserved (Koop, 1995; Hardison et al., 1997). However, noncoding and intergenic regions demonstrate considerable variation in the degree of sequence conservation. Sequences upstream of the first exon and in the first intron are among the most highly conserved noncoding regions. Numerous studies have suggested the relevance of cross-species sequence comparisons for the identification of regulatory elements (Koop & Hood, 1994; Oeltjen et al., 1997; Ansari-Lari et al., 1998). Regulatory elements identified by such an approach would be potential targets of noncoding variants such as the rare variant allele of *BRCA1* identified in Case 72.

In summary, my research specifically addressed the types and frequencies of large genomic rearrangements responsible for inactivation of *BRCA1*. I characterized the types of large germline rearrangements that occur within the *BRCA1* region and investigated the contribution of two large germline rearrangements to breast cancer in a population-based series of breast cancer patients. Although the structure of the *BRCA1* genomic region was generally well-conserved, one variant allele containing multiple large scale alterations of the *BRCA1* genomic region was identified. The existence of a variant with multiple large rearrangements in *cis* indicates that an investigation of the types and frequencies of noncoding variation in the *BRCA1* genomic region may yield a broader understanding of noncoding variation within the human genome.

In order to determine whether *BRCA1* is inactivated somatically by large rearrangement of *BRCA1*, I analyzed 92 breast carcinomas for genomic loss in the *BRCA1* region of chromosome 17q. I investigated genomic loss using a combination of loss of heterozygosity (LOH), Long PCR, and Southern analysis. Although two large germline rearrangements were detected in our breast tumor series using Long PCR and Southern analysis, no large somatic rearrangements were identified. The results of LOH analysis were correlated with *BRCA1* protein immunohistochemistry data generated by Rachel Gonzalez-Hernandez in order to test whether LOH is a mechanism for inactivating *BRCA1* in sporadic breast cancer. Reduced *BRCA1* protein in sporadic breast carcinomas was significantly associated with loss of the most 5' *BRCA1* intragenic marker. LOH at the remaining, more 3' *BRCA1* intragenic markers was not associated with reduced *BRCA1* protein. Interestingly, 57% breast carcinomas retaining all three *BRCA1* intragenic markers showed reduced *BRCA1* protein. Thus, there are likely other mechanisms for inactivation of *BRCA1* in sporadic breast cancer.

LIST OF REFERENCES

- American Cancer Society, Cancer Facts and Figures. 1997. Atlanta, GA:ACS.
- Anbazhagan R, Fuji H and Gabrielson E. 1999. Microsatellite instability is uncommon in breast cancer. *Clin Cancer Res*, 839-844.
- Anderson LA, Friedman L, Osborne-Lawrence S, Lynch E, Weissenbach J, Bowcock A and King M-C. 1993. High-density genetic map of the BRCA1 region of chromosome 17q12-q21. *Genomics* 17: 618-23.
- Anderson SF, Schlegel BP, Nakajima T, Wolpin ES and Parvin JD. 1998. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet*, 19: 254-256.
- Ansari-Lari MA, Oeltjen JC, Schwartz S, Zhang Z, Muzny DM, Lu J, Gorrell JH, Chinault AC, Belmont JW, Miller W and Gibbs RA. 1998. Comparative sequence analysis of a gene-rich cluster at human chromosome 12p13 and its syntenic region in mouse chromosome 6. *Genome Res*, 8: 29-40.
- Antonarakis SE and the Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Human Mutation*, 11:1-3.
- Barker DF, Liu X and Almeida ER. 1996. The BRCA1 and 1A1.3B promoters are parallel elements of a genomic duplication at 17q21. *Genomics*, 38: 215-222.
- Baylin SB, Herman JG, Graff JR, Vertino PM and Issa J-P. 1998. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv. Cancer Res*, 72: 141-196.
- Beckmann MW, Picard F, An HX, van Roeyen CR, Dominik SI, Mosny DS, Schnurch HG, Bender HG and Niederacher D. 1996. Clinical impact of detection of loss of heterozygosity of BRCA1 and BRCA2 markers in sporadic breast cancer. *Br J Cancer*, 73: 1220-6.
- Benachenchou N, Guiral S, Gorska-Flipot I, Labuda D and Sinnott D. 1999. Frequent loss of heterozygosity at the DNA mismatch-repair loci hMLH1 and hMSH3 in sporadic breast cancer. *Br J Cancer*, 79: 1012-1017.
- Berchuck A, Heron K-A, Carney ME, Lancaster JM, Fraser EG, Vinson VL, Deffenbaugh AM, Miron A, Marks JR, Futreal PA and Frank TS. 1998. Frequency of germline and somatic *BRCA1* mutations in ovarian cancer. *Clin. Cancer Res*, 4: 2433-2437.

Bergthorsson JT, Egilsson V, Gudmundsson J, Arason A and Invarsson S. 1995. Identification of a breast tumor with microsatellite instability in a potential carrier of the hereditary non-polyposis colon cancer trait. *Clin Genet*, 47: 305-310.

Bernard-Gallon DJ, De Latour MP, Rio PG, Penault-Llorca FM, Favvy DA, Hizez C, Chassagne J and Bignon YJ. 1999. Subcellular localization of BRCA1 protein in sporadic breast carcinoma with or without allelic loss of BRCA1 gene. *Int J Oncol*, 14: 653-61.

Bieche I, and Lidereau R. 1995. Genetic alterations in breast cancer. *Gen. Chromosom. Cancer*, 14: 227-251.

Bignell GR, Barfoot R, Seal S, Collins N, Warren W and Stratton MR. 1998. Low frequency of somatic mutations in the LKB1/Peutz-Jeghers syndrome gene in sporadic breast cancer. *Cancer Res*, 58: 1384-1386.

Boyd J, Rhei E, Federici MG, Borgen PI, Watson P, Franklin B, Karr B, Lynch J, Lemon SJ and Lynch HT. 1999. Male breast cancer in the hereditary nonpolyposis colorectal syndrome. *Breast Cancer Res Treat*, 53: 87-91.

Breast Cancer Information Core. 1999.

http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/BIC,

Breslow NE and Day NE 1993. *Statistical Methods in Cancer Research: The Analysis of Case-Control Studies*. Vol. 1 Breslow NE; Day NE New York : Oxford University Press.

Brown MA, Xu CF, Nicolai H, Griffiths B, Chambers JA, Black D and Solomon E. 1996. The 5' end of the BRCA1 gene lies within a duplicated region of human chromosome 17q21. *Oncogene*, 12: 2507-2513.

Brzovic PS, Meza J, King M-C and Klevit RE. 1998. The cancer-predisposing mutation C61G disrupts homodimer formation in the NH₂-terminal BRCA1 RING finger domain. *J Biol Chem*, 273: 7795-7799.

Callebaut I and Mornon J-P. 1997. From BRCA1 to RAP1: A widespread BRCT module closely associated with DNA repair. *FEBS Letters*, 400: 25-30.

Carson N, Gilpin C, Hunter A, Allanson J and Aubry H. 1999. An in-frame deletion of BRCA1 exon 20 in a family with early onset breast cancer and ovarian cancer. *Am J Hum Genet*, 65 Suppl: A1610.

Castilla LH, Couch FJ, Erdos MR, Hoskins KF, Calzone K, Garber JE, Boyd J, Lubin MB, Deshano ML, Brody LC, Collins FS and Weber BL. 1994. *Nat Genet*, 8: 387-391.

- Catteau A, Harris WH, Xu C-F and Solomon E. 1999. Methylation of the *BRCA1* promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. *Oncogene*, 18: 1957-1965.
- Chapman MS and Verma IM. 1996. Transcriptional activation by BRCA1. *Nature*, 382: 678-679.
- Christians RC and Hanawalt PC. 1992. Inhibition of transcription and strand-specific DNA repair by alpha-amanitin in Chinese hamster ovary cells. *Mutat Res*, 274: 93-101.
- Coles C, Condie A, Chetty U, Steel CM, Evans HJ and Prosser J. 1992. p53 mutations in breast cancer. *Cancer Res*, 52: 5291-5298.
- Collins FS, Brooks LD and Chakravarti A. 1998. A DNA Polymorphism Discovery Resource for research on human genetic variation. *Genome Res*, 8: 1229-1231.
- Devilee P, and Cornelisse CJ. 1994. Somatic genetic changes in human breast cancer. *Biochem. Biophys. Acta*, 1198: 113-130.
- Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, and Ruschhoff J. 1997. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res*, 57: 4749-56.
- Dunning AM, Chiano M, Smith NR, Dearden J, Gore M, Oakes S, Wilson C, Stratton M, Peto J, Easton D, Clayton D and Ponder BAJ. 1997. Common BRCA1 variants and susceptibility to breast and ovarian cancer in the general population. *Hum Mol Genet*, 6: 285-289.
- Fero ML, Randel E, Gurley K, Roberts JM and Kemp CJ. 1998. The murine gene *p27^{kip1}* is haploinsufficient for tumour suppression. *Nature*, 396: 177-180.
- Freemont PS. 1993. The RING finger: A novel protein sequence motif related to the zinc finger. *Ann NY Acad Sci*, 684: 174-192.
- Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch E, Rowell SE, King, M-C. 1994. Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat Genet* 8:399-404.
- Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, Eddington K, McClure M, Frye C, Weaver-Feldhaus J, Ding W, Gholami Z, Söderkvist P, Terry L, Skolnick MH, Kamb A and Wiseman R. 1994. *BRCA1* mutations in primary breast and ovarian carcinomas. *Science* (Washington DC), 266: 120-122.

- Gowen LC, Avrutskaya AV, Latour AM, Koller BH and Leadon SA. 1998. BRCA1 required for transcription-coupled repair of oxidative DNA damage. *Science*, 281: 1009-1012.
- Greenblatt MS, Bennett WP, Hollstein M and Harris C. 1994. Mutations in the p53 tumor suppressor gene: Clues to cancer etiology and molecular pathogenesis. *Cancer Res*, 54: 4855-4878.
- Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B and King M-C. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science*, 250:1684-1689.
- Hanawalt PC. 1994. Transcription-coupled repair and human disease. *Science*, 266: 1957-1958.
- Hardison RC, Oeltjen J and Miller W. 1997. Long human mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res*, 7: 959-966.
- Hosking L, Trowsdale J, Nicolai H, Solomon E, Foulkes W, Stamp G, Signer E and Jeffreys A. 1995. A somatic *BRCA1* mutation in an ovarian tumor. *Nature Genet.*, 9: 343-344.
- Jarvis EM, Kirk JA and Clarke CL. 1998. Loss of nuclear BRCA1 expression in breast cancers is associated with a highly proliferative tumor phenotype. *Cancer Genet. Cytogenet.*, 101: 109-115.
- Jones PA and Laird PW. 1999. Cancer epigenetics comes of age. *Nature Genet*, 21: 163-167.
- Khoo U-S, Ozcelik H, Cheung ANY, Chow LWC, Ngan HYS, Done SJ, Liang ACT, Chan VWY, Au GKH, Hg W-F, Poon CSP, Leung Y-F, Loong F, Ip P, Chan GSW, Andrulis IL, Lu J and Ho FCS. 1999. Somatic mutations in the BRCA1 gene in Chinese sporadic breast and ovarian cancer. *Oncogene*, 18: 4643-4646.
- Koonin Ev, Altschul SF and Bork P. 1996. BRCA1 protein products: functional motifs. *Nat Genet*, 13: 266-267.
- Koop BF. 1995. Human and rodent DNA sequence comparisons: A mosaic model of genomic evolution. *Trends Genet*, 11: 367-371.
- Koop BF and Hood L. 1994. Striking sequence similarity over almost 100 kilobases of human and mouse T-cell receptor DNA. *Nature Genet*, 7: 48-53.

- Krajinovic M, Richer C, Gorska-Flipot I, Gaboury L, Novakovic I, Labuda D and Sinnett D. 1998. Genomic loci susceptible to replication errors in cancer cells. *Br J Cancer*, 78: 981-985.
- Lancaster JM, Wooster R, Mangion J, Phelan CM, Cochran C, Gumbs C, Seal S, Barfoot R, Collins N, Bignell G, Patel S, Hamoudi R, Larsson C, Wiseman R, Berchuck A, Iglehart JD, Marks JR, Ashworth A, Stratton M and Futreal PA. 1996. BRCA2 mutations in primary breast and ovarian cancers. *Nat Genet*, 13: 238-240.
- Leadon SA and Lawrence DA. 1992. Strand-selective repair of DNA damage in the yeast GAL7 gene requires RNA polymerase II. *J Biol Chem*, 267: 23175-23182.
- Lee W-Y, Jin Y-T, Chang T-W, Lin P-W and Su I-J. 1999. Immunolocalization of BRCA1 protein in normal breast tissue and sporadic invasive ductal carcinomas: a correlation with other biological parameters. *Histopathology*. 34: 106-112.
- Li J, Yen C, Liaw D, Podyspanina K, Bose S, Wang SI, Puc J, Miliareis C, Rodgers L, McCombie R, Bigner SH, Giovanellis BC, Ittmann M, Tyoko B, Hibshoosh H, Wigler MH, Parsons R. 1997. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*, 275: 1943-1947.
- Lin KM, Ternent CA, Adams DR, Thorson AG, Blatchford GJ, Christensen MA, Watson P and Lynch HT. 1999. Colorectal cancer in hereditary breast cancer kindreds. *Dis Col Rect*, 42: 1041-1045.
- Lovering R, Hanson IM, Borden KL, Martin S, O'Reilly NJ, Evan GI, Rahman D, Pappin DJ, Trowsdale J and Freemont PS. 1993. Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc Natl Acad Sci USA*, 90: 2112-2116.
- Lu M, Conzen SD, Cole CN and Arrick BA. 1996. Characterization of functional messenger RNA splice variants of BRCA1 expressed in nonmalignant and tumor-derived breast cells. *Cancer Res*, 56: 4578-4581. 1-16.
- Magdinier F, Ribieras S, Lenoir GM, Frappart L and Dante R. 1998. Down-regulation of BRCA1 in human sporadic breast cancer: analysis of DNA methylation patterns of the putative promoter region. *Oncogene*, 17: 3169-3176.
- Mancini DN, Rodenhiser DI, Ainsworth PJ, O'Malley FP, Singh S, Xing W and Archer TK. 1998. CpG methylation within the 5' regulatory region of the *BRCA1* gene is tumor specific and includes a putative CREB binding site. *Oncogene*, 16: 1161-1169.
- Matsushima M, Kobayashi K, Emi M, Saito H, Saito J, Suimori K and Nakamura Y. 1995. Mutation analysis of the *BRCA1* gene in 76 Japanese ovarian cancer patients: four

germline mutations, but no evidence of somatic mutation. *Hum Mol Genet*, 4: 1953-1956.

Mazzarella R and Schlessinger D. 1998. Pathological consequences of sequence duplications in the human genome. *Genome Res*, 8: 1007-1021.

Mellon I, Spivak G and Hanawalt PC. 1987. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*, 51: 241-249.

Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C and Frank TS. 1995. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. *Nature Genet.*, 9:439-443.

Meza JE, Brzovic, King M-C and Klevit RE. 1999. Mapping the functional domains of BRCA1. Interaction of the RING finger domains of BRCA1 and BARD1. *J Biol Chem*, 274: 5659-5665.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hyssey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C, Bogden R, Dayananth P, Ward J, Tonin P, Narod S, Bristow PK, Norris FH, Helvering L, Morrisson P, Rosteck P, Lai M, Barrett JC, Lewis C, Neuhausen S, Albright LC, Goldgar D, Wiseman R, Kamb A and Skonick M.H. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science (Washington DC)*, 266: 66-71.

Monnat RJ, Hackmann AFM and Chiaverott, TA. 1992. Nucleotide sequence analysis of human hypoxanthine phosphoribosyltransferase (*HPRT*) gene deletions. *Genomics*, 13: 777-787.

Montagna M, Santacatterina M, Torri A, Menin C, Zullato D, Chieco-Bianchi L and D'Andrea E. 1999. Identification of a 3 kb Alu-mediated *BRCA1* gene rearrangement in two breast/ovarian cancer families. *Oncogene*, 18: 4160-4165.

Monteiro ANA, August A and Hanafusa H. 1996. Evidence for a transcriptional activation function of *BRCA1* C-terminal region. *Proc Natl Acad Sci USA*, 93: 13595-13599.

Narod SA, Ford D, Devilee P, Barkardottir RB, Lynch HT, Smith SA, Ponder BJ, Weber BL, Garber JE, Birch JM, Cornelis RS, Kelsell DP, Spurr NK, Smyth E, Haites N, Sobol H, Bignon Y-J, Chang-Claude J, Hamann U, Lindblom A, Borg A, Piver S, Gallion HH, Struewing JP, Whittemore A, Tonin P, Goldgar DE, Easton DF, the Breast Cancer Linkage Consortium. 1995. An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am J Hum Genet* 56: 254-264.

Neuhausen SL, Mazoyer S, Friedman L, Stratton M, Offit K, Caligo A, Tomlinson G, Cannon-Albright L, Bishop T, Kelsell D, Solomon E, Weber B, Couch F, Struewing J, Tonin P, Durocher F, Narod S, Skolnick MH, Lenoir G, Serova O, Ponder B, Stoppa-Lyonnet D, Easton D, King M-C and Goldgar DE. 1996. Haplotype and phenotype analysis of six recurrent BRCA1 mutations in 61 families: results of an international study. *Am J Hum Genet*, 58: 271-80.

Newman, B., Austin, M, Lee, M and King, M-C. 1988. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc. Natl. Acad. Sci. USA*, 85: 3044-3048.

Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, Liu ET. 1995. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res Treat*, 35:51-60.

Newman B, Mu H, Butler LM, Millikan RC, Moorman PG and King M-C. 1998. Frequency of breast cancer attributable to BRCA1 in a population-based series of American women. *JAMA*, 279: 915-921.

Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengård J, Salomaa V, Vartiainen E, Boerwinkle E and Sing CF. 1998. DNA sequence diversity in a 9.7 kb region of the human lipoprotein lipase gene. *Nature Genet*, 18: 233-240.

Niederacher D, Picard F, van Roeyen C, An HX, Bender HG and Beckmann MW. 1997. Patterns of allelic loss on chromosome 17 in sporadic breast carcinomas detected by fluorescent-labeled microsatellite analysis. *Genes Chromosom Cancer* 18:181-192.

Oeltjen JC, Malley TM, Muzny DM, Miller W, Gibbs RA and Belmont JW. 1997. Large scale comparative sequence analysis of the human and murine Bruton's tyrosine kinase loci reveals conserved regulatory domains. *Genome Res*, 7: 315-329.

Osborne RJ, Merlo GR, Mitsudomi T, Venesio T, Liscia DS, Cappa AP, Chiba I, Takahashi T, Nau MM Callahan R and Minna JD. 1991. Mutations in the p53 gene in primary human breast cancers. *Cancer Res*, 51: 6194-6198.

Peltomaki P. 1997. DNA mismatch repair gene mutations in human cancer. *Environ Health Perspect*, 105 Suppl: 775-780.

Perren A, Weng L-P, Boag AH, Ziebold U, Thakore K, Dahia PLM, Komminoth P, Lees JA, Mulligan LM, Mutter GL and Eng C. 1999. Immunohistochemical evidence of loss of PTEN expression in primary ductal adenocarcinomas of the breast. *Am J Pathol*, 155: 1253-1260.

Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drüsedau M, Hogervorst FBL, Hageman S, Arts PJW, Ligtenberg MJL, Meijers-Heijboer H, Klijn, JGM, Vasen HFA, Cornelisse CJ, van't Veer LJ, Egbert-Bakker, van Ommen G-J.B, Devilee P. 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet*, 17: 341-345.

Puget N, Torchard D, Serova-Sinilkova OM, Lynch HT, Feunteun J, Lenoir GM, Mazoyer, S. 1997. A 1-kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. *Cancer Res*, 57: 828-831.

Puget N, Stoppa-Lyonnet D, Sinilkova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S. 1999a. Screening for germ-line rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. *Cancer Res*, 59: 455-461.

Puget N, Sinilkova OM, Stoppa-Lyonnet D, Audoynaud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer, S. 1999b. An alu-mediated 6-kb duplication in the *BRCA1* gene: a new founder mutation? *Am J Hum Genet*, 64: 300-302.

Purandare SM and Patel PI. 1997. Recombination hotspots and human disease. *Genome Res*. 7: 773-786.

Rebbeck TR, Couch FJ, Kant J, DeShano M, Peng Y, Chen K, Garber JE, Weber BL. 1997. Genetic heterogeneity in hereditary breast cancer: Role of *BRCA1* and *BRCA2*. *Am J Hu. Genet* 60: 547-553.

Rice JC, Massey-Brown K.S and Futscher BW. 1998. Aberrant methylation of the *BRCA1* CpG island promoter is associated with decreased *BRCA1* mRNA in sporadic breast cancer cells. *Oncogene*, 17: 1807-1812.

Rio PG, Pernin D, Bay JO, Albuissou E, Kwiatkowski F, De Latour M, Bernard-Gallon DJ and Bignon YJ. 1998. Loss of heterozygosity of *BRCA1*, *BRCA2* and *ATM* genes in sporadic invasive ductal breast carcinoma. *Int J Oncol*, 13: 849-853.

Risinger JI, Barrett JC, Watson P, Lynch HT and Boyd J. 1996. Molecular genetic evidence of the occurrence of breast cancer as an integral tumor in patients with the hereditary nonpolyposis colorectal carcinoma syndrome. *Cancer*, 77: 1836-1843.

Rhei E, Kang L, Bogomolny F, Federici MG, Borgen PI and Boyd J. 1997. Mutation analysis of the putative tumor suppressor gene *PTEN/MMAC1* in primary breast carcinomas. *Cancer Res*, 57: 3657-3659.

Rohlf's EM, Yang Q, Skrzynia C, Graham ML and Silverman LM. 1999. Complex rearrangement in *BRCA1* results in the in-frame loss of exons in the BRCT domain. *Am J Hum Genet*, 65 *Suppl*: A1610.

- Saurin AJ, Borden KL, Boddy MN and Freemont PS. 1996. Does this have a familiar RING? *Trends Biochem Sci*, 21: 208-214.
- Schubert EL, Lee MK, Mefford HC, Argonza RH, Morrow JE, Hull J, Dann JL, King, M-C. 1997. BRCA2 in American families with four or more cases of breast or ovarian cancer: Recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet* 60: 1031-1040.
- Scully R, Chen J, Plug A, Xiao Y, Weaver D, Feunteun J, Ashley T and Livingston DM. 1997a. Association of BRCA1 with RAD51 in mitotic and meiotic cells. *Cell*, 88: 265-275.
- Scully R, Chen J, Ochs RL, Keegan K, Hoekstra, Feunteun J and Livingston DM. 1997b. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell*, 90: 425-435.
- Scully R, Anderson SF, Chao DM, Wei W, YE L, Young RA, Livingston DM and Parvin JD. 1997c. BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc Natl Acad Sci USA*, 94: 5605-5610.
- Sherr CJ. 1996. Cancer cell cycles. *Science*, 274: 1672-1677.
- Silva JM, Gonzalez R, Provencio M, Dominguez G, Garcia JM, Gallego I, Palacios J, Espana P and Bonilla F. 1999. Loss of heterozygosity in BRCA1 and BRCA2 markers and high-grade malignancy in breast cancer. *Breast Cancer Res Treat*. 53: 9-17.
- Smith SA, Easton DF, Evans DG and Ponder BA. 1992. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nature Genet.*, 2: 128-131.
- Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M, Taylor M, Hood L and King M-C. 1996. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene *BRCA1*. *Genome Res.*, 6: 1029-1049.
- Somasundaram X, Zhang H, Zeng Y-X, Houvras Y, Peng Y, Zhang H, Wu GS, Licht JD, Weber BL and El-Deiry WS. 1997. Arrest of the cell cycle by tumour-suppressor BRCA1 requires the CDK-inhibitor p21^{WAF1/CIP1}. *Nature*, 389: 187-190.
- Steck PA, Pershouse MA, Jasser SA, Yung WKA, Lin H, Ligon AH, Langford LA, Baumgard ML, Hattier T, Davis T, Frye C, Hu R, Swedlund B, Teng DHF, Tavtigian SV. 1997. Identification of a candidate tumor suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet*, 15: 356-362.

- Stratton M. 1996. Molecular biology of hereditary breast cancer. *Hum Mol Genet*, 5: 515-518.
- Sun H, Treco D, Schultes NP and Szostak JW. 1989. Double-strand breaks at an initiation site for meiotic gene conversion. *Nature*, 338: 87-90.
- Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. 1997. Identification of a 14 kb deletion involving the promoter region of BRCA1 in a breast cancer family. *Hum Mol Genet*, 6: 1513-1515.
- Szabo CI and King M-C. 1995. Inherited breast and ovarian cancer. *Hum Mol Genet*, 4: 1811-1817.
- Szabo CI and King M-C. 1997. Population genetics of BRCA1 and BRCA2. *Am J Hum Genet*, 60: ?.
- Takahashi H, Behbakht K, McGover, PE, Chiu H-C, Couch FJ, Weber B, Friedman LS, King M-C, Furusato M, LiVlosi VA, Menzin AW, Liu PC, Benjamin I, Morgan MA, King SA, Rebane BA, Cardonick A, Mikuta JJ, Rubin SC and Boyd J. 1995. Mutation analysis of the *BRCA1* gene in ovarian cancers. *Cancer Res.*, 55: 2998-3002. 1995.
- Takata M et al. 1998. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J*, 17: 5497-5508.
- Taylor J, Lymboura M, Pace PE, A'Hern RP, Desai AJ, Shousha S, Coombes RC and Ali S. 1998. An important role for BRCA1 in breast cancer progression is indicated by its loss in a large proportion of non-familial breast cancers. *Int. J. Cancer*, 79: 334-342.
- Thakur S and Croce CM. 1999. Positive regulation of the BRCA1 promoter. *J Biol Chem*, 274: 8837-43
- Thompson ME, Jensen RA, Obermiller PS, Page DL and Holt JT. 1995. Decreased expression of *BRCA1* accelerates growth and is often present during sporadic breast cancer progression. *Nature Genet.*, 9: 444-450.
- Van der Vliet PC. 1996. Roles of transcription factors in DNA replication. In *DNA replication in eukaryotic cells*, pp87-118. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Wilson CA, Payton MN, Elliott GS, Buas FW, Cajulis EE, Grosshans D, Ramos L, Reese DM, Slamon DJ and Calzone FJ. 1997. Subcellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-delta11b. *Oncogene*, 14: 1-16.

Wilson CA, Ramos L, Villasenor MR, Anders KH, Press MF, Clarke K, Karlan B, Chen JJ, Scully R, Livingston D, Zuch RH, Kanter MH, Cohen S, Calzone F.J and Slamon DJ. 1999. Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. *Nature Genet.*, 21: 236-240.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G, Barfoot R, Hamoudi R, Patel S, Rice C, Biggs P, Hashim Y, Smith A, Connor F, Arason A, Gudmundsson J, Ficenec D, Kelsell D, Ford D, Tonin P, Bishop DT, Spurr NK, Ponder BAJ, Eeles R, Peto J, Devilee P, Cornelisse C, Lynch H, Narod S, Lenoir G, Egilsson V, Barkadottir RB, Easton DF, Bentley DR, Futreal PA, Ashworth A, and Stratton MR. 1995. Identification of the breast cancer susceptibility gene BRCA2. *Nature*, 378: 789-792.

Xu C-F, Chambers JA, Nicolai H, Brown MA, Hujeriat Y, Mohammed S, Hodgson S, Kelsell DP, Spurr NK, Bishop DT and Solomon E. 1997a. Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families. *Genes, Chromosomes, Cancer*, 18:102-110.

Xu, C-F 1997. Complex regulation of the BRCA1 gene. *J Biol Chem*, 272: 20994-20997.

Xu X, Weaver Z, Linke SP, Li C, Gotay J, Wang XW, Harris CC, Ried T and Deng CX. 1999. Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell*, 3: 389-95.

Yarden RI and Brody LC. 1999. BRCA1 interacts with components of the histone deacetylase complex. *Proc Natl Acad Sci USA*, 96: 4983-4988.

APPENDIX ONE

Complex germline rearrangement of BRCA1 associated with breast and ovarian cancer.

Shannon R. Payne¹, Beth Newman^{2,3}, and Mary-Claire King⁴

¹ Department of Genetics, University of Washington, Seattle, WA, USA 98195

² Lineberger Cancer Center, University of North Carolina, Chapel Hill, North Carolina
27599

³ Current Address: School of Public Health, Queensland University of Technology,
Brisbane, Queensland, 4059, Australia

⁴ Departments of Genetics & Medicine, Division of Medical Genetics, University of
Washington, Seattle, WA, USA 98195

Corresponding Author: Mary-Claire King

Telephone: 206-616-4294

Fax: 206-616-4295

e-mail: mcking@u.washington.edu

Running Title: Complex rearrangement of BRCA1

ABSTRACT

Germline mutations in *BRCA1* predispose to breast and ovarian cancer. Most germline *BRCA1* mutations are small insertions, deletions or single base pair (bp) substitutions. These mutation classes are rarely found as somatic mutations in *BRCA1*. Conversely, somatic deletions of multiple megabase pairs (Mb) including *BRCA1*, as reflected by loss of heterozygosity, occur frequently in both inherited and sporadic breast and ovarian cancer. In order to determine whether deletions or rearrangements of hundreds to thousands of bp might contribute to inherited mutation in *BRCA1*, we developed a Long PCR strategy for screening the entire genomic *BRCA1* locus in high-risk families. We evaluated genomic DNA from one high-risk family of Western European ancestry with *BRCA1*-linked cancer in which no genomic mutations had been detected using conventional methods. Long PCR revealed a complex mutation, g.12977 ins10 del 1039 (based on GenBank L78833) comprising an inverted duplication and deletion in *BRCA1* that removes portions of exon 3 and intron 3, including the 5' splice site for intron 3. As a result of the deletion, exon 3 is skipped, leading to a truncated protein and disease predisposition. Unlike previously reported large germline deletions in *BRCA1*, neither breakpoint resides within an Alu element. The g.12977 ins10 del1039 mutation was not detected among eleven other breast cancer families, nor among 406 breast cancer patients unselected for family history.

Keywords: BRCA1, large deletion, RNA splicing, exon skipping

Germline protein-truncating mutations in *BRCA1* cause hereditary predisposition to breast and ovarian cancer (Hall et al., 1990; Ford et al., 1994). *BRCA1* mutations detectable by commonly employed PCR-based techniques are small insertions, deletions or single base pair (bp) substitutions which comprise most germline mutations in high-risk families (Breast Cancer Information Core, 1999), but rarely occur as somatic mutations (Futreal et al., 1994; Hosking et al., 1995; Matsushima et al., 1995; Merajver et al., 1995; Takahashi et al., 1995; Berchuck et al., 1998). Conversely, megabasepair (Mb) somatic deletions including *BRCA1*, as reflected by loss of heterozygosity (LOH), occur frequently in both inherited and sporadic breast and ovarian cancer (Devilee et al., 1994; Bieche et al., 1995). We have developed a protocol to screen for an intermediate class of mutations, involving deletions or rearrangements of hundreds to thousands of bp. Several large deletions and one large duplication involving Alu repetitive elements have been reported previously as germline mutations in *BRCA1* (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b).

We report here the identification and characterization of a large genomic deletion in a family with inherited breast and ovarian cancer. The mutation results in skipping of exon 3 in the mature *BRCA1* message. Although the genomic mutation was not detectable using conventional PCR-based *BRCA1* screening procedures, it was identified using a Long PCR strategy designed to detect intermediate length genomic deletions and rearrangements.

Previous analysis of lymphoblast cDNA from Family 5 revealed a transcript lacking exon 3 in several *BRCA1*-linked family members (Friedman et al., 1994; Fig.1). Deletion of exon 3 maintains the reading frame, but creates a stop at codon 27. Splice junctions of exons 2, 3, and 5 (the exon joined to the 3' splice site of exon 3) were sequenced from PCR products using primers immediately flanking the exons, yielding only wild-type sequence (Friedman et al., 1994). The genomic basis of the variant remained unknown.

For the analysis described here, gDNA from family members was amplified using the fifteen PCR primer pairs indicated in Table 1. Primers were designed from

BRCA1 genomic sequence (GenBank L78833). Primer pairs were designed to amplify 3 to 11 kb genomic fragments with at least 1 kb of overlap. PCR products were evaluated for differences from predicted size before and after digestion with appropriate restriction enzymes.

Using primer pair 3 designed to amplify a 5531 bp product encompassing exon 3, a variant Long PCR product co-segregated with the *BRCA1*-linked haplotype of Family 5 (Fig. 1, 2). All family members heterozygous for transcripts lacking exon 3 were also heterozygous for the variant PCR product.

Wild-type and variant Long PCR products were cloned and sequenced from multiple family members. Sequence of all variant clones revealed the complex mutation shown in Figure 3. Based on *BRCA1* genomic sequence L78833, the mutation is a 10bp inverted duplication of 12965-12974 from *BRCA1* exon 3 (bp 210-219 of *BRCA1* cDNA, HSU14680) and deletion of 1039bp. The notation for the mutation (Antonarakis et al., 1998) is g.12977 ins10 del1039 based on GenBank L78833. The net deletion is 1029 bp, yielding an amplified product of 4502 bp which is consistent with the electrophoretic mobility of the variant PCR product. The 1039 bp genomic deletion in Family 5 results in skipping of *BRCA1* exon 3 in the mRNA and premature protein truncation at codon 27.

Three other families with breast cancer linked to *BRCA1* were tested for the g.12977 ins10 del1039 mutation. None of the families contained the g.12977 ins10 del1039 mutation. Families with multiple cases of breast and or ovarian cancer and with negative LOD scores for linkage to *BRCA1* might nonetheless harbor mutations in one of these genes if some cases are sporadic (ie. phenocopies). Seven such families were screened for g.12977 ins10 del1039. All were wild-type at this site. Finally, we investigated the contribution of the g.12977 ins10 del1039 mutation to breast cancer in a population-based series of 242 white and 164 African-American breast cancer patients unselected for family history. Patients were ascertained previously as part of the Carolina Breast Cancer Study (Newman et al., 1995; Newman et al., 1998). A PCR primer pair was designed to detect a 514 bp breakpoint junction fragment for the g.12977 ins10 del1039 mutation. Primers used to amplify the breakpoint junction

fragment were: (forward) 5'-TTT-TTC-TCC-CCC-CCT-ACC-CTG-3'; (reverse) 5'-GCT-CAG-CAT-TTG-TTA-CTC-AAG-CTG-3'. No mutation carriers were detected in 406 cases.

The mutation in Family 5 differs from previously reported large germline deletions in *BRCA1*. Many large *BRCA1* mutations involve Alu sequences (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b). For two of these, both the 5' and 3' breakpoints reside in Alu sequence. In one of these families, a 1 kb deletion of exon 17 is mediated by oppositely oriented Alu elements, removing most parts of both Alu elements from the mutant allele (Puget et al., 1997). In another family, a 14 kb deletion involving the *BRCA1* promoter region creates a new Alu element (Swensen et al., 1997). In contrast, the 5' breakpoint for the Family 5 mutation occurs within *BRCA1* exon 3 and the 3' breakpoint occurs in intron 3 at position +1008. Neither of these breakpoints occurs in an Alu element. Further, there is no more than 4 bp identity between sequence near the 5' breakpoint and sequence near the 3' breakpoint, suggesting that this mutation is the result of a nonhomologous event.

Combinations of inversions and deletions have been seen in both somatic and germline mutations of several genes including *TP53* (Greenblatt et al., 1996) and *Factor IX* (Ketterling et al., 1994; Sommer, 1995). Studies of somatic mutations in *HPRT* reveal alterations with interesting similarity to the naturally occurring *BRCA1* germline mutation of Family 5. Somatic deletions in *HPRT* often include inserted bases at the breakpoint junctions that are inverted complements of sequences found at the breakpoints (Rainville et al., 1995). In *HPRT*, topoisomerase I and topoisomerase II have been implicated in generating the free DNA ends that may result in these nonhomologous recombination events (Rainville et al., 1995; Monnat et al., 1992). Topoisomerase II sites also occur near rearrangement breakpoints in the *Dystrophin* gene (Hu et al., 1991) and ring chromosome 21 (Wong et al., 1989). *In vitro* and *in vivo* evidence for the role of topoisomerase II in nonhomologous recombination is well established in prokaryotes (O'Connor et al., 1985; Ikeda, 1986) and a role for vertebrate topoisomerase II in nonhomologous recombination has been demonstrated *in vitro* (Bae et al., 1988). Topoisomerase II boxes are in close proximity to the 3' deletion breakpoint

of the complex mutation in Family 5. Sequences matching 13 of 15 bp of the topoisomerase II d box, GTN(A/T)A(T/C)ATTNATNN(A/G), occur in *BRCA1* intron 3 (Sander and Hsieh, 1985). Either the sequence tTAAgTATTTATGTA at position +948 or the sequence aTGTACATTTtTCTG at +1225 might serve as a recognition site for cleaving the double stranded DNA at +1008 of *BRCA1* intron 3.

Mutations in *BRCA1* and *BRCA2* account for at least 80% of hereditary breast and ovarian cancer (Narod et al., 1995; Rebbeck et al., 1997). Complex *BRCA1* and *BRCA2* mutations such as that in Family 5 account for a still unknown fraction of unexplained families with inherited breast and ovarian cancer (Narod et al., 1995; Rebbeck et al., 1997; Schubert et al., 1997). Additionally, there may be other as yet uncharacterized genes that predispose to hereditary breast cancer (Rebbeck et al., 1997; Schubert et al., 1997).

The mutation in Family 5 adds to a growing class of mutations not detectable by conventional screening methods. Amplification followed by genomic SSCP analysis would invariably amplify the wild-type allele in Family 5. Direct sequencing of splice junctions would not identify the genomic mutation because the primer routinely used to analyze the 5' splice site for intron 3 was deleted in the variant allele. Reports of genomic deletions of *BRCA1* that were undetected by conventional methods further illustrate the limitations in standard *BRCA1* and *BRCA2* screening procedures.

As more large genomic deletions and rearrangements are characterized in the *BRCA1* region, several questions remain to be addressed. Are Alu-Alu recombination events prominent in the *BRCA1* region, as seen for deletions in the β -globin gene cluster (Henthorn et al., 1990), or no more likely to be involved than random sequences in the region? Likewise, which is more representative of large deletions and rearrangements in the *BRCA1* region: the complex inverted duplication with deletion observed in Family 5 in this study or the simple deletions reported previously? Of primary importance is whether such mutations occur somatically as well as in the germline. If such mutations do exist somatically, it will be interesting to determine whether they are similar in Alu-dependence and complexity to those observed in the germline.

ACKNOWLEDGEMENTS

We wish to thank Jan Morrow for technical assistance, Piri Welch, Elizabeth Schubert, and Raymond Monnat for critical comments and discussion, and Roddy Grant for assistance with figures. This work was supported in part by NIH grant RO1-CA27632. Shannon Payne is supported in part by NIH training grant T32-GMO7735.

REFERENCES

- Antonarakis SE and the Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Human Mutation* 11:1-3.
- Bae YS, Kawasaki I, Ikeda H, Liu LF. 1988. Illegitimate recombination mediated by calf thymus DNA topoisomerase II *in vitro*. *Proc Natl Acad Sci USA* 85: 2076-2080.
- Berchuck A, Heron K-A, Carney ME, Lancaster JM, Fraser EG, Vinson VL, Deffenbaugh AM, Miron A, Marks JR, Futreal PA, Frank TS. 1998. Frequency of germline and somatic *BRCA1* mutations in ovarian cancer. *Clin Cancer Res* 4:2433-2437.
- Bieche I, Lidereau R. 1995. Genetic alterations in breast cancer. *Genes Chromosomes Cancer* 14: 227-251.
- Breast Cancer Information Core. 1999.
http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/BIC
- Devilee P, Cornelisse CJ. 1994. Somatic genetic changes in human breast cancer. *Biochem Biophys Acta* 1198: 113-130.
- Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE and the Breast Cancer Linkage Consortium. 1994. Risks of cancer in *BRCA1*-mutation carriers. *Lancet* 343:692-695.
- Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch E, Rowell SE, King, M-C. 1994. Confirmation of *BRCA1* by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nature Genet* 8:399-404.
- Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, Eddington K, McClure M, Frye C, Weaver-Feldhaus J, Ding W, Gholami Z, Söderkvist P, Terry L, Skolnick MH, Kamb A, Wiseman R. 1994. *BRCA1* mutations in primary breast and ovarian carcinomas. *Science* 266:120-122.
- Greenblatt MS, Grollman AP, Harris CC. 1996. Deletions and insertions in the p53 tumor suppressor gene in human cancers: Confirmation of the DNA polymerase slippage/misalignment model. *Cancer Res* 56: 2130-2136.
- Hall JM, Lee MK, Newman B, Morrow J, Anderson LA, Huey B, King, M-C. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250:1684-1689.
- Henthorn PS, Smithies O, Mager DL. 1990. Molecular analysis of deletions in the human β -globin gene cluster: Deletion junctions and locations of breakpoints. *Genomics* 6: 226-237.
- Hosking L, Trowsdale J, Nicolai H, Solomon E, Foulkes W, Stamp G, Signer E, Jeffreys A. 1995. A somatic *BRCA1* mutation in an ovarian tumor. *Nature Genet* 9: 343-344.
- Hu X, Ray PN, Worton RG. 1991. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J* 10:2471-2477.

- Ikeda H. 1986. Bacteriophage T4 DNA topoisomerase mediates illegitimate recombination *in vitro*. Proc Natl Acad Sci USA 83: 922-926.
- Ketterling RP, Vielhaber EL, Lind TJ, Thorland EC, Sommer SS. 1994. The rates and patterns of deletions in the human Factor IX gene. Am J Hum Genet 54: 201-213.
- Matsushima M, Kobayashi K, Emi M, Saito H, Saito J, Suimori K, Nakamura Y. 1995. Mutation analysis of the *BRCA1* gene in 76 Japanese ovarian cancer patients: four germline mutations, but no evidence of somatic mutation. Hum Mol Genet 4:1953-1956.
- Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C, Frank, T.S. 1995. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. Nature Genet 9: 439-443.
- Monnat RJ, Hackmann AFM, Chiaverotti TA. 1992. Nucleotide sequence analysis of human hypoxanthine phosphoribosyltransferase (HPRT) gene deletions. Genomics 13: 777-787.
- Narod SA, Ford D, Devilee P, Barkardottir RB, Lynch HT, Smith SA, Ponder BJ, Weber BL, Garber JE, Birch JM, Cornelis RS, Kelsell DP, Spurr NK, Smyth E, Haites N, Sobol H, Bignon Y-J, Chang-Claude J, Hamann U, Lindblom A, Borg A, Piver S, Gallion HH, Struewing JP, Whitternore A, Tonin P, Goldgar DE, Easton DF, the Breast Cancer Linkage Consortium. 1995. An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. Am J Hum Genet 56: 254-264.
- Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, Liu ET. 1995. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. Breast Cancer Res Treat 35:51-60.
- Newman B, Mu H, Butler L, Millikan RC, Moorman PG, King M-C. 1998. Frequency breast cancer attributable to *BRCA1* in a population-based series of American women. JAMA. 279:915-921.
- O'Connor MB, Malamy MH. 1985. Mapping of DNA gyrase cleavage sites *in vivo*: oxolinic acid induced cleavages in plasmid pBR322. J Mol Biol 181: 545-550.
- Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drüsedau M, Hogervorst FBL, Hageman S, Arts PJW, Ligtenberg MJL, Meijers-Heijboer H, Klijn, JGM, Vasen HFA, Cornelisse CJ, van't Veer LJ, Egbert-Bakker, van Ommen G-J.B, Devilee P. 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. Nature Genet 17: 341-345.
- Puget N, Torchard D, Serova-Sinilkova OM, Lynch HT, Feunteun J, Lenoir GM, Mazoyer, S. 1997. A 1-kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. Cancer Res 57: 828-831.
- Puget N, Stoppa-Lyonnet D, Sinilkova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S. 1999a. Screening for germ-line rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. Cancer Res 59: 455-461.
- Puget N, Sinilkova OM, Stoppa-Lyonnet D, Audouyraud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer, S. 1999b. An alu-mediated 6-kb duplication in the *BRCA1* gene: a new founder mutation? Am J Hum Genet 64: 300-302.
- Rainville IR, Albertini RJ, Nicklas JA. 1995. Breakpoints and junctional regions of intragenic deletions in the HPRT gene in human T-cells. Som Cell Mol Genet 21: 309-326.

- Rebbeck TR, Couch FJ, Kant J, DeShano M, Peng Y, Chen K, Garber JE, Weber BL. 1997. Genetic heterogeneity in hereditary breast cancer: Role of BRCA1 and BRCA2. *Am J Hu. Genet* 60:547-553.
- Sander M, Hsieh T-S. 1985. *Drosophila* topoisomerase II double-strand DNA cleavage: analysis of DNA sequence homology at the cleavage site. *Nuc Acids Res* 13: 1057-1072.
- Schubert EL, Lee MK, Mefford HC, Argonza RH, Morrow JE, Hull J, Dann JL, King, M-C. 1997. BRCA2 in American families with four or more cases of breast or ovarian cancer: Recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet* 60: 1031-1040.
- Sommer SS. 1995. Recent human germline mutation: Inferences from patients with hemophilia B. *Trends Genet* 11: 141-147.
- Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. 1997. Identification of a 14 kb deletion involving the promoter region of BRCA1 in a breast cancer family. *Hum Mol Genet* 6: 1513-1515.
- Takahashi H, Behbakht K, McGovern PE, Chiu H-C, Couch FJ, Weber B, Friedman LS, King M-C, Furusato M, LiVlosi VA, Menzin AW, Liu PC, Benjamin I, Morgan MA, King SA, Rebane BA, Cardonick A, Mikuta JJ, Rubin SC, Boyd J. 1995. Mutation analysis of the *BRCA1* gene in ovarian cancers. *Cancer Res* 55: 2998-3002.
- Wong C, Kazazian HH, Stetten G, Earnshaw WC, Van Keuren ML, Antonarakis SE. 1989. Molecular mechanism in the formation of a human ring chromosome 21. *Proc Natl Acad Sci USA* 86: 1914-1918.

Legends to Figures

Figure 1. Pedigree for Family 5. Three generation pedigree for Family 5 indicating markers used to establish linkage. Below the linkage markers are displayed the results of both SSCP analysis of cDNA and Long PCR analysis of genomic DNA using primer pair no. 3 designed to amplify a genomic region including *BRCA1* exon 3. Some unaffected family members are not included.

Figure 2. Variant allele reflecting *BRCA1* g.12977 ins10 del1039 in Family 5 detected by Long PCR of genomic DNA. Numbers above lanes are identification numbers from the pedigree. The wild-type genomic product extends from *BRCA1* intron 2 to intron 3 and is 5531 bp. Amplification from the mutant allele yields a 4502 bp product. The variant band is present in all individuals with the *BRCA1*-linked haplotype and who are heterozygous for deletion of exon 3 in *BRCA1* mRNA.

Figure 3. Exon skipping resulting in a truncated BRCA1 protein in Family 5. (A) Wild-type BRCA1 sequence and splicing of exon 3. nt 210-240 of HSU14680 shown. (B) Breakpoint sequence of the 1039 bp deletion and effects on splicing in Family 5. nt 210-221 of HSU14680 and 10 bp inverted duplication, followed by intron 3 sequence.

Table 1: Long PCR primers used for analysis of BRCA1 gene.

Pair No.	Primer Name	Temp.*	Size (bp)	GenBank L78833 ^b	Sequence
1	n2.1.br BA2	65	2364	2309-4672	5'-GCTA ATG CAA AGA CCG TCC GCT G-3' 5'-GTA CTT CTT CAA CCG GAA GAG CAG ATA AAT C-3'
2	1aF 3R	65	9936	3028-12963	5'-TTT GGA CAA TAG GTA GCG ATT CTG ACC TTC-3' 5'-AAC TCC AGA CTA GCA GGG TAG GGG GGG-3'
3	3F 4R	61	5531	12792-18322	5'-TCC TGA CAC AGC AGA CAT TTA-3' 5'-CCC GTC TCT ACA GAA AAC AC-3'
4	ivs3F 7R	65	10857	13715-24571	5'-TGT GAA GAC AGG AAA GGA CTT GAT ACC AGT TTC-3' 5'-CAC GGT TTC TGT AGC CTA TAC TTT GGA TGA TAG-3'
5	6I-2 ivs7R	65	3828	23839-27666	5'-GCT TTT CAG CTT GAC ACA GGT TTG G-3' 5'-CCY CAG CAC TCC TAA GAA CAT TTA GTA TAG G-3'
6	8F 11R	65	8361	28853-37213	5'-CAG GAA ACC AGT CTC AGT GTC CAA CTC TCT AAC CTT G-3' 5'-TGT CAC TCA GAC CAA CTT CTT GGC TTT CAG AC-3'
7	11F 13R	65	9518	36783-46300	5'-CCA TAC ACA TTT GGC TCA GGG TTA CCG AAG AGG G-3' 5'-TTC GCA GGT CTT CAA GGG CAG AAG AGT CAC-3'
8	12F ivs13R	65	8927	37715-46641	5'-GRT GTG AGA GTG AAA CAA GCG TCT CTG AAG ACT GCG-3' 5'-GCT TGT CAC CAA TTT CTC CCA TTC CAC TTA GCT TC-3'
9	13F 15R	65	8071	46235-54275	5'-GGA GCT AGR CTT CTA ACA GCT ACC CTT CCA TC-3' 5'-GAC TCC CAG AGC AAC TGT GCA TGT ACC ACC TAT C-3'
10	15F-2 18R-2	65	10613	54241-64852	5'-TGA TAG GTG GTA CAT GCA CAG TTG CTC TGG-3' 5'-GCT AAC TAC CCA TTT TCC TCC CCG AAT TCC-3'
11	17F 19R	61	4581	60922-65502	5'-GTG TAG AAC GTG CAG GAT TG-3' 5'-CAT TGT TAA GGA AAG TGG TGG-3'
12	18F 20R	65	6897	64777-71673	5'-TGC AGA TGC TGA GTT TGT GTG TGA ACG GAC-3' 5'-CCT GGG AAT CTC TTG CTC GCT TTG GAC C-3'
13	ivs19F 21R	65	7495	70154-77648	5'-TCC CAG TGA GGT GAA AAG CCG APT GTT AAG TTC-3' 5'-CCY ATA GCA ACA GAT TTC TAG CCG CTT GAG G-3'
14	20F 23R	61	9662	71518-81179	5'-ATA TGA CTT GTC TGC TCC AC-3' 5'-ACT GTG CTA CTC AAG CAC CA-3'
15	22F ivs24R	63	8536	79560-88095	5'-AGC TGT GTG GTG CTT CTG TGG TGA AGG-3' 5'-AGA GCT AGR AAG ATC AGA TGG TCT ACA GGA C-3'

a. Temp. = Annealing Temperature

b. Region of GenBank L78833 amplified by primer pair

Figure 1

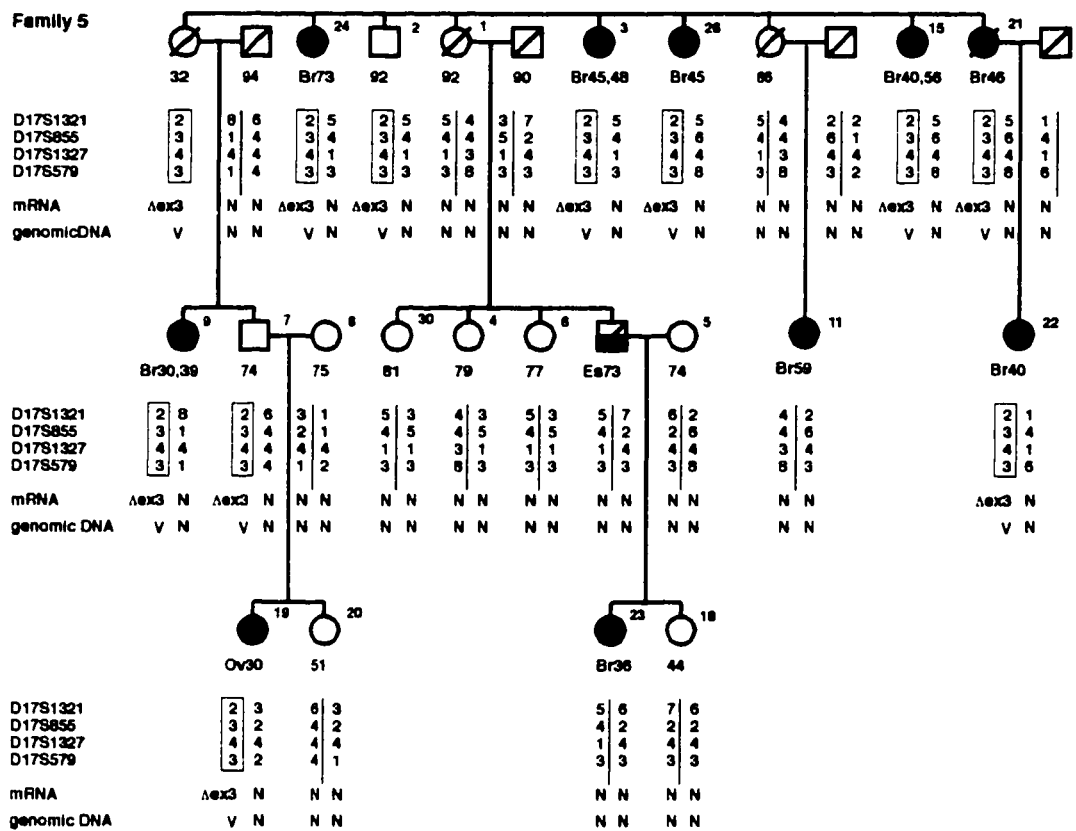
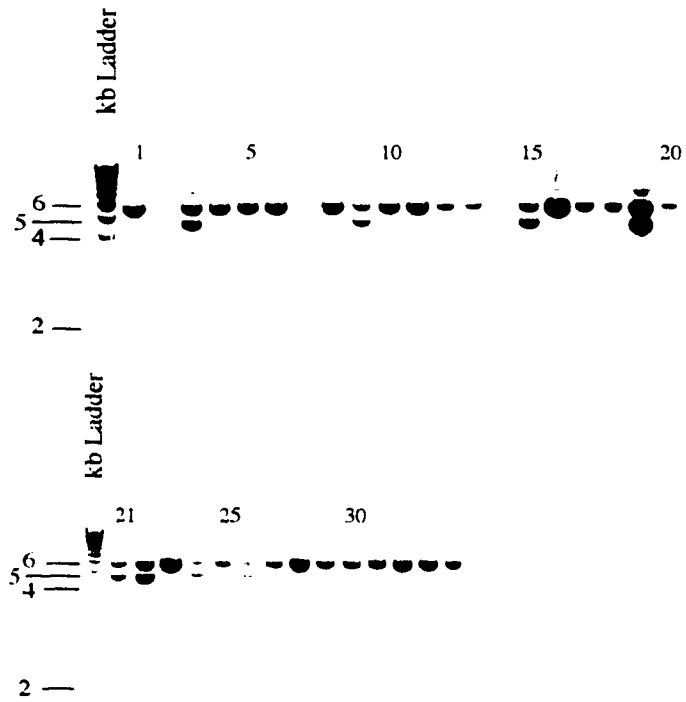


Figure 2



VITA

Shannon Renée Payne was the first child born to Ruth Ann Payne and Jerry Oliver Payne in Mt. Carmel, Illinois on September 13, 1970. To Shannon's great delight, her little sister joined her in this world on August 24, 1974. Shortly thereafter, Shannon's parents returned to Kentucky and settled in the southwest edge of the "Big City", Louisville, Kentucky. Shannon spent most of her childhood playing fantastic imagination games with anyone she could convince to join her. In 1979, she began classical study of the violin, a pursuit that was to give her both a leveling perspective on the world of academics and some of her closest lifelong friendships. At the age of 14 years, she and her sister became heads of the household in the care of their beloved Aunt Brenda and her three young children. It was a trying time, but gave her resources upon which she could fall back time and time again.

The absurdity of the high school education system led Shannon to choose an undergraduate education with an emphasis on learning and synthesis as opposed to grade point average. Her experiences at New College of the University of South Florida were among the most liberating to date. It was there that she met her husband, Roddy Grant, who shared her perspective on education and her love of new experiences. A required genetics course with Professor Sandra Gilchrist changed the course of Shannon's life. In genetics, she found a logic that was strong and pure. Although Shannon considered leaving the study of science for the Teacher's College of Columbia University, she could not abandon her love of genetics.

In the fall of 1993, Shannon entered a Ph.D. program in the Department of Botany in Seattle, Washington. Through a series of unexpected events, Shannon transferred to the Department of Genetics in the fall of 1995 where she found a sympathetic community of curious souls. Her positive experience with model systems via interactions with fellow genetics department researchers led her to give the mouse world "a go" after the completion of her Ph.D. in the Department of Genetics at the University of Washington. Stay tuned...