

Molecular Genetics in Breast Cancer — minisymposium

Transcript identification in the BRCA1 candidate region

Barbara L. Weber, MD^{1,*}, Kenneth J. Abel, PhD¹, Fergus J. Couch, PhD¹, Sofia Merajver, MD, PhD¹, Lucio Castilla, MS², Lawrence C. Brody, PhD³, Francis S. Collins, MD, PhD³
*Departments of Internal Medicine*¹, *and Biology*², *University of Michigan, Ann Arbor, MI; and National Center for Human Genome Research*³, *Bethesda, MD*

Key words: BRCA1, breast/ovarian cancer susceptibility gene, cDNA, exon amplification, molecular cloning, tumor suppressor gene

Summary

Chromosome 17q12-21 is known to contain a gene (or genes) which confers susceptibility to early-onset breast cancer and ovarian cancer (BRCA1). Identification and isolation of BRCA1 will likely provide the basis for increased understanding of the pathogenesis of breast and ovarian cancer, the development of targeted diagnostic and therapeutic approaches, and a means of screening women at risk of being BRCA1 mutation carriers. Genetic and physical maps of the BRCA1 candidate region have been largely completed and efforts are being directed at identification of candidate genes from within this region. We have begun the task of identifying transcripts from this region employing three complementary strategies. These include: 1) direct cDNA screening with cosmids derived from the BRCA1 region; 2) exon amplification; and 3) magnetic bead capture. Transcripts identified using these approaches are being characterized for: 1) tissue expression pattern; 2) the presence of genomic rearrangement in DNA derived from affected members of families believed to show linkage between breast cancer and genetic markers in the BRCA1 candidate interval; 3) altered size and/or expression pattern in RNA prepared from such individuals; and 4) homology to known genes or functional motifs. Germline mutations in affected individuals from these families will serve as presumptive evidence of BRCA1 identity.

Introduction

Family history has long been known to be a major contributor to breast cancer risk, but has frequently been assumed to be the result of complex interactions between multiple factors. This may in fact be true for a significant percentage of breast cancer cases. However, a landmark study,

published in 1990, demonstrated the existence of a breast cancer susceptibility gene on the long arm of human chromosome 17 [1]. Using genetic linkage analysis, Hall et al. convincingly demonstrated that alterations in this putative gene, now termed BRCA1, were inherited in an autosomal dominant manner and conferred a high risk of developing early onset breast cancer in a subset of

Presented by Barbara L. Weber at the 16th Annual San Antonio Breast Cancer Symposium, San Antonio TX, USA, November 4, 1993; Minisymposium on "Molecular Genetics in Breast Cancer".

* *Address for correspondence and offprints:* Barbara L. Weber, M.D., Department of Medicine, University of Pennsylvania, Room 1009 BRB I, 422 Curie Blvd., Philadelphia PA 19104, USA; *Tel:* 215-898-0247; *Fax:* 215-662-7617.

families. A subsequent report confirmed these findings and extended the observation to include susceptibility to ovarian cancer [2]. More recent studies indicate that about half of families with apparent autosomal dominant transmission of breast cancer susceptibility show linkage to genetic markers in this region. For families with both breast and ovarian cancer, the linked fraction may be greater than 90% [3]. Although estimates of the percentage of all breast cancer cases which occur as a direct result of germline mutations in BRCA1 vary from five to twenty percent [4-7], what is certain is that the estimates must be a function of age. Thus, although less than 5% of all breast cancers in elderly women are likely to result from BRCA1 germline mutations, this fraction may be much greater for women diagnosed with breast cancer under the age of 45 [3].

The findings of Hall et al. opened the door to the use of positional cloning aimed at isolation of BRCA1. The cloning of BRCA1 will have vast significance for several reasons. The most immediate effect will likely be the development of predictive testing for breast cancer. At present, genetic counseling based on linkage analysis is possible for only very few families. Once this gene has been isolated and common mutations characterized, identification of carriers will likely be able to be performed on individuals, without the requirement for study of entire large families. Perhaps of even greater significance, isolation and cloning of BRCA1 should ultimately lead to a more detailed understanding of the biology of breast cancer. It is hoped that the insights derived from the study of BRCA1 will give us new and more effective means to diagnose, treat, and ultimately prevent breast cancer.

BRCA1 is a putative tumor suppressor gene

Loss of heterozygosity (LOH) at a specific chromosomal location in tumors is widely interpreted as a marker for the absence of a functional tumor suppressor gene. Non-random deletions, as

detected by LOH at specific chromosomal regions, have been identified in a number of common human cancers, including carcinoma of the breast, supporting this hypothesis. In the case of inherited mutations in putative tumor suppressor genes, it is generally accepted that the germline mutation constitutes the "first hit" required for malignant transformation. In this setting, the subsequent somatic loss of the non-mutant allele represents the "second hit" [8]. Thus germline mutations in tumor suppressor genes create a highly susceptible group of individuals for whom only one somatic event is required to produce inactivation of tumor suppressor activity at a given locus. Recent evidence [9,10] demonstrating loss of non-mutant alleles flanking BRCA1 in breast tumors from affected members of linked families strongly supports the theory that BRCA1 is a tumor suppressor gene.

Identification of LOH at specific chromosomal locations has proven to be a useful tool not only for locating tumor suppressor genes, but for narrowing the candidate interval which must contain the sought-after gene. The location of the DCC gene, thought to play an important role in the development of colorectal cancer, was identified solely on the basis of LOH studies [11]. The candidate interval for the retinoblastoma gene was also significantly narrowed by defining small interstitial deletions in familial tumor specimens [12]. We have similarly used LOH analysis to decrease the size of the BRCA1 candidate region as defined by genetic linkage analysis.

Positional cloning

Functional cloning of a disease-related gene requires information on the gene product. Unfortunately, in the case of BRCA1, all that is known about the gene is the approximate chromosomal location and the phenotype of carrier females. As such, we have used a strategy initially referred to as reverse genetics, now termed positional cloning [13], in our attempts to

isolate BRCA1. Positional cloning has been successfully employed to clone a number of genes mediating important genetic diseases, including cystic fibrosis [14], neurofibromatosis [15], myotonic dystrophy, and Huntington disease [16]. It was also useful in identifying the familial adenomatous polyposis (APC) gene on chromosome 5q21 [17].

The positional cloning approach, as applied to the BRCA1 effort, is as outlined: 1) Families with apparent autosomal dominant transmission of breast cancer were studied with a large number of polymorphic genetic markers representing the entire human genome. A 20 cM region of chromosome 17q21 was identified as the original candidate interval [1]. 2) Additional families were studied to confirm the finding and demonstrate susceptibility to ovarian cancer in the same families [2]. 3) The candidate interval was saturated with polymorphic markers, and the order of these markers along the chromosome was determined by the complementary methods of radiation hybrid mapping, fluorescence in situ hybridization mapping, and genetic mapping [18-20]. 4) The candidate interval was narrowed from 20 cM to less than 4 cM by genotyping families believed to show linkage between breast/ovarian cancer and genetic markers on chromosome 17q with additional polymorphic markers in the region looking for recombination events [10,21,22]. 5) A physical map was constructed using yeast artificial chromosomes, pulsed-field gel electrophoresis (PFGE), and cosmid clones from a flow-sorted human chromosome 17 cosmid library provided by the Los Alamos National Laboratory. 6) Genes within the bounded region are now being identified by exon amplification [23], direct screening of cDNA libraries with cosmid clones from the region, and magnetic bead capture [24,25]. These transcripts are being examined for tissue distribution, mutations in affected individuals, and alterations between normal and malignant breast tissue. Demonstration of germline mutations in affected individuals from linked families will be accepted as presumptive evidence of gene identity.

Candidate genes

The ultimate goal of positional cloning is the identification of a disease-related gene. As noted above, several efficient methods are now available to isolate genes contained within genomic DNA fragments. However, a logical first step is to examine all known genes within the candidate region. The original BRCA1 candidate interval contained several previously-cloned genes which could be candidates for BRCA1. Any gene with biological properties suggesting that it could play a role in the development of breast and ovarian cancer needs to be excluded by recombination events or examined for the presence of germline mutations in affected members of linked families before it can be excluded as a potential candidate. Genes encoding nm-23 H1 and H2, (NME1 and NME2), retinoic acid receptor α -subunit (RARA), topoisomerase II α -subunit (TOP2A), gastrin (GAS), 17-hydroxysteroid dehydrogenase (EDH-17B), a small nuclear ribonucleoprotein (RNU2), and pancreatic polypeptide (PPY) were at one time all within the BRCA1 candidate region. All but GAS, TOP2A, EDH-17B, RNU2, and PPY have been excluded by genetic analysis of linked families. Of these, EDH-17B seems the most plausible candidate. This enzyme, which is expressed in many normal and neoplastic tissues, catalyzes the interconversion of estrone and estradiol. The latter is a hormone known to play a role in the growth regulation of breast cancer cell lines *in vitro*. However, to date, no germline mutations of the EDH-17B gene have been described in affected members of families in which the disease is thought to be linked to BRCA1 [22]. As no convincing candidate genes have been identified, we have turned to strategies which allow identification of novel genes within the candidate interval.

Transcript identification

The final step in positional cloning of any disease-related gene involves the isolation and

characterization of transcribed sequences from the genomic region that defines the candidate interval. Until recently, methods of finding transcripts tended to be laborious and cumbersome when the candidate interval spanned several hundred thousand to a few million base pairs. With the advent of several new transcript identification techniques, the effort required for isolation of transcripts from large genomic regions has been considerably lessened. In the past, isolation of coding regions from large genomic fragments were limited to strategies designed to identify sequences which are highly conserved among species (and therefore likely to represent coding regions) and the identification of sequence motifs present at the 5' end of many genes (CpG islands). Optimization of hybridization conditions and attempts to exploit splicing signals within transcripts have allowed us to apply several more efficient methods to the isolation of BRCA1.

Direct cDNA screening

Direct screening of cDNA libraries is conceptually the most straight forward of the three methods presented here. In brief, direct screening involves radioactive labeling of large regions of cloned genomic DNA (40-200 kb) and the use of this material to probe cDNA libraries. The result should be the isolation of transcripts from the region of interest. Although appealing in its simplicity, the method has enough drawbacks that it best serves to complement the other methods described here rather than supplant them.

It is estimated transcribed genes make up no more than 10% of the human genome. The remainder of the human genome consists of non-transcribed DNA between and within genes, and many families of repetitive elements. This noncoding DNA is not useful for screening cDNA libraries because it is either not transcribed, or transcribed but not unique to a given region of the genome (i.e. transcribed repetitive elements). Early positional cloning methods stripped the

genomic DNA of these "extraneous" sequences by first isolating non-repetitive DNA from genomic clones; however, this method is cumbersome and time-consuming for large chromosomal regions. If the candidate interval is defined by genetic recombination as is the case for BRCA1, it is likely to be a megabase or larger. For genomic regions this large, it is most efficient to directly label large clones such as cosmids (approximately 50 kb) or yeast artificial chromosomes (YACs, 100-1000 kb) and eliminate the time consuming step of subcloning single copy sequences. The tradeoff is that complex probes are more difficult to use for cDNA library screens. YACs have the additional complication that they are difficult to purify from the contaminating yeast genomic DNA. In the case of the BRCA1 region, we are using cosmids which have been individually labeled and pooled in groups of two to six cosmids each. To eliminate hybridization to transcribed repetitive sequences represented in a cDNA library, the probes are preannealed to a vast excess of unlabeled human genomic DNA. The preannealing conditions are chosen such that any highly repetitive probe sequences are bound to the cold competitor DNA and therefore are not available to hybridize to the filter-bound cDNAs.

On average, the genome contains an expressed sequence about every 50 kb. If human breast tissue expresses half of all genes, we would expect to find an expressed sequence with each cosmid probe used; in actuality we have been isolating approximately half that number. This may represent the actual distribution of genes in this region of chromosome 17 or a lack of sensitivity of the direct screening method. Genes may be missed for the following reasons: 1) Direct screening is biased towards the isolation of genes with large exons. In most cases the 3' terminal exon is the largest single exon in a gene. We have chosen to screen oligo dT-primed cDNA libraries to maximize the number of clones containing a complete 3' terminal exon. 2) The low "specific" activity of the probes with respect to coding sequences leads to a low signal to noise ratio when compared to traditional probes. In

many cases positive phage plaques are barely above the background; therefore, cDNAs hybridizing at or just below this threshold are also missed.

Falsely positive clones are primarily repetitive transcribed sequences. These sequences might not be sufficiently blocked by the preannealing procedure and will be present in many cDNA libraries. For this reason, every candidate clone is hybridized to a somatic cell hybrid panel to confirm that it maps to the BRCA1 region of chromosome 17 with a hybridization pattern suggestive of single copy DNA.

Exon amplification

Exon amplification, also called exon trapping, is a powerful technique for isolating gene sequences from genomic clones. This technique relies on the normal processing of most primary RNA transcripts, whereby noncontiguous sequences destined to make up the mature messenger RNA (exons) are accurately spliced together and the intervening sequences (introns) are excised. The exons are flanked by specific sequences which serve as signposts for the cellular splicing machinery, identifying the precise sites where cutting and joining of the exons occurs. Exon amplification permits the selective recovery of sequences present in cloned genomic DNA which are flanked by functional splice sites, yielding subclones from genomic DNA which are highly enriched for gene sequences.

We have successfully used exon amplification [23] to isolate coding sequences from cosmids derived from the BRCA1 candidate region. Briefly (see Figure 1), cosmid DNAs are digested and ligated into a plasmid splicing vector. The original vector, pSPL1, contained only a single cloning site; recent modifications to this vector include multiple cloning sites, permitting greater flexibility in the types of restriction fragments which can be used. The fragments are inserted into an intron of the HIV *tat* gene which is flanked by functional donor and acceptor splice

sites. The resulting constructs are transfected into COS-7 cells where *in vivo* processing pairs exons present in the genomic fragment with the *tat* donor and acceptor splice sites in the vector. The *tat* intron sequences are excised and the exon is now present in the mature mRNA. After a brief period in culture, cytoplasmic RNA is isolated from the cells and used as a template to generate cDNA using reverse transcriptase. The first strand cDNAs are made double stranded and amplified by PCR using nested, vector-specific primers, then cloned into plasmids and propagated in bacteria. Within one week, large numbers of candidate exon clones can be generated by this method. Starting with approximately 120 cosmid DNAs organized into pools of 6-10 each, we have generated more than 1300 exon clones derived from the BRCA1 region.

In order to eliminate redundancy, we have arrayed the exon clones on nylon filters, and probed the filters by hybridization with PCR-amplified inserts from individual clones. We have observed an average redundancy of 5 to 10-fold, although the range has varied considerably above and below this average. Non-redundant, representative clones which are shown to be single copy in hybridization tests with genomic DNA, and which map back to the correct chromosomal region, are then used as hybridization probes to screen cDNA libraries for larger corresponding clones. Approximately one-third of tested clones have yielded cognate cDNAs in our screening of libraries derived from normal breast. It is expected that a much higher fraction of clones would identify cDNAs in a comprehensive survey of a variety of tissues, as it has been reported that greater than 90% of clones generated by this method are likely to represent "true" exons [26].

The technique of exon amplification has a number of inherent advantages. First, large numbers of candidate gene sequences can be generated from regionally localized genomic DNA clones in a relatively short period of time. Second, coding sequences can be obtained irrespective of the sites or levels of expression of the

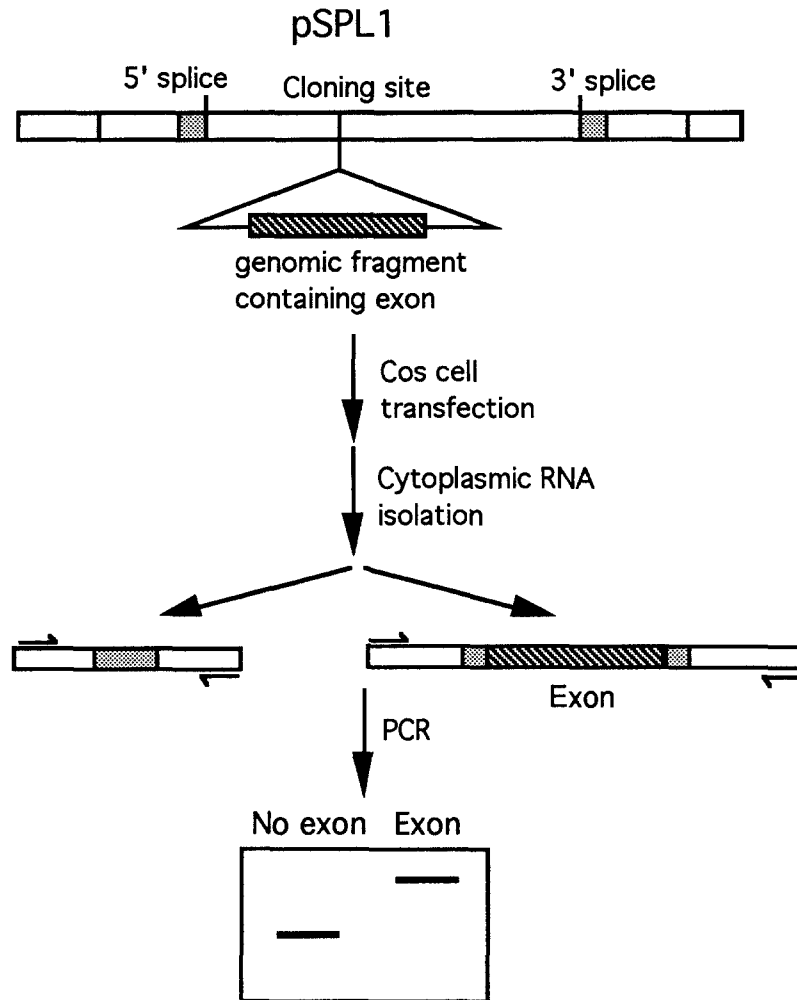


Figure 1. Outline of exon amplification method for isolating coding sequences.

corresponding gene, theoretically permitting isolation of genes expressed at very low levels within a highly restricted set of cell types. Third, more than 90% of clones have been free of repetitive sequences and are therefore useful as hybridization probes without additional competitor DNA. Fourth, the exon clones are typically short enough (average <200 base pairs) to permit rapid sequence analysis of the cloned inserts, enabling high volume sequencing of many clones simultaneously. In an effort to predict potential functions of the corresponding genes, the sequences determined for each representative clone can be used to screen databases of nucleic acid and protein sequences for homologies with known genes.

Finally, the ability to pool cosmids means that large genomic regions can rapidly be surveyed for genes by this technique.

There are also a number of limitations unique to this method. First, several types of artifactual products can arise from aberrant splicing events including pairing of only the vector splice sites, cloning of noncoding sequences which are flanked by "cryptic" splice sites within the insert and/or vector, or cloning of sequences which derive entirely from the HIV *tat* intron. In the original system using pSPL1, these cloning artifacts represented nearly half of the clones generated. However recent modifications to the vector [25] have greatly reduced the frequency of vector-only

and cryptic splicing artifacts, and efficient ligation of genomic fragments and vector can nearly eliminate the background of *tat* intronic clones. Second, while only a few percent of our clones were found to contain highly repetitive sequences, it appears that highly repetitive Alu sequences may continue to be cloned as a result of splicing events. We and others [27] have observed that specific sequences within Alu repeats can act as donor and acceptor splice junctions. Third, because of the short average length of the cloned exons, some of the clones may be difficult to use as hybridization probes. Finally, while the requirement for two flanking splice sites provides greater reliability for amplifying internal exons, this technique is not designed to isolate the 5'- or 3'-most exons of genes, and thus is not expected to efficiently amplify sequences from genes with less than three exons. However, a version of this technique has been reported recently which targets the 3' exons of genes [28].

Magnetic bead capture of cDNA

Magnetic bead capture is based on solution hybridization of cDNAs to immobilized region-specific genomic DNA fragments. Originally genomic DNA was immobilized on nylon filters [24]; this technique has subsequently been modified to take advantage of magnetic beads as the genomic fragment support system [25]. Briefly (see Figure 2), cosmid DNA from the BRCA1 candidate interval is digested with a frequent-cutter restriction enzyme and ligated to linkers which contain the appropriate restriction site and unique sequence complementary to the biotinylated PCR primers being used for subsequent amplification. cDNAs from appropriate libraries (breast and ovary in this case) are also PCR-amplified using vector primers and hybridized at high stringency to the biotinylated genomic fragments. The biotinylated genomic-cDNA complexes are captured with streptavidin-coated paramagnetic beads (Dyna), and washed to remove

unbound and non-specific cDNAs. The captured cDNAs, enriched for sequences from the chromosomal region the cosmids are derived from, are eluted by boiling and PCR-amplified with vector primers. Following purification and size selection of fragments greater than 200 bp in length, cDNAs are subcloned as a region-specific sub-library. Larger cDNAs are analyzed to determine if they map back to the original cloned genomic DNA or contain repetitive and/or vector sequences. Each clone which maps back to the region can then be used to: 1) determine the size of the full length cDNA by Northern blot analysis; 2) isolate the full length transcript from cDNA libraries; and 3) generate cDNA sequence from the gene for homology searches in genome databases.

We have successfully utilized this technique to isolate several hundred cDNA fragments from the BRCA1 candidate region. In this case, small pools of 7-10 overlapping cosmids served as the genomic DNA template. Of those tested, 72% of cDNAs have been shown to map back to the original genomic clones, 12% contained only repetitive sequences, 6% were cosmid vector fragments, and 10% did not map back, suggesting that these clones originated in other genomic regions and were captured non-specifically. As both genomic and cDNA clones are PCR-amplified, a singular advantage of this technique is that only small quantities of starting materials are required. An additional advantage of this approach is the speed with which large DNA segments can be screened for the presence of transcribed genes. In general, the technique is highly efficient, and may be applied to several sources of cDNA simultaneously or in parallel reactions. When random-primed cDNA libraries are used as starting material, cDNA fragments obtained are randomly distributed over each mRNA, aiding in rapid generation of full length cDNA clones. Furthermore, the method may normalize transcript levels, with enrichment of rare messages.

Several limitations are associated with this method. First, the small size of the captured subclones necessitates re-screening of cDNA

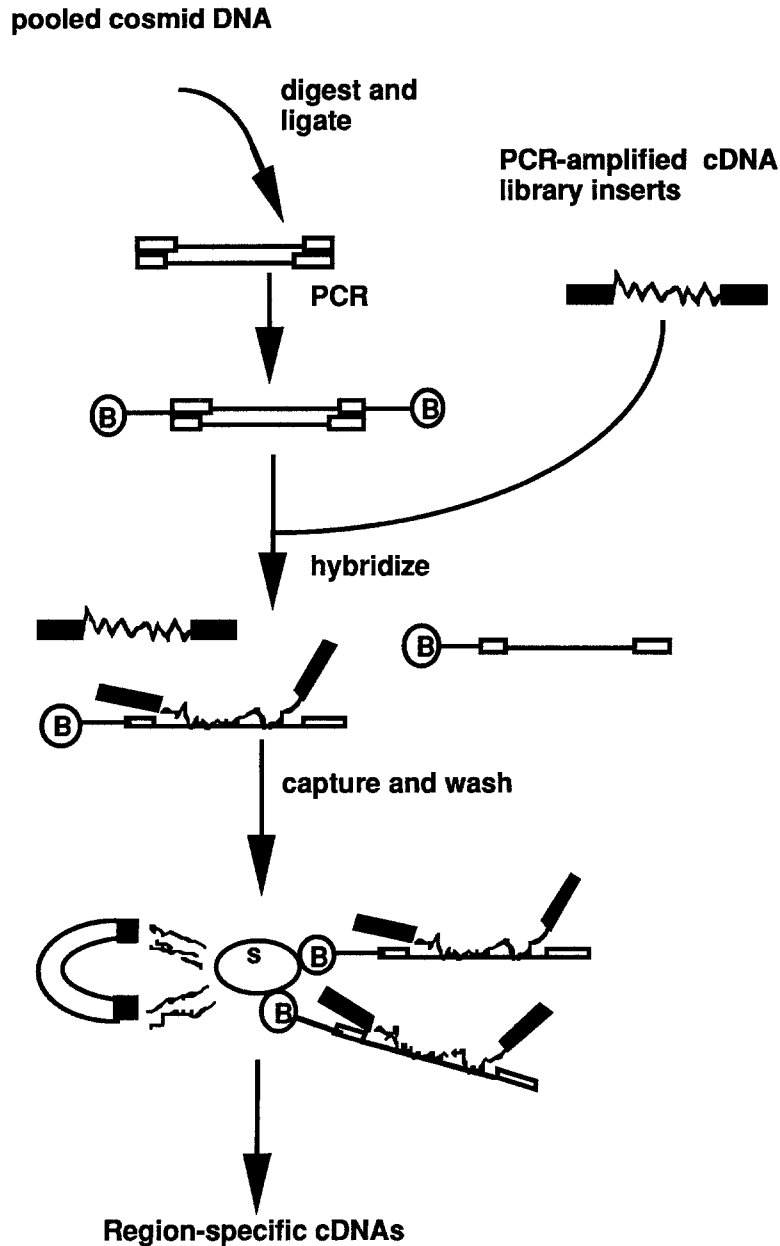


Figure 2. Outline of magnetic bead capture of specific cDNA fragments.

libraries to obtain full length cDNA. Secondly, genomic clones which contain expressed repeat sequences can result in capture of a large number of cDNAs containing these repeats. Once identified, these repetitive sequences can be used to pre-screen the sublibrary of captured clones or can be added as a blocking agent to any future

capture experiments in the region. Thirdly, because the technique is based on hybridization, pseudogenes and multigene family members may be enriched. These clones can be excluded at an early stage by the more stringent hybridization conditions associated with mapping the cDNA back to the original parent genomic clones.

Transcript characterization

BRCA1 candidate genes identified by any of these methods are being used to screen RNA blots from a variety of normal tissues, including breast and ovary. The results of these experiments provide information on the expected size of the full-length transcript. Primary tumors and tumor cell lines derived from these tissues are also being screened as BRCA1 may be differentially expressed in normal and malignant tissue. It is also possible that transcripts from breast tumors and/or obligate mutation carriers may be altered with respect to size. The proof of correctness of a candidate gene will be the identification of specific DNA abnormalities in germline DNA of affected individuals from linked families. Accordingly, genomic DNA from families in which breast cancer is linked to genetic markers on chromosome 17q is being surveyed for mutations in candidate genes using Southern blotting and more sensitive methods such as chemical mismatch cleavage [29], single-strand conformation polymorphism screening (SSCP) [30] or PCR sequencing. Somatic mutations in tumors will provide supporting evidence, but as has been recently demonstrated for the MCC gene [31] in familial polyposis and the prohibitin gene in breast cancer [32], these may not always indicate the gene responsible for inherited disease.

Conclusion

BRCA1 is a highly penetrant breast/ovarian cancer susceptibility gene on the long arm of human chromosome 17. We and others are attempting to isolate BRCA1 using a positional cloning approach. Following publication of the location of this gene as defined by genetic linkage analysis [1], we have narrowed the candidate interval by further genetic analysis of affected families and studies of allelic loss in tumors from these families, generated and ordered new polymorphic markers within and flanking the candidate interval and developed a physical map of the

region. We now have turned our efforts to transcript identification within the candidate interval. We are currently using three complementary approaches dependent on region-specific cosmids, including direct screening of cDNA libraries, exon amplification, and magnetic bead capture. It is hoped that this intensive approach will lead to the isolation of BRCA1 in the near future, the identification of which is likely to provide important insights into the pathophysiology of breast cancer.

References

1. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King M-C: Linkage of early onset breast cancer to chromosome 17q21. *Science* 250:1684-1689, 1990.
2. Narod SA, Feuteun J, Lynch HT, Watson P, Conway T, Lynch J, Lenoir GM: Familial breast-ovarian cancer locus on chromosome 17q12-23. *Lancet* 338: 82-83, 1991.
3. Easton DF, Bishop DT, Ford D, Crockford GP, and the Breast Cancer Linkage Consortium: Genetic linkage analysis in familial breast and ovarian cancer — results from 214 families. *Am J Hum Genet* 52:678-701, 1993.
4. Claus EB, Risch N, Thompson WD: Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 48:232-242, 1991.
5. Colditz GA, Willett WC, Hunter DJ, Stampfer MJ, Manson JE, Hennekens CH, Rosner BA: Family history, age and risk of breast cancer. *JAMA* 270: 338-343, 1993.
6. Slattery ML, Kerber RA: A comprehensive evaluation of family history and breast cancer risk: the Utah Population Database. *JAMA* 270:1563-1568, 1993.
7. Weber BL, Garber JE: Family history and breast cancer: probabilities and possibilities. *JAMA* 270: 1602-1603, 1993.
8. Knudson AG: Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA* 68:820-823, 1971.
9. Smith SA, Easton DF, Evans DGR, Ponder BAJ: Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nature Genetics* 2:128-131, 1991.
10. Chamberlain JS, Boehnke M, Frank TS, Kioussis S, Xu J, Guo S-W, Hauser ER, Norum RA, Helmbold EA,

- Markel DS, Keshavarzi SM, Jackson CE, Calzone K, Garber J, Collins FS, Weber BL: BRCA1 maps proximal to D17S579 on chromosome 17q21 by genetic analysis. *Am J Hum Genet* 52:792-798, 1993.
11. Fearon ER, Cho KR, Nigro JM, Kern SE, Simons JW, Ruppert JM, Hamilton SR, Preisinger AC, Thomas G, Kinzler KW: Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science* 247:49-56, 1990.
 12. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EY: Human retinoblastoma susceptibility gene: cloning, identification and sequence. *Science* 235:1394-1399, 1987.
 13. Collins FS: Positional cloning: Let's not call it reverse any more. *Nature Genetics* 1:3-6, 1992.
 14. Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N: Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* 245:1059-1065, 1989.
 15. Wallace MR, Marchuk DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, Fountain JW, Brereton A, Nicholson J, Mitchell AL: Type I neurofibromatosis gene: Identification of a large transcript disrupted in three NF1 patients. *Science* 249:181-186, 1990.
 16. The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's Disease chromosome. *Cell* 72:971-983, 1993.
 17. Kinzler KW, Nilbert MC, Su L-K, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hedge P, McKechnie D: Identification of FAP locus genes from chromosome 5q21. *Science* 253:661-665, 1991.
 18. Abel KJ, Boehnke M, Prahallad M, Ho P, Flejter WL, Watkins M, VanderStoep J, Chandrasekharappa SC, Collins FS, Glover TW, Weber BL: A radiation hybrid map of the BRCA1 region of chromosome 17q12-21. *Genomics* 17:632-641, 1993.
 19. Flejter WL, Barcroft CL, Guo S-W, Lynch ED, Boehnke M, Chandrasekharappa SC, Hayes S, Collins FS, Weber BL, Glover TW: Multi-color FISH mapping with Alu-PCR amplified YAC clone DNA determines the order of markers in the BRCA1 region on chromosome 17q12-12. *Genomics* 17:624-631, 1993.
 20. Anderson LA, Friedman L, Osborne-Lawrence S, Lynch E, Weissenbach J, Bowcock A, King M-C: High-density genetic map of the BRCA1 region of chromosome 17q12-21. *Genomics* 17:618-623, 1993.
 21. Bowcock AM, Anderson LA, Black DM, Lawrence SO, Hall JM, King M-C: Polymorphisms in THRA1 and D17S183 flank less than 4 cM interval for the breast-ovarian cancer gene on chromosome 17q. *Am J Hum Genet* (in Press).
 22. Simard J, Feuteun J, Lenoir G, Tonin P, Normand T, Luu The V, Vivier A, Lasko D, Morgan K, Rouleau GA, Lynch H, Labrie F, Narod SA: Genetic mapping of the breast-ovarian cancer syndrome to a small interval on chromosome 17q12-21: exclusion of candidate genes EDH17B2 and RARA. *Human Molecular Genetics* 2:1193-1199, 1993.
 23. Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA, Housman DE: Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci USA* 88:4005-4009, 1991.
 24. Lovett M, Kere J, Hinton LM: Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proc Natl Acad Sci USA* 88:9628-9632, 1991.
 25. Tagle DA, Swaroop M, Lovett M, Collins FS: Magnetic bead capture of expressed sequences within large genomic segments. *Nature* 161:751-756, 1993.
 26. Buckler AJ, Statler CJ, Rutter JL, Murrell J, Trofatter JA, McCormick MK, Deaven L, Mayzis RK, Gusella JF: Chromosome specific libraries of exons and high resolution transcription maps [abstract]. *Am J Hum Genet* 53:A192.
 27. North MA, Sanseau P, Bucker AJ, Church D, Jackson A, Patel K, Trowsdale J, Lehrach H: Efficiency and specificity of gene isolation by exon amplification. *Mammalian Genome* 4:466-474, 1993.
 28. Krizman DB, Berget SM: 3'-terminal exon trapping: Identification of genes from vertebrate DNA. *Focus* 15:106-108, 1993.
 29. Cotton RG, Rodrigues NR, Campbell RD: Reactivity of cytosine and thymine in single base pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc Natl Acad Sci USA* 85:4397-4401, 1988.
 30. Orita M, Suzuki Y, Sekiya T, Hayashi K: Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5:874-879, 1989.
 31. Kinzler KW, Nilbert MC, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hamilton SR, Hedge P, Markham A, Carlson M, Joslyn G, Groden J, White R, Miki Y, Miyoshi Y, Nishishio I, Nakamura Y: Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. *Science* 251:1366-1370, 1991.
 32. Sato T, Saito H, Swensen J, Olifant A, Wood C, Danner D, Sakamoto T, Takita K, Kasumi F, Miki Y, Skolnik M, Nakamura Y: The human prohibitin gene located on chromosome 17q21 is mutated in sporadic breast cancer. *Cancer Res* 52:1643-1646, 1992.