AD NUMBER

ADB251634

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info; Jul 99 Other requests shall be referred to USAMRMC, Ft Detrick, MD 21702-5012

AUTHORITY

USAMRMC ltr, 1 Jun 2001.

AD_____

GRANT NUMBER DAMD17-97-1-7140

TITLE: Prognostic Significance of Loss-of-Heterozygosity of the CUTL1 Putative Tumor Suppressor Gene in Breast Cancers

PRINCIPAL INVESTIGATOR: Alain Nepveu, Ph.D.

CONTRACTING ORGANIZATION: McGill University
Montreal, Quebec H3A 2T2, Canada

REPORT DATE: July 1999

TYPE OF REPORT: Annual

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

NOTICE

USING GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA INCLUDED IN THIS DOCUMENT FOR ANY PURPOSE OTHER THAN GOVERNMENT PROCUREMENT DOES NOT IN ANY WAY OBLIGATE THE U.S. GOVERNMENT. THE FACT THAT THE GOVERNMENT FORMULATED OR SUPPLIED THE DRAWINGS, SPECIFICATIONS, OR OTHER DATA DOES NOT LICENSE THE HOLDER OR ANY OTHER PERSON OR CORPORATION; OR CONVEY ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY RELATE TO THEM.

LIMITED RIGHTS LEGEND

Award Number: DAMD17-97-1-7140
Organization: McGill University

Those portions of the technical data contained in this report marked as limited rights data shall not, without the written permission of the above contractor, be (a) released or disclosed outside the government, (b) used by the Government for manufacture or, in the case of computer software documentation, for preparing the same or similar computer software, or (c) used by a party other than the Government, except that the Government may release or disclose technical data to persons outside the Government, or permit the use of technical data by such persons, if (i) such release, disclosure, or use is necessary for emergency repair or overhaul or (ii) is a release or disclosure of technical data (other than detailed manufacturing or process data) to, or use of such data by, a foreign government that is in the interest of the Government and is required for evaluational or informational purposes, provided in either case that such release, disclosure or use is made subject to a prohibition that the person to whom the data is released or disclosed may not further use, release or disclose such data, and the contractor or subcontractor or subcontractor asserting the restriction is notified of such release, disclosure or use. This legend, together with the indications of the portions of this data which are subject to such limitations, shall be included on any reproduction hereof which includes any part of the portions subject to such limitations.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

Patricia C. Modrow _____ 2/11/02 _____

_____ _____

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1999	Annual (1 Jul 98 - 30 Jun 99)

4. TITLE AND SUBTITLE
Prognostic Significance of Loss-of-Heterozygosity of the CUTL1 Putative Tumor Suppressor Gene in Breast Cancer

5. FUNDING NUMBERS
DAMD17-97-1-7140

6. AUTHOR(S)
Nepveu, Alain, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
McGill University
Montreal, Quebec H3A 2T2, Canada

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution authorized to U.S. Government agencies only (proprietary information, Jul 99). Other requests for this document shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, Maryland 21702-5012.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A fraction (~18%) of spontaneous breast cancers exhibited LOH of 7q22, and the smallest deleted region included the CUTL1 gene. We found no correlation between LOH at 7q22 and tumor type or grade, ER or PR expression, nodal status, or age of onset. However, the average sizes of breast tumors with or without LOH at 7q22 were 4.99 and 2.93 cm respectively. We conclude that LOH of 7q22 will not have much prognostic value. Other results forced us to reconsider the hypothesis that CUTL1 is a candidate tumor suppressor gene. Although overexpression of the full length CDP/Cut protein is cytotoxic, elevated CDP/Cut DNA binding activity in mammary epithelial cell lines correlated with higher saturation density. Moreover, proteins shorter than the full length CDP/Cut protein were found to be responsible for CDP/Cut DNA binding activity. Shorter CDP/Cut protein can be generated by processing of the full length protein. In addition, our studies uncovered a novel, shorter, CDP/Cut mRNA that was expressed at higher levels in two breast carcinoma with lobular features. In tissue culture, overexpression of this mRNA in the MCF-12A mammary epithelial cells caused some morphological changes reminiscent of epithelial-mesenchymal transition.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer tumor suppressor gene; loss-of-heterozygosity; transcription; repression; prognostic markers	50 16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

A.N. For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

A.N. In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

A.N. In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

A.N. In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Alain Nguyen 7.29.99

PI - Signature Date

TABLE OF CONTENT

TABLE OF CONTENT (page 2)

INTRODUCTION

Background

Several years ago, we isolated the cDNA for a protein that binds to the promoter of the c-Myc proto-oncogene and represses its expression (1). This cDNA was found to encode the human homologue of *Drosophila* Cut, a homeodomain protein believed to determine cell-type specificity in several tissues. The human protein has been called the CCAAT-displacement protein (CDP) or CDP/Cut, and the gene encoding it, CUTL-1 (Cut-like 1).The CUTL1 gene was mapped to chromosome 7, band 7q22, a chromosomal region that is frequently rearranged in some human cancers (2, 3). Two sets of data suggested that CUTL1 may be altered in some breast cancers. First, female transgenic mice expressing the Polyomavirus (PyV) Large T (LT) antigen under the control of the mouse mammary tumor virus long terminal repeat (MMTV-LTR) frequently develop mammary tumors and uterine leiomyomas (4). The results of coimmunoprecipitation analyses revealed that specific complexes of CDP/Cut and PyV LT antigen could be detected in both leiomyomas and mammary tumors (4). Although the functional significance of this interaction remains to be determined, the existence of complexes between CDP/Cut and a viral oncoprotein suggested that alterations in the function of CDP/Cut may be an important event in the etiology of breast cancer. In agreement with this, in a pilot study, we identified LOH at CUTL1 in 6 of 63 informative patients (9.5%). Using polymorphic markers covering the entirety of 7q, we found that the region of LOH was very large and included 7q31 as well as 7q22 in 4 of these 6 cases, while in the 2 remaining cases only 7q22 was deleted. These results suggested that two regions of the long arm of chromosome are deleted in a fraction of breast cancers: 7q22 and 7q31.

Hypothesis

In our proposal of 1996, we hypothesized that the human Cut gene, CUTL1, may function as a tumor suppressor gene whose deletion and/or mutation represents an important event in the etiology of some breast tumors.

Specific Aims

Our **specific aims** were to use polymorphic markers within and around CUTL1 to confirm LOH of 7q22 in breast cancers, to determine whether CUTL1 is situated in the minimal region of 7q22 that is deleted in some breast tumors, to verify whether CUTL1 is mutated in these tumors and to determine whether LOH of CUTL1 can be used as a prognostic marker for breast tumors.

BODY

Task 1: Months 1-12 Determine the frequency of LOH of CUTL1 in patients with sporadic breast cancer using tumor samples from the Royal Victoria Hospital, the Manitoba Breast Tumor Bank and Dr. Rosette Lidereau.

This task was accomplished mostly during year 1, completed during the first months of year 2 and the results of this study were published in April 1999 in the journal Oncogene (5) (see Appendix 1). Briefly, we found LOH of CUTL1 in 12 out of 66 (18%) sporadic breast tumors (Appendix 1, Fig. 1 and 2, and Table 2). Among these 12 tumors, the LOH region encompassed the 7q31 region in 5 cases only. Thus, our results defined a second region of LOH on the long arm of chromosome 7. The smallest deleted region in these 12 tumors included CUTL1. The boundary of the smallest deleted region on the telomeric side was located between markers D7S578 and D7S515, and on the centromeric side between markers D7S515 and D7S666. It should be noted that the last two markers are positioned within introns 3 and 6 of CUTL1, respectively. These results were consistent with the notion that CUTL1 is a tumor suppressor gene. If CUTL1 is not the tumor suppressor gene, at the minimum our results provide mapping information that will be useful for an eventual positional cloning strategy.

Task 2 Months 12-24 Establish correlation between LOH of CUTL1 and other clinical or pathological parameters. Initiate retrospective studies to confirm any correlation.

To determine whether loss of genetic material at 7q22 may be associated with pathological features of breast tumors, we analyzed available clinical data including the tumor type, grade and size, estrogen and progesterone receptor (ER and PR) status, occurrence of lymph node metastasis and age of onset (Appendix 1, Table 3). Statistical analysis was performed using the Mann-Whitney test (6). We found no correlation between LOH at 7q22 and tumor type or grade, ER or PR expression, nodal status, or age of onset. However, there was a significant association (p value = 0.0082) between LOH at 7q22 and tumor size (Appendix 1, Table 3). The average sizes of breast tumors with or without LOH at 7q22 were 4.99 and 2.93 cm respectively. The difference in size was further accentuated in the group of tumors with LOH at 7q22 but retention of 7q31 (6.5 Vs 2.99 cm; p value = 0.0074). In contrast, no significant correlation was found between LOH at 7q31 and any clinical feature.

Task 3 Months 1-12 Set up a PCR-based approach to measure c-Myc gene copy number in breast tumors. Analyze c-Myc gene status in parallel to CUTL1 in all samples.

Two sets of primers derived from the c-Myc gene intron 1 were used to measure c-Myc gene copy number in breast tumors. As a control, a primer set from the glyceraldehyde-phosphate dehydrogenase was included in the PCR-reactions. In addition, another strategy was used to verify whether one allele of the c-Myc gene was amplified. The primers were chosen so as to flank a CA repeat that is found at the 3' end of c-Myc intron 1. Since CA repeats are highly polymorphic, in many samples these primers allow the amplification of two products that can be distinguished on the basis of their difference in size. If one allele is present in several copies in a tumor, PCR-amplification would produce a band of greater intensity as compared to the other allele.

As hypothesized in our proposal of 1996, c-Myc gene amplification was not found among the 12 tumors with CUTL1 LOH. However, in breast tumor cell lines (see Tasks 5 and 6), we did not find any correlation between c-Myc expression and CDP/Cut DNA binding activity. For this reason, we decided not to analyze more breast tumors for c-Myc amplification.

We did not find any tumor with c-Myc amplification in our panel of breast tumors. Moreover, in breast tumor cell lines (see Tasks 5 and 6), we did not find any correlation between c-Myc expression and CDP/Cut DNA binding activity. For this reason, we decided not to analyze more breast tumors for c-Myc amplification.

Task 4 Months 12-24 Compare the prognostic values of CUTL1 LOH and c-Myc amplification. Changes in either c-Myc or CUTL1 will be related with various clinical and biological parameters including the histologic types (alveolar or ductal), tumor grades, lymph node infiltration, labeling index (or tumor cell proliferation), erbB2, estrogen and progesterone receptor status, disease-free survival and overall survival rates.

As stated above (Task 2), we found no correlation between LOH of CUTL1 and tumor type or grade, ER or PR expression, nodal status, or age of onset. Since LOH of CUTL1 was found in breast tumors of all grades, we concluded that this genetic change occurs early in the development of breast tumors. As a consequence, CUTL1 LOH will not have useful prognostic value. However, in the course of our study, we identified a novel CUTL1 mRNA that was found to be expressed at higher level in two breast tumors with an invasive phenotype (discussed in Task 7, see Fig. 3). In a separate document, I will submit a Revised Statement of Work in which I will propose to test whether expression of this novel mRNA may have useful prognostic value.

Task 5 Months 1-24 Investigate Cut expression as well as Cut DNA binding and repression activity in a panel of cell lines derived from breast tumors. Steady-state levels of Cut and c-Myc mRNAs and proteins will be measured by Northern and Western blot analyses. Evaluate Cut DNA binding activity by

electrophoretic mobility shift assay (EMSA). Assess Cut repression activity by cotransfection studies following RT-PCR cloning of Cut cDNAs from these cell lines.

CDP/Cut expression and DNA binding activity have been analyzed in a panel of breast tumor cell lines, nontumorigenic immortalized mammary epithelial cell lines and non immortalized human mammary epithelial cells (HMEC). In parallel, c-Myc mRNAs were measured to verify whether elevated CDP/Cut DNA binding activity correlates with lower c-Myc mRNA levels. Several findings were made.

CDP/Cut DNA binding activity varied greatly among cell lines (Fig. 1A, EMSA of CDP/Cut in human mammary cells; Fig. 2A, EMSA of CDP/Cut). Unexpectedly, there was no correlation between CDP/Cut activity and c-Myc mRNA levels.

CDP/Cut protein expression was analyzed by Western blot analysis (Fig. 1B and 2B). CDP/Cut has an expected M.W. of 160 kDa and migrates in SDS-PAGE with an apparent M.W. of 180-200 kDa. In breast tumor cell lines, a large number of bands were observed with anti-CDP/Cut antibodies raised against various regions of CDP/Cut. Whether all these bands represent true CDP/Cut proteins or are unrelated cross-reacting proteins is not entirely clear at this point. In light of these results, we have decided to develop novel antibodies raised against short regions of CDP/Cut (see Task 11 in the revised Statement of Work). Nevertheless, one important findings was made: there was no correlation between DNA binding and the abundance of the full length CDP/Cut protein, even when cell extracts were treated with phosphatases (Fig. 2A and B). Additional results in fact suggested that a protein shorter than full length CDP/Cut is responsible for CDP/Cut DNA binding activity. This issue will be discussed further in the conclusion section, under the title "A Shorter CDP/Cut Protein Is Responsible for CDP/Cut DNA Binding Activity".

Since the Cdc25A phosphatase had been found to stimulate CDP/Cut DNA binding, we measured Cdc25A expression by Western blot analysis. Although, Cdc25A was expressed at high levels in several breast tumor cell lines, we did not find a correlation between Cdc25A expression and CDP/Cut DNA binding activity (Fig. 1C). Since overexpression of CDC25B was detected in 32 percent of human primary breast cancers tested, it is possible that CDP/Cut activity in breast tumor cell lines correlates with the latter (14). This remains to be verified.

CDP/Cut contains four DNA binding domains: the Cut homeodomain and three regions called Cut repeats (7-10). In other studies performed in parallel to this project, we demonstrated that CDP/Cut DNA binding activity can be inhibited following phosphorylation of Cut repeats by casein kinase II (CKII) or protein kinase C (PKC), or of the Cut homeodomain most likely by a cyclin-dependent kinase (11-13). Moreover, CDP/Cut DNA binding activity can be stimulated following dephosphorylation by the Cdc25A phosphatase (11). In practice, inhibition of CDP/Cut DNA binding activity can be monitored by treating cellular extracts with a non-specific phosphatase prior to EMSA .(11-13) . As expected, in most cell lines, CDP/Cut DNA binding activity was down-modulated by phosphorylation. An example is shown in Fig. 2A.

Task 6 Months 1-24 Correlate Cut activity with c-Myc expression and gene copy number in cell lines derived from breast tumors.

c-Myc mRNA levels were investigated in a panel of breast tumor cell lines and compared with CDP/Cut DNA binding activity. We found no correlation between c-Myc expression and CDP/Cut DNA binding activity.

We originally cloned the human CDP/Cut cDNA using a protein binding site from the promoter of the c-Myc proto-oncogene and we accumulated biochemical evidence to show that CDP/Cut can repress c-Myc (1). Our recent studies, however, indicated that there is no physiological correlation between CDP/Cut DNA binding activity and repression of c-Myc (11). In G0, when c-Myc transcription was off, CDP/Cut DNA binding activity was also very weak. No change in CDP/Cut was observed when c-Myc transcription was upregulated as cells reentered the cell cycle in early G1. CDP/Cut DNA binding increased later at the G1/S transition, a time at which c-Myc transcription was stable. Thus, we are compelled to conclude that in fibroblastic cells and breast tumor cell lines, CDP/Cut does not contribute to the repression of c-Myc. We cannot exclude the possibility that in certain differentiated cells CDP/Cut may contribute to the down-regulation of c-Myc, but the evidence for this so far is lacking.

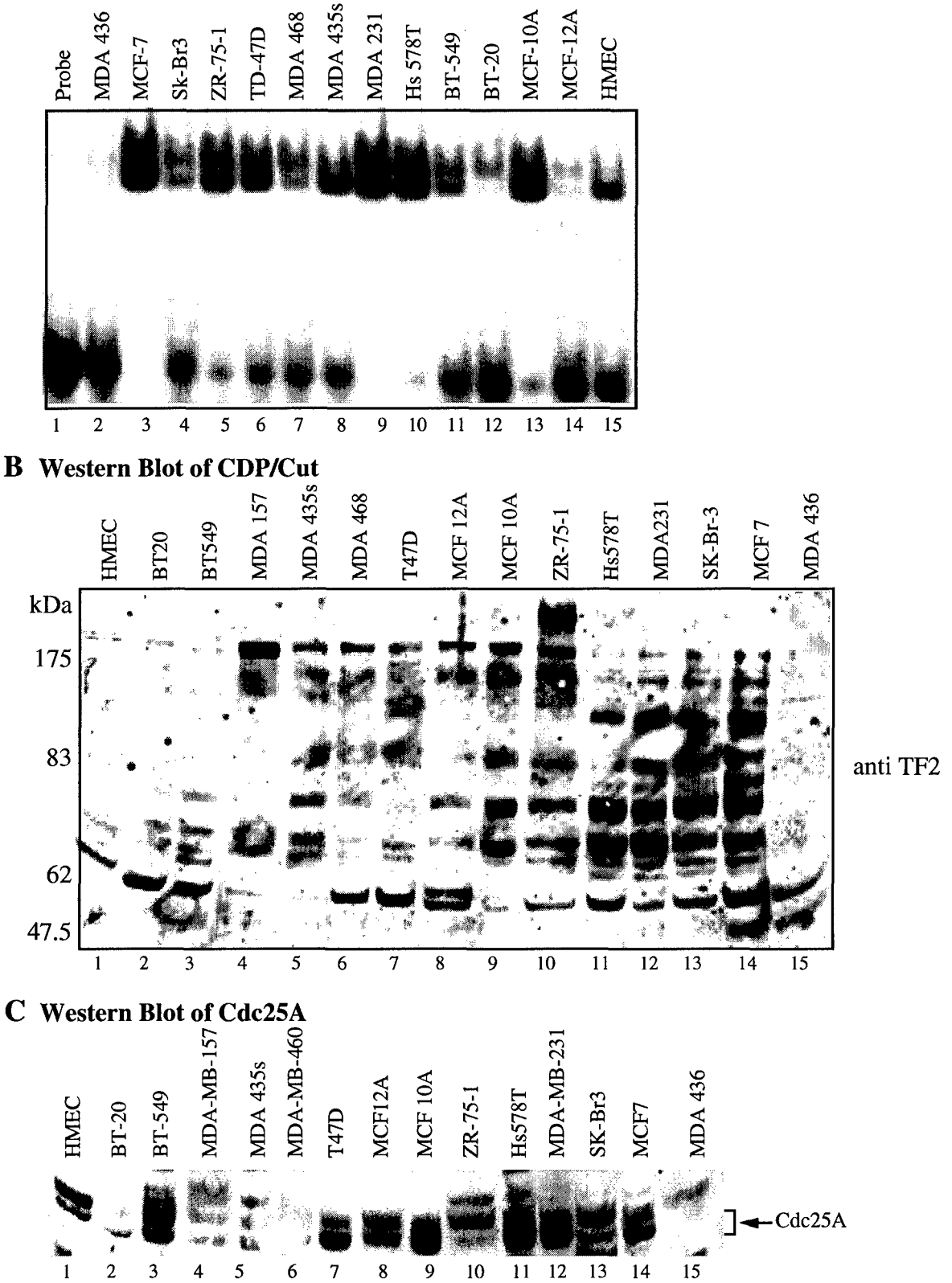
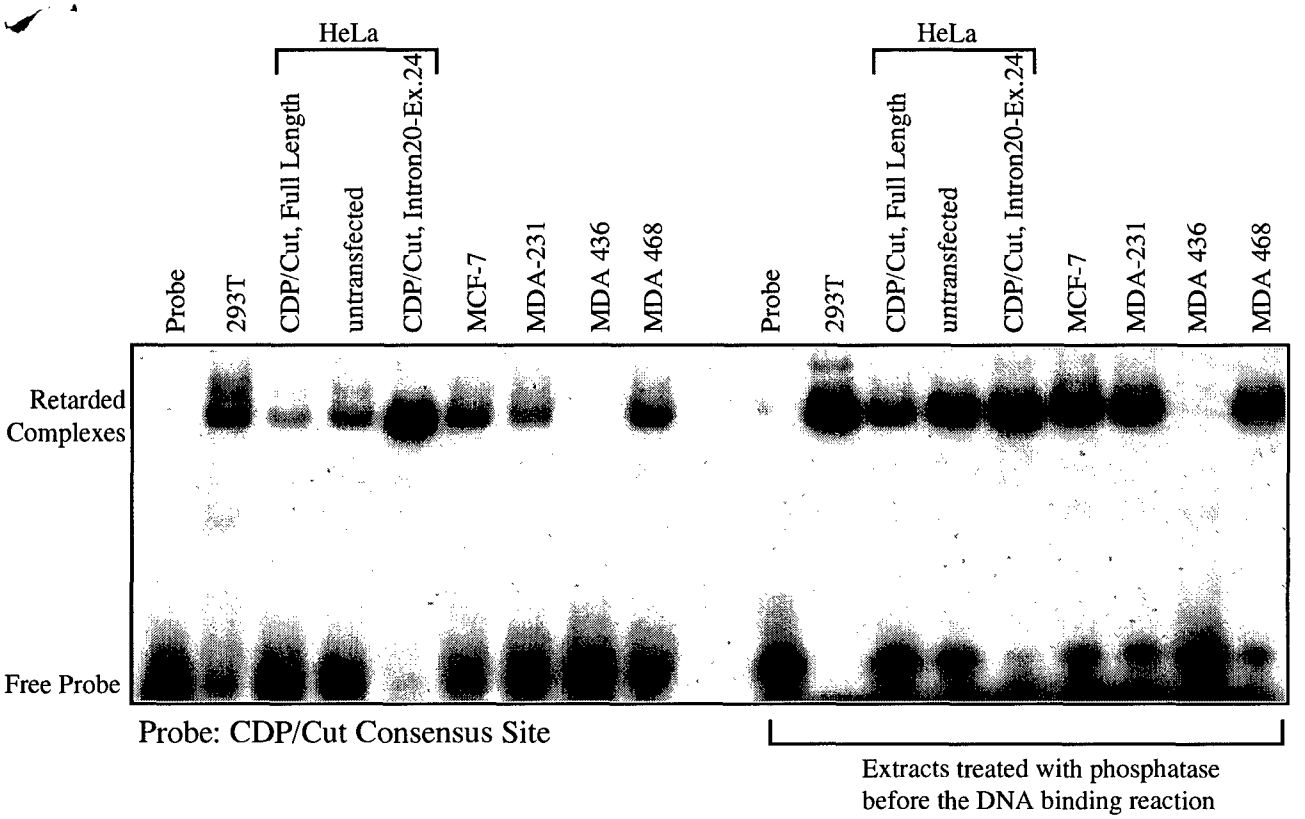
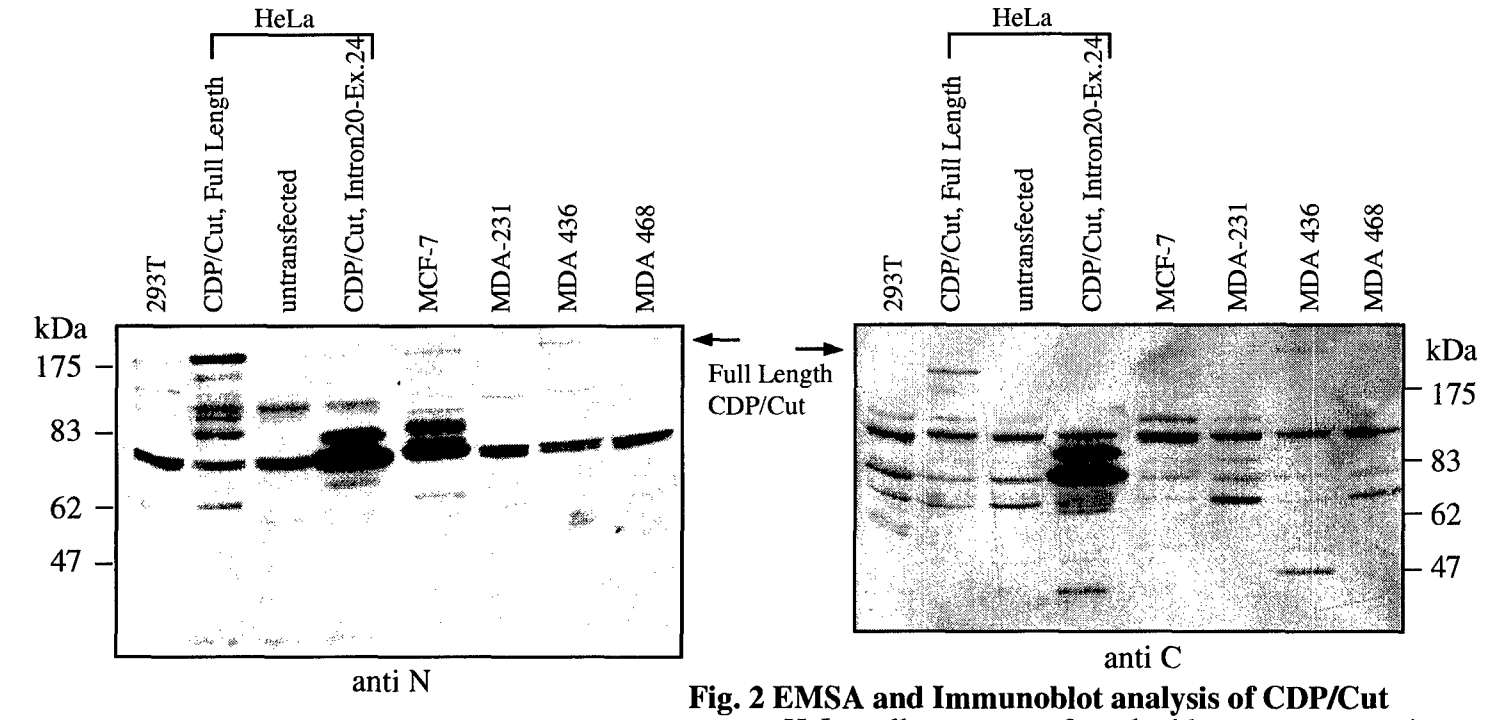


Fig. 1 Electrophoretic mobility shift assay (EMSA) and immunoblotting analysis of CDP/Cut in human mammary cell extracts

Cellular extracts were prepared from human mammary cells and used in electrophoretic mobility shift assays (EMSA) using a CDP/Cut consensus binding site **(A)** and in immunoblotting using anti-CDP/Cut antibodies **(B)**, or anti-Cdc25A antibodies **(C)**. Note that there is no correlation between DNA binding and the abundance of the full length CDP/Cut protein. Which of the bands revealed by anti-CDP/Cut antibodies is responsible for CDP/Cut DNA binding activity cannot be determined with TF2 antibodies (see Fig. 2 for the diagram of TF2 Ab).





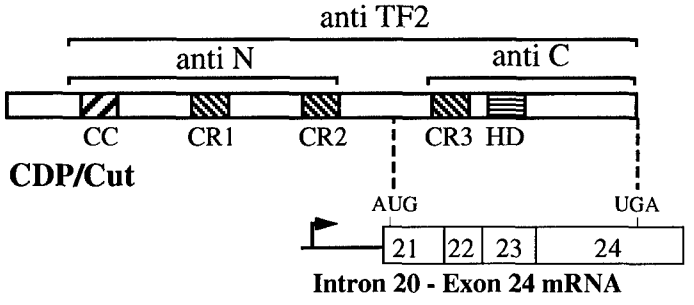


Fig. 2 EMSA and Immunoblot analysis of CDP/Cut

HeLa cells were transfected with vectors expressing either full length CDP/Cut or a mRNA initiated within intron 20. Cellular extracts were prepared from HeLa and breast tumor cell lines and used in electrophoretic mobility shift assays (EMSA) using a CDP/Cut consensus binding site (A) and in immunoblotting using anti-CDP/Cut antibodies (B).

Note that extracts from transfected cells generated retarded complexes that migrated close to each other. Also, which of the bands revealed by anti-CDP/Cut antibodies is responsible for CDP/Cut DNA binding activity cannot be determined with certitude using these antibodies.

Task 7 Months 12-24 Investigate expression of CUTL1 in breast tumors.
In tumors with LOH, establish if the remaining CUTL1 allele is expressed.

Using reverse-transcriptase polymerase chain amplification (RT-PCR), we analyzed CDP/Cut mRNA expression in tumors with LOH of CUTL1 (Fig. 3). The results demonstrated that CDP/Cut mRNA are expressed in breast tumors with LOH of CUTL1. In addition to the full length CDP/Cut mRNA transcript, we detected expression of a CDP/Cut mRNA which is initiated within intron 20. Interestingly, expression of this mRNA was elevated in two breast tumors with an invasive phenotype (Fig. 3A, lanes 7 and 13). In normal tissues, RNase mapping analysis indicated that this novel mRNA is expressed in placenta (Fig. 3B).

Task 8 Months 12-36 In breast tumors with CUTL1 LOH, determine whether the remaining allele is mutated.
Months 12-24 For the tumors for which RNA is available, clone CUTL1 cDNAs by RT-PCR.

Using RT-PCR, we have cloned CUTL1 cDNAs from 5 tumors. DNA sequencing was performed over the regions encoding the CDP/Cut functional domains. From the 5' end (or amino-terminal end), these regions include the coiled-coil region, Cut repeat 1, Cut repeat 2, Cut repeat 3, the Cut homeodomain, and the carboxy-terminal repression domains (Fig. 4). Two strategies were used for DNA sequencing. Either PCR-sequencing or isolation of individual cDNA clones followed by DNA sequencing of 3 individual clones for each breast tumor. Similar results were obtained by these two strategies. No mutation was identified, however, a number of polymorphisms were identified.

Months 12-24 Determine the CUTL1 exon/intron structure, including the sequence at the 5' and 3' end of introns. Design and test primers for PCR-amplification of CUTL1 exons.

We have completed the CUTL1 exon/intron structure, including the sequence at the 5' and 3' end of introns. A manuscript has been sent for publication (see Appendix 2). In short, the gene spans at least 340 Kbp and contains 33 exons (Appendix 2, Fig. 1). Synthesis of five different transcripts involves 2 promoter regions, 2 polyadenylation sites and 7 alternative splicing events. The two polyadenylation sites are located at the ends of exons 24 and 33 and are separated by approximately 40 Kbp. Transcription is initiated in two genomic regions, giving rise to alternate first exons which are spliced to a common exon 2. All transcripts contain exons 2 to 14, but differ in their 3' regions. Exon 14 can be spliced alternatively to the beginning or the middle of exon 15 (CDP/Cut mRNAs), or to exon 25 (CASP mRNA), generating transcripts with exons 15 to 24 or exons 25 to 33. Moreover, exon 16 can be spliced out from the mature transcripts that contain exons 15 to 24. Overall, 5 distinct transcripts are generated as a result of alternative transcription initiation, splicing and polyadenylation.

Months 24-36 Sequence analysis of CUTL1 genomic sequences and cDNAs.
Screening of mutations in CUTL1 genomic sequences and cDNAs, using dideoxy fingerprinting (ddF).
Confirmation of mutations by direct sequencing.

Our plan was first to analyze the cDNA sequence from a number of breast tumors with CUTL1 LOH in order to verify whether mutations were present which would alter the coding information. We were hoping to identify a region of the molecule that was a "hotspot" for mutations. We were going then to analyze the same region in other breast tumors this time using genomic DNA and performing PCR-sequencing of the appropriate exons. As mentioned above, so far no mutation has been found. As a result, we have not performed sequence analysis of CUTL1 genomic sequences.

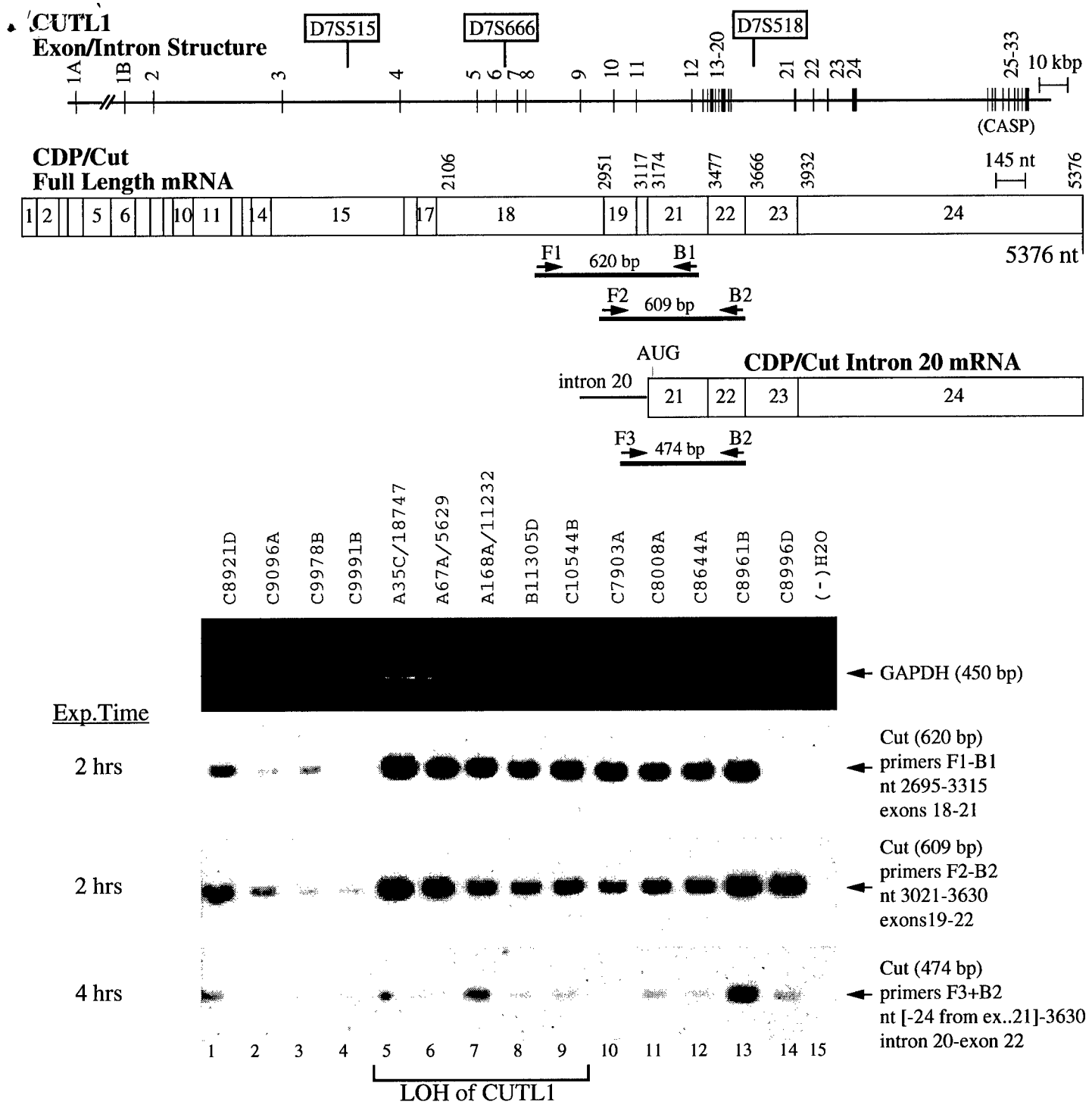
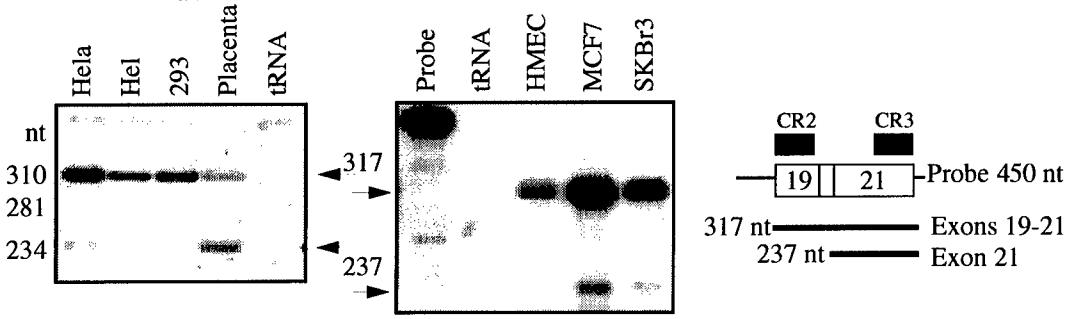


Fig. 3 (A) RT-PCR analysis of CDP/Cut mRNA in breast tumors

Total RNA was prepared from slices of breast tumors and used in reverse transcription-polymerase chain reaction (RT-PCR) using primers derived from exon or intron sequences as indicated. In the third panel, a novel mRNA that is initiated within intron 20 is revealed using primers from intron 20 and exon 22. Note that the large size of intron 21 excludes that the amplified fragments originate from genomic DNA or unspliced transcripts.

B RNase Mapping of CDP/Cut mRNAs



(B) RNase mapping using a cDNA probe encompassing exons 19 to 21. A protected fragment corresponding to a transcript that does not include exon 19 and 20 is detected in placenta and in a number of tumor cell lines.

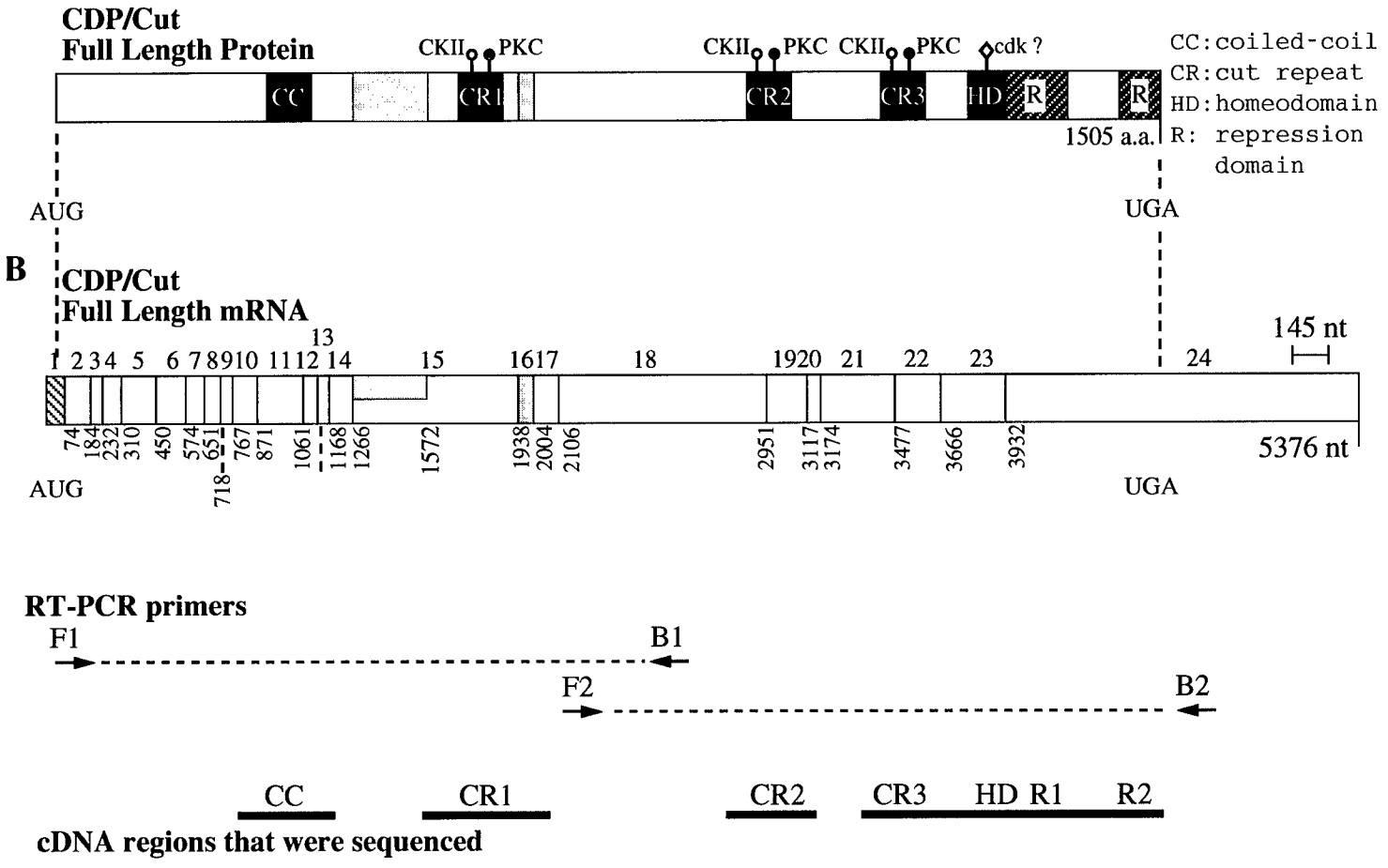


Fig. 4 DNA sequence analysis of CDP/Cut cDNA isolated from breast tumors with CUTL1 LOH

CDP/Cut cDNA fragments were generated by RT-PCR using primers F1-B1 and F2-B2, as indicated. The cDNA fragments were either cloned into the pKS plasmid or used as substrates in PCR-sequencing using a second set of primers. For each cDNA fragment cloned, three individual clones were analyzed by DNA sequencing. The regions that were sequenced are indicated at the bottom.

Task 9 Months 12-36 Preparation of expression vectors to test the functionality of CUTL1 products expressed in breast tumors.

Test the DNA binding activity of CUTL1 products in breast tumors: perform transfections and test cell extracts in electrophoretic mobility shift assays.

Test repression potential of CUTL1 products in breast tumors: perform cotransfection studies using appropriate effector and reporter constructs.

As described in Task 8, using RT-PCR, we have cloned CUTL1 cDNAs from 5 tumors. Since the length of the coding information is slightly more than 4500 nucleotides, the cDNAs were cloned in two pieces of approximately 2500 bp. PCR primers were chosen such that the 5' and 3' fragments would overlap and include the unique HindIII site at position 2273. The full length coding information was reconstituted following ligation of the 5' and 3' fragments. These fragments were then inserted into a mammalian expression plasmid. The expression vectors were introduced into NIH3T3 cells by transient transfection in conjunction with reporter constructs (either MEC-ME1a1-CAT or p21-luciferase. Protein extracts were prepared and processed to measure luciferase activity and CDP/Cut DNA binding activity by EMSA. All CDP/Cut cDNA clones behaved in a similar manner to our "normal" CDP/Cut cDNA isolated from placenta. These results confirmed and extended our mutation screening results: not only were no mutation found, but also the CDP/Cut protein expressed in breast tumors with LOH of CUTL1 behaved like the wild type CDP/Cut.

KEY RESEARCH ACCOMPLISHMENTS

1. We have defined a second region of LOH on the long arm of chromosome 7, within 7q22. These results were consistent with the notion that CUTL1 is a tumor suppressor gene. If CUTL1 is not the tumor suppressor gene, at the minimum our results provide mapping information that will be useful for an eventual positional cloning strategy.

2. We found no correlation between LOH at 7q22 and tumor type or grade, ER or PR expression, nodal status, or age of onset. However, there was a significant association (p value = 0.0082) between LOH at 7q22 and tumor size. The average sizes of breast tumors with or without LOH at 7q22 were 4.99 and 2.93 cm, respectively.

3. We demonstrated that CDP/Cut mRNA was expressed in breast tumors with LOH of CUTL1. In 5 breast tumors with LOH of CUTL1, the other allele was found to be expressed. When CDP/Cut cDNAs were obtained from these tumors and expressed in NIH3T3, the CDP/Cut proteins encoded by these cDNAs appeared to be functional in EMSA and reporter assays: they bound DNA and repress transcription just like the wild type CDP/Cut protein.

4. In breast tumor cell lines, CDP/Cut DNA binding did not correlate with the level of full length CDP/Cut protein, even when extracts were treated by phosphatase. CDP/Cut DNA binding activity was found to be due to a protein that is shorter than the full length CDP/Cut protein.

5. We found no correlation between c-Myc expression and CDP/Cut DNA binding activity.

6. We identified a novel CUTL1 mRNA, initiated within intron 20, that was found to be expressed at higher level in two breast tumors with an invasive phenotype and features of lobular carcinomas.

7. We have completed the CUTL1 exon/intron structure, including the sequence at the 5' and 3' end of introns.

REPORTABLE OUTCOMES

Manuscripts

Zeng, W. R., P. Watson, J. Lin, S. Jothy, R. Lidereau, M. Park, and A. Nepveu. Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene. **Oncogene** 18: 2015-2021. 1999.

Zeng, W. R., E. Soucie, N.S. Moon, N. Martin-Soudant, G. Bérubé and A. Nepveu. Exon/intron structure and alternative transcripts of the CUTL1 gene. Submitted to Gene.

Degrees/Training

Wendy R. Zeng obtained her Ph.D. in May 1999. Although she received a studentship from a local agency within McGill University and was not funded by this grant, some of the reagents she used were paid for with money form this grant. She finished her Ph.D. with three first author papers and two co-author papers. Work for two of her first author papers (listed above) was performed as part of this project.

CONCLUSIONS

LOH of CUTL1 Will Not Have Much Prognostic Value

Our work has established that a fraction (~18%) of spontaneous breast cancers exhibit LOH of 7q22. We were hoping that LOH of 7q22 could serve as a prognostic marker that would be useful in predicting the outcome of node-negative breast cancers. We tested this idea and, unfortunately, our results demonstrated that LOH of 7q22 correlates only with increased tumor size. I therefore conclude that LOH of 7q22 does not have much prognostic value.

Another hypothesis we tested was whether LOH of CUTL1, or reduced CDP/Cut activity, correlated with increased c-Myc expression. Our results demonstrated that this is not the case.

In the course of this study, a number of novel findings were made which raised important issues. Below, I will summarize these results and discuss their significance. Although some of these experiments relate to the basic mechanisms of CDP/Cut function, and as such are funded by another agency (MRC Canada), the knowledge we acquired through these studies has important implications on our investigation of CDP/Cut in breast tumors. Therefore, I will summarized some of our results and discuss how these affect our investigation of CDP/Cut and c-Myc in breast tumors. In a separate document, I will submit a Revised Statement of Work for the third year of this grant.

Different CUTL1 Gene Products May Have Distinct Biological Functions

We have shown that CUTL1 is present within the 7q22 chromosomal region that is deleted in 15% of human uterine leiomyomas and 18% of breast cancers(5, 15). Moreover, in breast tumors that develop in MMTV-Large T transgenic mice, CDP/Cut and Large T were found to form a complex in co-immunoprecipitation experiments. Finally, in transfection assays we have previously found that forced expression of the full length CDP/Cut protein greatly reduced the number of stable G-418 resistant colonies. In addition, when transformants were analyzed for the presence and expression of exogenous CDP/Cut, it was found that the CDP/Cut expressing vector was either not present or not expressed. Altogether, these results were consistent with the notion that CUTL1 is a candidate tumor suppressor gene. Recent findings, however, suggested that this issue is more complex than expected. In particular, the CUTL1 gene was found to encode for several gene products and the effect of cellular proliferation appears to depend on which gene product is expressed.

A number of results suggested that CDP/Cut activity may not have adverse effects on cellular proliferation. First, CDP/Cut DNA binding activity was found to be elevated in proliferating cells and, in most cell lineages, it decreased as cells became terminally differentiated (16-19). Secondly, in fibroblastic cells CDP/Cut DNA binding activity is minimal in G0, is activated at the end of G1 and is maximal during S phase (11). One target of CDP/Cut repression was shown to be the gene encoding for the $p21^{WAF1/CIP1/SDI1}$ cyclin kinase inhibitor which, when expressed at high levels, can block cell cycle progression in G1. Thirdly, we compared the growth properties of mammary epithelial cells with weak and strong CDP/Cut DNA binding activity. Four pairs of human mammary epithelial cell lines with various tumorigenic potential were analyzed. Pairwise comparison of breast tumor cell lines revealed that those cell lines with higher CDP/Cut DNA binding activity proliferated faster and to a higher saturation density (Fig. 5).

A Shorter CDP/Cut Protein Is Responsible for CDP/Cut DNA Binding Activity

In breast tumor cell lines, CDP/Cut DNA binding did not correlate with the level of full length CDP/Cut protein, even when extracts were treated by phosphatase (described under Task 5, Fig. 2). These findings led us to investigate further the relationship between CDP/Cut protein expression and DNA binding activity.

Our recent results revealed that CDP/Cut DNA binding activity, which we monitored in electrophoretic mobility shift assays (EMSA) using a CDP/Cut consensus binding site and specific anti-CDP/Cut antibodies, is due to a shorter protein than the full length CDP/Cut. Using the method of Orchard and May, we found that the molecular weight of the retarded complex generated by the endogenous CDP/Cut was lower than that of CDP/Cut itself. These results led us to transfect NIH3T3 cells with a series of vectors expressing CDP/Cut proteins of variable lengths in order to investigate the retarded complexes generated by these proteins. As seen in Fig. 6A, a protein truncated upstream of the Cut repeat 2 (CR2) produced a retarded complex that comigrated with that of the endogenous CDP/Cut protein, while a protein truncated upstream of CR3 generated a slightly faster migrating protein-DNA

Growth Curves of Breast Cell Lines with High and Low CDP/Cut DNA Binding Activity

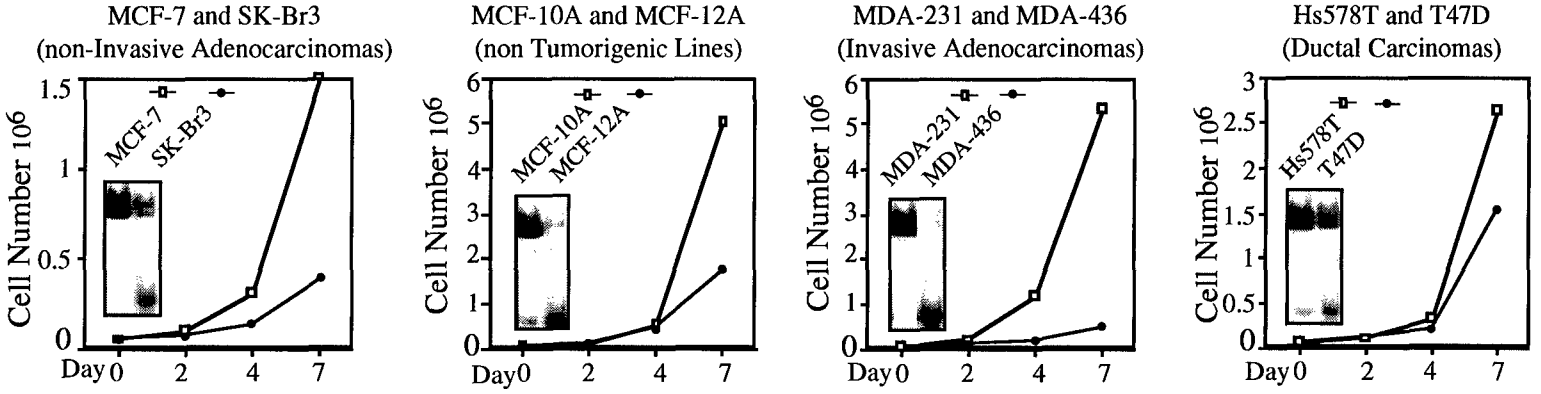


Fig. 5 Growth Curves of Breast Cell Lines with High and Low CDP/Cut DNA Binding Activity

Four pairs of breast cell lines with similar tumorigenic properties but different CDP/Cut DNA binding activity were cultured until saturation. Cells were counted on the indicated days. Cellular extracts were prepared from exponentially growing cells and used in electrophoretic mobility shift assays (EMSA) using a CDP/Cut consensus binding site. Invariably, those cell lines with higher CDP/Cut DNA binding activity were less contact inhibited and reached a higher saturation density.

A

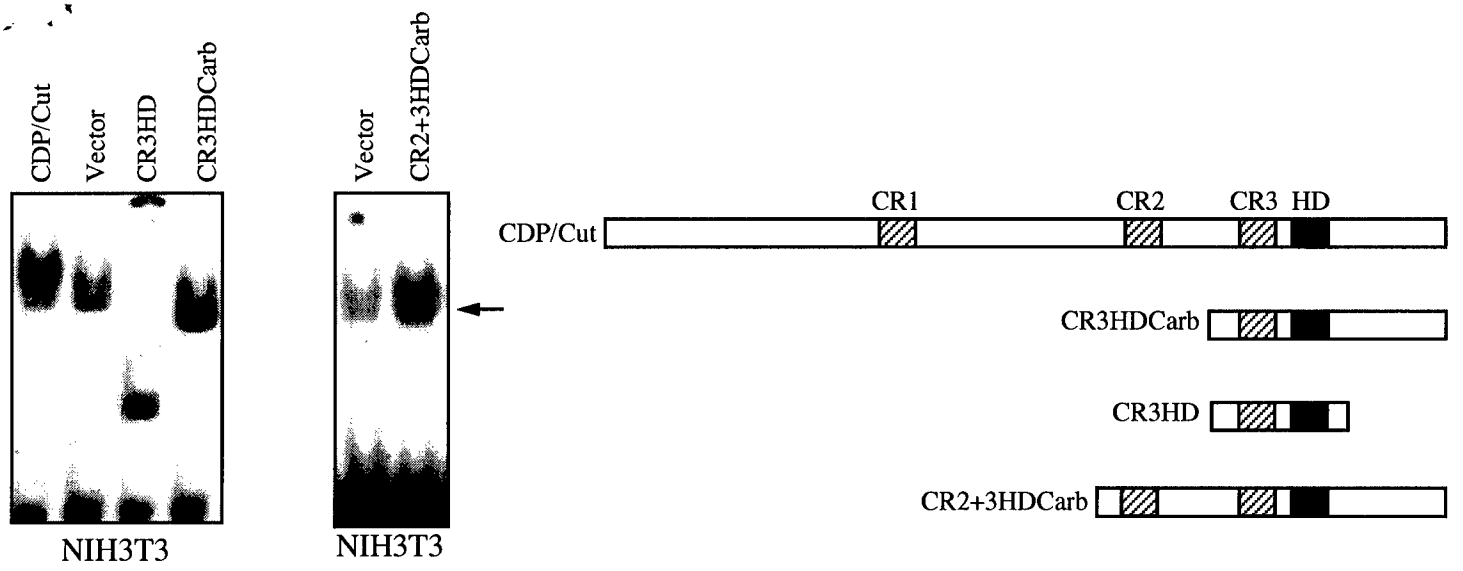


Fig. 6A A shorter CDP/Cut protein is responsible for CDP/Cut DNA binding activity

NIH3T3 cells were transfected with a series of vectors expressing CDP/Cut proteins of variable lengths. Cellular extracts were prepared and tested in EMSA with a CDP/Cut consensus binding site. Note that the retarded complexes generated by the "CR2CR3HDCarb" protein comigrates with that from the endogenous CDP/Cut protein.

B **Processing of Epitope-Tagged CDP/Cut Proteins in NIH3T3**

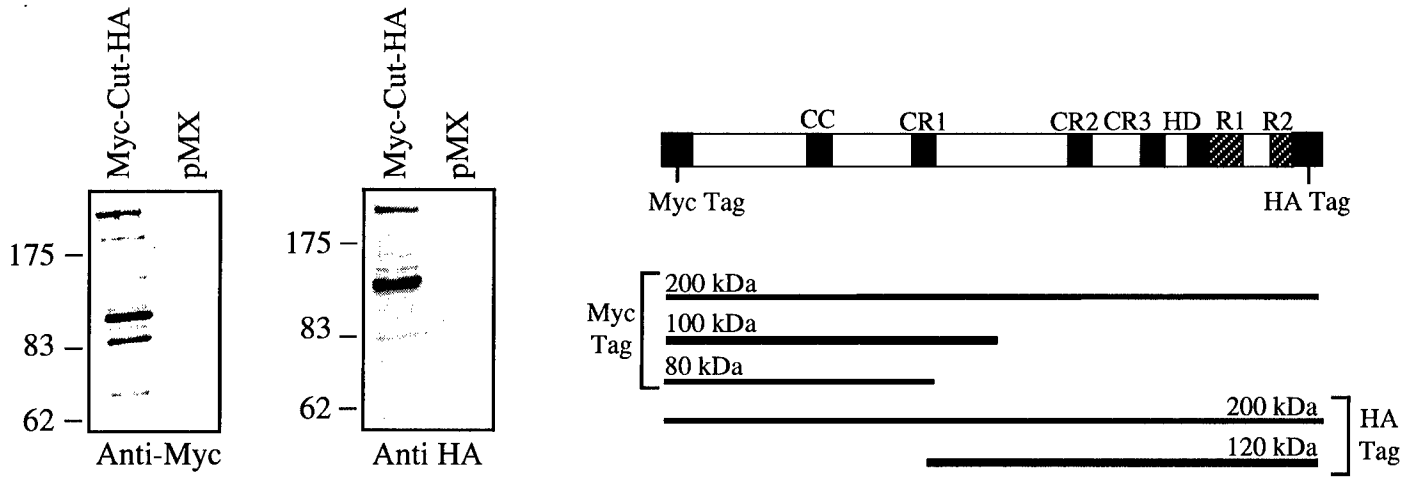


Fig. 6B Processing of epitope-tagged CDP/Cut proteins in NIH3T3

NIH3T3 cells were transfected with an empty vector (pMX) or a vector expressing a CDP/Cut protein with epitope tags at its amino and car boxy-terminal ends. Extracts were prepared and analyzed by immunoblotting with anti-Myc and anti-HA antibodies. Shown on the diagram are the approximate sizes of the proteins detected by immunoblotting. We conclude that shorter CDP/Cut proteins are generated by processing of the full length protein. However, in the case of the 120 kDa protein detected with the anti-HA antibody, we cannot exclude the possibility that this protein is generated by translation initiation at an internal AUG site since the HA-tag is located at the carboxy-terminus.

complex (see also Fig. 2). These results are in accordance with the findings that when purified from a baculovirus expression system the full length CDP/Cut protein could bind only weakly to DNA, while a truncated protein bound DNA with high affinity. We are currently testing the hypothesis that the amino-terminal domain allosterically inhibits the DNA binding activity of CR3HD.

Two Alternative Mechanisms May Lead to the Production of Shorter CDP/Cut Proteins

In fibroblastic cells, shorter CDP/Cut proteins appear to be generated by proteolytic processing of the full length protein as demonstrated by the expression of a CDP/Cut protein with epitope tags at its amino and carboxy-terminal ends (Fig. 6B). This findings, together with our results from DNA binding studies, lead us to hypothesize that the activation of CDP/Cut DNA binding activity in S phase would result from two events: not only does the Cut homeodomain need to be dephosphorylated by the Cdc25A phosphatase, but in addition the full length CDP/Cut protein must be proteolytically processed into a shorter form that is truncated at its amino-terminus.

Processing of the full length CDP/Cut protein, however, is not the only mechanism leading to the expression of shorter CDP/Cut proteins. In placenta, which is composed largely of trophoblasts undergoing endomitosis, a novel mRNA was found to be initiated within intron 20 (Fig. 3B). This mRNA encodes for a short CDP/Cut protein corresponding to a little less than the carboxy-terminal half (see diagram in Fig. 3A). Similar transcripts with an opened reading frame starting within exon 21 have been described in mouse testis (20).

In summary, shorter CDP/Cut proteins can be expressed by at least two mechanisms: processing of the full length protein or synthesis of a mRNA that is shorter at its 5' end.

A Novel CDP/Cut mRNA Initiated From Intron 20 Is Expressed in Some Breast Tumor Cell Lines

Some tumor cell lines express a shorter CDP/Cut mRNA initiated within intron 20 (Fig. 3B). A cDNA corresponding to this novel mRNA was cloned and tested in transient transfections. A protein of an apparent molecular weight of 80 kDa was expressed, demonstrating that mRNA initiated in intron 20 are efficiently translated to produce a shorter CDP/Cut protein (Fig. 2B). In electrophoretic mobility shift assay, the retarded complex generated by this short CDP/Cut protein migrated slightly below that from the endogenous CDP/Cut protein in various normal or breast tumor cell lines (Fig. 2A).

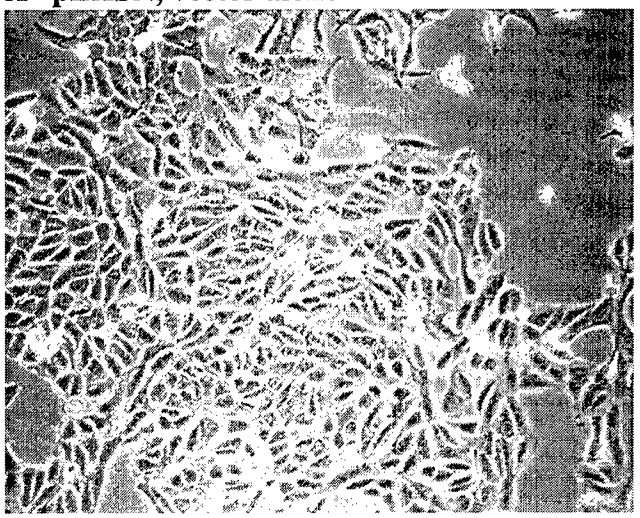
The Intron 20-Initiated CDP/Cut mRNA May Induce a Transition from an Epithelial to Mesenchymal Phenotype

CDP/Cut DNA binding activity correlated with higher saturation density in breast tumor cell lines (Fig. 2). On the other hand, we found that proteins shorter than the full length CDP/Cut protein are responsible for CDP/Cut DNA binding activity (Fig. 6). Since the intron 20-initiated mRNA encodes for a CDP/Cut protein that is active in DNA binding, we asked whether overexpression of this transcript would confer higher proliferative potential to those breast tumor cell lines with low CDP/Cut DNA binding activity and low saturation density. We engineered a retroviral vector encoding for the intron 20-initiated mRNA, produced amphotropic retroviruses, infected the MCF-12A breast cell line and selected for G-418 resistance. To our surprise, this vector seemed to change the phenotype of the infected cells (Fig. 7). Clones obtained following infection with an empty retroviral vector grew as tightly packed, organized, colonies. In contrast, approximately half of the G-418 resistant colonies obtained with the vector expressing the intron 20-initiated mRNA exhibited a morphology different from that of the parental cells. Colonies were not as densely packed, because cells were not tightly attached to each other. Cell dissociation is likely to result from the breakdown of cell:cell adherens junction (E-cadherin). These changes correspond to a transition from an epithelial morphology towards a more mesenchymal fibroblastic phenotype, referred to as epithelial-mesenchymal transition (E-M transition) (21, 22).

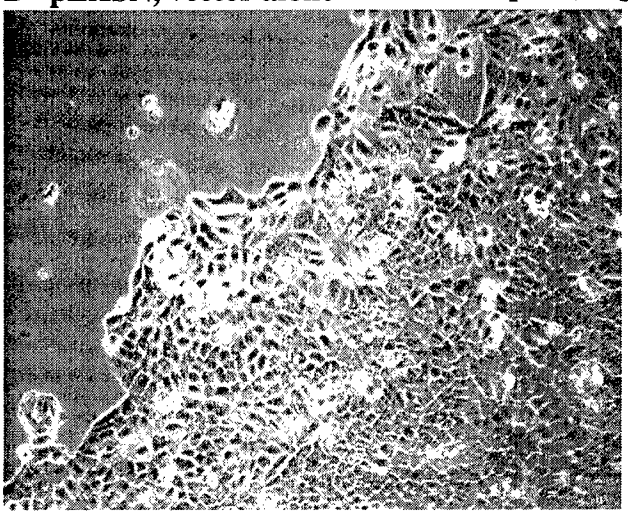
The Intron 20-Initiated CDP/Cut mRNA Is Expressed at Higher Level in Two Breast Tumors With Properties of Lobular Carcinoma

Using RNase mapping analysis, we detected the presence of a novel CDP/Cut mRNA in some breast tumor cell lines (Fig. 3B). Cloning of the cDNA for this mRNA, followed by DNA sequencing analysis, indicated that it is initiated within intron 20 and has the potential to encode a shorter CDP/Cut protein containing the Cut repeat 3, the Cut homeodomain and the carboxy-terminal domain. We designed primers to detect this mRNA in breast tumors using RT-PCR. We found that the intron 20

A pLXSN, vector alone

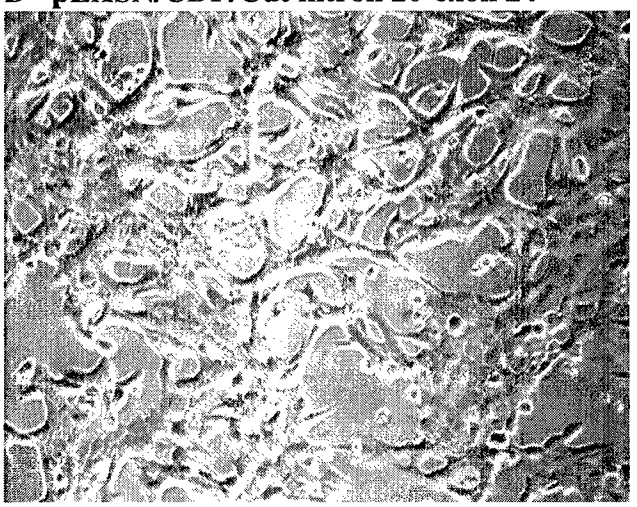


B pLXSN, vector alone

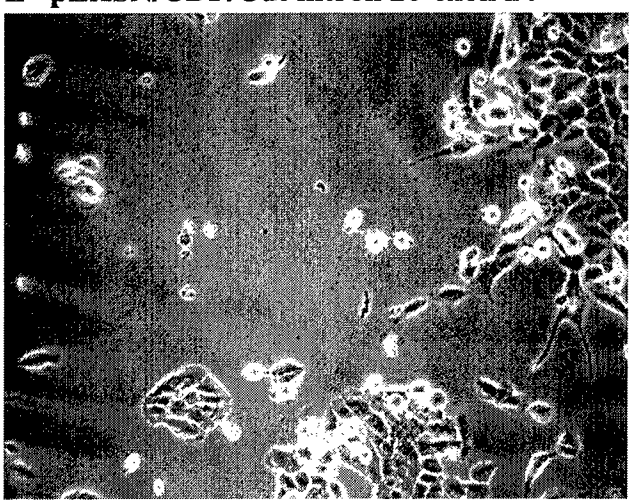


C pLXSN/CDP/Cut intron 20-exon 24

D pLXSN/CDP/Cut intron 20-exon 24



E pLXSN/CDP/Cut intron 20-exon 24



F pLXSN/CDP/Cut intron 20-exon 24

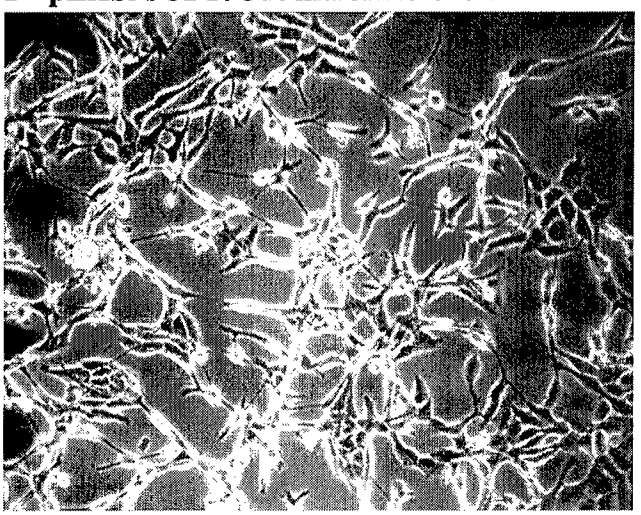


Fig. 7 Overexpression of the intron 20-initiated CDP/Cut cDNA causes morphological changes in the MCF-12A cell line.

Cells from a non tumorigenic mammary epithelial cell line, MCF-12A, were infected with a retroviral vector expressing the intron 20-initiated CDP/Cut cDNA (panels C to F). As a control, cells were infected with an empty retroviral vector (panels A and B). The medium was supplemented with G-418 to select for resistant colonies. After 14 days in culture, photographs were taken at 32x magnification. Note that cells infected with the vector alone form densely packed foci, whereas those infected with the vector expressing the intron 20-initiated CDP/Cut cDNA grow as isolated cells.

mRNA was expressed at higher levels in two tumors: A168A and C8961B. Interestingly, A168 is an invasive ductal carcinoma, but with a very diffuse growth pattern like that of an invasive lobular carcinoma, and C8961B is an invasive lobular carcinoma.

These observations together with the change in phenotype of MCF-12A infected with a retrovirus expressing the intron 20 CDP/Cut mRNA suggest that the expression of this form of CDP/Cut may be associated with distinct biological properties.

REFERENCES

1. Dufort, D. and Nepveu, A. The human cut homeodomain protein represses transcription from the c-myc promoter, Molecular & Cellular Biology. *14:* 4251-7, 1994.
2. Lemieux, N., Zhang, X. X., Dufort, D., and Nepveu, A. Assignment of the human homologue of the Drosophila Cut homeobox gene (CUTL1) to band 7q22 by fluorescence in situ hybridization, Genomics. *24:* 191-3, 1994.
3. Scherer, S. W., Neufeld, E. J., Lievens, P. M., Orkin, S. H., Kim, J., and Tsui, L. C. Regional localization of the CCAAT displacement protein gene (CUTL1) to 7q22 by analysis of somatic cell hybrids, Genomics. *15:* 695-696, 1993.
4. Webster, M. A., Martinsoudant, N., Nepveu, A., Cardiff, R. D., and Muller, W. J. The Induction Of Uterine Leiomyomas and Mammary Tumors In Transgenic Mice Expressing Polyomavirus (Pyv) Large T (Lt) Antigen Is Associated With the Ability Of Pyv Lt Antigen to Form Specific Complexes With Retinoblastoma and Cutl1 Family Members, Oncogene. *16:* 1963-1972, 1998.
5. Zeng, W. R., Watson, P., Lin, J., Jothy, S., Lidereau, R., Park, M., and Nepveu, A. Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene, Oncogene. *18:* 2015-2021, 1999.
6. Mann, H. B. and Whitney, D. R. On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other, Annals of Mathematical Statistics. *18:* 50-60, 1947.
7. Andres, V., Chiara, M. D., and Mahdavi, V. A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins, Genes & Development. *8:* 245-57, 1994.
8. Aufiero, B., Neufeld, E. J., and Orkin, S. H. Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein, Proc. Natl. Acad. Sci. USA. *91:* 7757-7761, 1994.
9. Harada, R., Dufort, D., Denis-Larose, C., and Nepveu, A. Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains, Journal of Biological Chemistry. *269:* 2062-7, 1994.
10. Harada, R., Berube, G., Tamplin, O. J., Denis-Larose, C., and Nepveu, A. DNA-binding specificity of the cut repeats from the human cut-like protein, Molecular & Cellular Biology. *15:* 129-40, 1995.
11. Coqueret, O., Berube, G., and Nepveu, A. The Mammalian Cut Homeodomain Protein Functions As a Cell-Cycle-Dependent Transcriptional Repressor Which Downmodulates P21(Waf1/Cip1/Sdi1) In S Phase, EMBO Journal. *17:* 4680-4694, 1998.
12. Coqueret, O., Martin, N., Berube, G., Rabbat, M., Litchfield, D. W., and Nepveu, A. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Casein Kinase II, Journal of Biological Chemistry. *273:* 2561-2566, 1998.
13. Coqueret, O., Berube, G., and Nepveu, A. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Protein Kinase C, Journal of Biological Chemistry. *271:* 24862-24868, 1996.
14. Galaktionov, K., Lee, A. K., Eckstein, J., Draetta, G., Meckler, J., Loda, M., and Beach, D. CDC25 phosphatases as potential human oncogenes, Science. *269:* 1575-7, 1995.
15. Zeng, W. R., Scherer, S. W., Koutsilieris, M., Huizenga, J. J., Filteau, F., Tsui, L. C., and Nepveu, A. Loss Of Heterozygosity and Reduced Expression Of the Cutl1 Gene In Uterine Leiomyomas, Oncogene. *14:* 2355-2365, 1997.
16. Lievens, P. M. J., Donady, J. J., Tufarelli, C., and Neufeld, E. J. Repressor activity of CCAAT displacement protein in HL-60 myeloid leukemia cells, J. Biol. Chem. *270:* 12745-12750, 1995.
17. Skalnik, D. G., Strauss, E. C., and Orkin, S. H. CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter, J Biol Chem. *266:* 16736-16744, 1991.

18. Heuvel, G. B. V., Bodmer, R., McConnell, K. R., Nagami, G. T., and Igarashi, P. Expression Of a Cut-Related Homeobox Gene In Developing and Polycystic Mouse Kidney, Kidney International. *50:* 453-461, 1996.
19. Guy, C. T., Muthuswamy, S. K., Cardiff, R. D., Soriano, P., and Muller, W. J. Activation of the c-Src tyrosine kinase is required for the induction of mammary tumors in transgenic mice, Genes Dev. *8:* 23-32, 1994.
20. Vandenheuvel, G. B., Quaggin, S. E., and Igarashi, P. A Unique Variant Of a Homeobox Gene Related to Drosophila Cut Is Expressed In Mouse Testis, Biology of Reproduction. *55:* 731-739, 1996.
21. Boyer, B., Valles, A. M., and Thiery, J. P. Model systems of epithelium-mesenchyme transitions, Acta Anatomica. *156:* 227-39, 1996.
22. Birchmeier, W., Behrens, J., Weidner, K. M., Hulsken, J., and Birchmeier, C. Epithelial differentiation and the control of metastasis in carcinomas, Current Topics in Microbiology & Immunology. *213:* 117-35, 1996.

Oncogene (1999) 18, 2015–2021

http://www.stockton-press.co.uk/onc

Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene

Wendy Rong Zeng[1], Peter Watson[2], Jenny Lin[1], Serge Jothy[3], Rosette Lidereau[4], Morag Park[1,5] and Alain Nepveu*[,1,5]

[1]*Molecular Oncology Group, Departments of Medicine, Oncology, McGill University, Montreal, Quebec, Canada;* [2]*Department of Pathology and Physiology, Faculty of Medicine, University of Manitoba, Canada;* [3]*Department of Laboratory Medicine and Pathobiology, University of Toronto, Ontario, Canada;* [4]*Laboratoire d'Oncogénétique, Centre René Huguenin, 35 rue Daily, F-92211 St-Cloud, France;* [5]*Biochemistry (A.N.), McGill University, Montreal, Quebec, Canada*

In breast cancer, loss of heterozygosity (LOH) has been described on the long arm of chromosome 7, at band q31, suggesting the presence of a tumor suppressor gene in this region. In this study, we have identified a second region of LOH on 7q, at band 7q22. Deletion of genetic material at 7q22 was found in all tumor types and grades and was associated with increased tumor size. The region of LOH at 7q22 in every case included one or more of three polymorphic markers that are located within the CUTL1 gene. LOH of 7q22 has also been documented in the case of human uterine leiomyomas (Zeng *et al.*, 1997; Ishwad *et al.*, 1997). Interestingly, in both leiomyomas and mammary tumors induced in transgenic mice expressing the Polyomavirus (PyV) large T (LT) antigen, immunocomplexes of CUTL1 and PyV LT antigen were detected (Webster *et al.*, 1998). Altogether, genetic data in human breast cancer and biochemical analyses in breast tumors from transgenic mice suggest that CUTL1 is a candidate tumor suppressor gene.

Keywords: breast cancer; tumor suppressor gene; chromosome 7q22; CUTL1; human Cut homeobox gene; loss of heterozygosity

Introduction

Breast cancer accounts for almost 30% of cancer diagnoses and almost 20% of cancer-related deaths in women. It is the leading cause of death among women aged 40–55. In the United States, it affects more than one in ten women (Harris *et al.*, 1992). Although little is known about the etiology and pathogenesis of breast neoplasias, alterations of oncogenes and tumor suppressor genes are considered to be critical in the multistep process leading to the development of breast cancer.

In human breast cancer, genetic deletions have been demonstrated to be one of the major genetic abnormalities. LOH has been documented at several chromosomal locations, including 1p, 1q, 2p, 3p, 6q, 7q, 8q, 9q, 11p, 11q, 13q, 15q, 16q, 17p, 17q, 18p, 18q and 22q (reviewed by Bieche and Lidereau, 1995). A few of these regions of LOH have been shown to include a known tumor suppressor gene implicated in breast cancer, such as *p53* at chromosome 17q13 (Isobe *et al.*, 1986; McBride *et al.*, 1986; Miller *et al.*, 1986), *BRCA-1* at 17q21 (Miki *et al.*, 1994) and *BRCA-2* at 13q12-13 (Wooster *et al.*, 1995). Genome-wide LOH analysis to generate allelotypes of human breast cancer revealed that several chromosomal regions could be deleted in a single tumor but no common sets of deletions could be identified between different tumors (Devilee *et al.*, 1997; Kerangueven *et al.*, 1997; Larsson *et al.*, 1990; Sato *et al.*, 1990).

From cytogenetic studies, deletions within the long arm of chromosome 7 have been found in breast cancer and other types of tumors, including myeloid leukemias (reviewed in Fischer *et al.*, 1998), kidney carcinomas, colon carcinomas, ovarian carcinomas, lung carcinomas, head-and-neck carcinomas (Berker-karauzum *et al.*, 1998; Dave *et al.*, 1995; Lundgren, 1991; Lundgren *et al.*, 1992; Mertens *et al.*, 1997; Solinas-Toldo *et al.*, 1996; Storto *et al.*, 1990) and uterine leiomyomas (Ozisik *et al.*, 1993; Sargent *et al.*, 1994; Xing *et al.*, 1997). LOH analyses confirmed that genetic material on 7q is frequently deleted in breast cancer, particularly in the 7q31 region (Bieche *et al.*, 1997; Callahan *et al.*, 1993; Champeme *et al.*, 1995; Deng *et al.*, 1994; Devilee *et al.*, 1997; Kristjansson *et al.*, 1997; Lin *et al.*, 1996; Tougas *et al.*, 1996). Altogether these data have been taken to suggest that a tumor suppressor gene is located in this region, however, this gene remains to be identified.

In previous studies, the loss of heterozygosity (LOH) at 7q22 has been reported in a subset of uterine leiomyomas (Ishwad *et al.*, 1997; Zeng *et al.*, 1997), which are benign tumors of smooth muscle origin most often referred to as fibroids. The smallest commonly deleted region in leiomyomas included markers that are located within CUTL1 (Cut-like-1), a gene that is homologous to the *Drosophila melanogaster cut* gene. Genetic studies in *Drosophila melanogaster* indicated that *cut* is involved in determination and maintenance of cell identity in several tissues. CUTL1 encodes a homeodomain transcription factor that functions as a transcriptional repressor. In humans, the Cut protein was first identified as the CCAAT displacement protein (CDP), a repressor of the gp-91 phox gene in undifferentiated myeloid cells (Skalnik *et al.*, 1991). A CUTL1 cDNA was also isolated by screening a cDNA expression library within a probe derived from the

*Correspondence: A Nepveu
Received 1 June 1998; revised 29 September 1998; accepted 23 October 1998

c-Myc gene promoter (Dufort and Nepveu, 1994). The human Cut protein was shown to bind to the c-Myc promoter and repress its expression (Harada *et al.*, 1994; Mailly *et al.*, 1996). Other *cut*-related cDNAs have been isolated from several mammalian species including dog, mouse and rat and were respectively termed Clox (Cut-like homeobox), Cux-1 and Cux-2 (Cut homeobox) and CDP-2 (Andres *et al.*, 1992; Dufort and Nepveu, 1994; Neufeld *et al.*, 1992; Quaggin *et al.*, 1996; Valarche *et al.*, 1993; Yoon and Chikaraishi, 1994).

Female transgenic mice expressing the Polyomavirus (PyV) Large T (LT) antigen under the control of the mouse mammary tumor virus long terminal repeat (MMTV-LTR) frequently develop, in addition to mammary tumors, uterine leiomyomas (Webster *et al.*, 1998). Since we had identified LOH at CUTL1 in human uterine leiomyomas, we examined whether PyV LT antigen formed specific complexes not only with members of the retinoblastoma (Rb) family (p105Rb, p107, p130), but also with the mammalian Cut protein. The results of coimmunoprecipitation analyses revealed that specific complexes of Cut and PyV LT antigen could be detected in both leiomyomas and mammary tumors. The existence of such complexes suggested that the alterations in the function of the Cut protein may be an important event in the etiology of breast cancer. These results lead us to hypothesize that genetic alterations within the 7q22 region may also occur in human breast tumors. We report here that LOH at 7q22 is observed in a subset of sporadic breast cancers and is associated with increased tumor size.

Results

A subset of breast tumors exhibits LOH of 7q22 and/or 7q31

The present study was initiated following the observation that genetic material at 7q22 is deleted in a fraction of uterine leiomyomas. The smallest deleted region at 7q22 included three polymorphic markers which since then have been located within the CUTL1 gene (Figure 2) (Ishwad *et al.*, 1997; Zeng *et al.* 1997; and Zeng *et al.*, manuscript in preparation). Secondly, the product of the CUTL1 gene was found to associate with the PyV LT oncoprotein both in uterine leiomyomas and in breast tumors that arise at high frequency in MMTV-PyV LT transgenic mice (Webster *et al.*, 1998). We thus postulated that genetic alterations within the 7q22 region may occur not only in human uterine leiomyomas but also in human breast tumors.

As a first step to verify whether LOH of 7q22 may occur in breast cancer, we analysed breast tumor DNA samples that had previously been characterized regarding LOH of 7q31, the chromosomal region adjacent to 7q22. We and others have shown that within 7q22 the marker D7S518 is the most frequently deleted in uterine leiomyomas. Thus, we first asked whether breast tumors with or without LOH of 7q31 could exhibit LOH of D7S518. This marker has been mapped to intron 20 of the CUTL1 gene (Zeng *et al.*, 1997; and Zeng *et al.*, manuscript in preparation). A total of 33 pairs of samples from Bièche *et al.* (1997) study were analysed. LOH of D7S518 was found in five of ten cases with LOH of 7q31 and in one out of 23 cases without LOH of 7q31 (Table 1). A total of 63 pairs of samples from the Lin *et al.* (1996) study were analysed. Among nine pairs of samples with LOH of 7q31, seven were informative for D7S518 and four exhibited LOH of this marker (Table 1). Among 54 pairs of samples without LOH of 7q31, two had suffered LOH of D7S518. Results obtained with these two sets of samples demonstrated that among breast tumors with LOH of the 7q31 region, 53% (9/17) also exhibit LOH of the D7S518 marker in 7q22. It is likely that the deleted region in these tumors is large and encompasses both 7q22 and 7q31. On the other hand, among breast tumors without LOH of 7q31, a small fraction suffered LOH of D7SD518. Thus, LOH of D7S518 can be observed independently of LOH at 7q31. We conclude that the long arm of chromosome 7 contains at least two tumor suppressor genes that can be inactivated in breast cancers, one at 7q22 and one at 7q31.

LOH of 7q22 and 7q31 in tumors from the NCIC-Manitoba breast tumor bank

To establish the incidence and extent of 7q22 deletions, we investigated LOH of 7q22 in samples from the National Cancer Institute of Canada (NCIC)-Manitoba Breast Tumor Bank (Watson *et al.*, 1996). We have analysed 66 pairs of tumors and adjacent normal control tissues representative of all classes of breast tumors. At least three microsatellite markers within 7q22, *D7S518, D7S515 and D7S666*, were used in all cases. In a previous study, we have mapped the marker *D7S518* within an intron of the CUTL1 gene and the markers *D7S515* and *D7S666* close to the 5′ end of this gene (Zeng *et al.*, 1997). Determination of the CUTL1 exon/intron structure and DNA sequencing of the long arm of chromosome 7 has since revealed that all three markers are located within the CUTL1 gene, in introns 20, 3 and 6 respectively (Figure 2) (Zeng *et al.*, manuscript in preparation; http://www.ncbi. nlm.nih.gov, accession number AF024533 and AF047825). Tumors with LOH of one or more of these markers were then further analysed to map the boundaries of the deleted regions. The physical order of all of the microsatellite markers and CUTl with respect to each other have previously been established (Zeng *et al.*, 1997). The number of cases studied for each marker, the number of informative

Table 1 Loss of heterozygosity of D7S518 in breast tumors with or without LOH of 7q31

Sources of samples	*7q31 status in breast cancers*	*Samples analysed*	*Informative cases for D7S518*	*LOH of D7S518*
Bièche *et al.*	LOH of 7q31	10	10	5 (50%)
Study	Retention of 7q31	23	23	1 (4%)
Lin *et al.*	LOH of 7q31	9	7	4 (57%)
Study	Retention of 7q31	54	54	2 (3.7%)

DNAs from breast cancers and matched normal tissues were amplified from patients with breast cancers using oligonucleotide primers for the polymorphic marker D7S518. The source of samples is indicated, together with the numbers of patients tested, informative patients and patients with LOH for markers on 7q31 or 7q22. The information regarding LOH at 7q31 was derived from the original studies

cases, and the number and percentage of cases that exhibit LOH are given in Table 2 and shown in Figure 2. Representative LOH results are shown in Figure 1. LOH was scored with at least one 7q22 marker in 12 of 66 (18.2%) of the tumors examined (patients 10544, 11305, 93-4635, 93-5199, 93-11232, 93-11747, 93-12017, 93-18747, 94-1663, 94-3808, 94-5629 and 94-133582, see Figure 2). The superposition of the overlapping deletions in the 12 tumors revealed two common regions of deletion, one in 7q22 and encompassing the CUTL1 gene, and one in 7q31 (Figure 2). The proximal (centromeric) and distal (telomeric) boundaries of the critical region in 7q22 were defined by breakpoints in tumors #10544, 11305, and #93-5199 that were flanked by the *D7S666* and *D7S658* markers, respectively (Figure 2). The proximal (centromeric) and distal (telomeric) boundaries of the critical region in 7q31 were flanked by the *D7S480* and *D7S650* markers, respectively (Figure 2).

Association between LOH at 7q22 and clinical parameters

To determine whether loss of genetic material at 7q22 or 7q31 may be associated with pathological features of breast tumors, we analysed available clinical data including the tumor type, grade and size, estrogen and progesterone receptor (ET and PR) status, occurrence of lymph node metastasis and age of onset. Statistical analysis was performed using the Mann–Whitney test (Mann and Whitney, 1947). We found no correlation between LOH at 7q22 and tumor type or grade, ER or PR expression, nodal status, or age of onset. However, there was a significant association (P value = 0.0082) between LOH at 7q22 and tumor size (Table 3). The average sizes of breast tumors with or without LOH at 7q22 were 4.99 and 2.93 cm respectively. The difference in size was further accentuated in the group of tumors with LOH at 7q22 but retention of 7q31 (6.5 Vs 2.99 cm; P value = 0.0074). In contrast, no significant correlation was found between LOH at 7q31 and any clinical feature.

Table 2 LOH analysis of 66 breast cancers using 11 polymorphic markers on chromosome 7

Markers	*Patients tested*	*Informative patients*	*Patients with LOH*
D7S524	12	5	2
D7S527	3	2	0
D7S518	66	59	6
D7S666	66	47	4
D7S515	66	48	8
D7S658	5	3	0
D7S471	1	1	0
D7S486	9	8	3
D7S522	66	45	3
D7S480	66	61	7
D7S650	66	48	7

DNAs from tumors and matched normal peripheral breast tissues were amplified from 66 patients with breast cancers using oligonucleotide primers for 11 polymorphic markers on chromsome 7q. The list of markers is presented, together with the numbers of patients tested, informative patients and patients with LOH for each marker. The level of informativeness observed for each of the markers in our cohort of patients was consistent with the published values

Discussion

LOH within the 7q31 region has previously reported in human breast carcinomas, and in several cases the deletions encompassed the adjacent region, 7q22 (Bieche *et al.*, 1992, 1997; Champeme *et al.*, 1995; Deng *et al*, 1994; Lin *et al.*, 1996; Zenklusen *et al.*, 1994). The results of the present study establish that in some breast tumors the loss of genetic material on the long arm of chromosome 7 can be limited to the 7q22 region. Thus, regarding chromosomal deletions on 7q, collectively the available data define three classes of breast tumors on the basis of whether the LOH region encompasses markers in only 7q22, only 7q31 or both 7q22 and 7q31 (Bieche *et al.*, 1992, 1997; Champeme *et al.*, 1995; Deng *et al.*, 1994; Lin *et al.*, 1996; Zenklusen *et al.*, 1994). These results suggest that 7q contains at least two tumor suppressor genes that can be inactivated in breast cancers, one at 7q22 and one at 7q31.

Three sets of tumor samples originating from different research centers have been analysed in this study. LOH of 7q22 and 7q31 have been found in tumors from each set, however, there was considerable variation in the proportion of LOH between the three sets especially for markers at 7q31. This variation is unlikely to be due to the differences in ethnic group composition of each set since the patients in all sets were almost exclusively of Caucasian origin. We think that differences in LOH frequencies can be attributed mainly to the fact that we have been very conservative in our appreciation of LOH. In particular, we did not count as LOH any sample where both alleles in the normal control were not clear. This bias almost certainly led us to underestimate LOH frequencies.

There was no correlation between LOH at 7q22 and tumor grade, suggesting that the loss of genetic material at 7q22 is probably an early event in tumor development and, most likely, is not associated with tumor progression. On the other hand, LOH of 7q22 was associated with increased tumor size, raising the possibility that the tumor suppressor gene at 7q22 could function to restrict cellular proliferation.

The smallest commonly deleted region on 7q22 includes polymorphic markers that are all located within the CUTL1 gene. Therefore, the tumor suppressor gene at 7q22 either is CUTL1 or is located close to it. Interestingly, in breast tumors that arise in MMTV-PyV LT transgenic mice, the murine Cut protein was found to form a complex together with the PyV LT antigen (Webster *et al.*, 1998). Moreover, in cotransfection studies, the CUTL1 gene product was also coimmunoprecipitated together with the SV40 Large T antigen (SV40 LT) (Martin *et al.*, manuscript in preparation). At present the effect of Large T oncoproteins on Cut function is not clear. However, these viral oncoproteins have previously been found to inactivate the function of the p53 and pRB tumor suppressor proteins. Thus, interactions with these viral oncoproteins suggest that Cut proteins may play an important role in the control of cellular proliferation. This hypothesis received further support from the finding that Cut DNA binding activity is regulated in a cell-cycle dependent manner (Coqueret *et al.*, 1998). Whereas Cut was expressed in all phases of the cell cycle, Cut DNA binding activity was the highest at the end of G1 and during S phase.

In conclusion, two sets of data, LOH mapping analysis and protein-protein interaction studies, strongly point towards the CUTL1 gene as a candidate tumor suppressor gene. It should be stressed, however, that the CUTL1 gene appears to cover a very large distance, over 200 Kbp, and that several introns are more than 10 Kbp (Zeng *et al.*, manuscript in preparation). It is thus possible that another transcription unit exists within the boundaries of the CUTL1 gene. In accordance with the recently

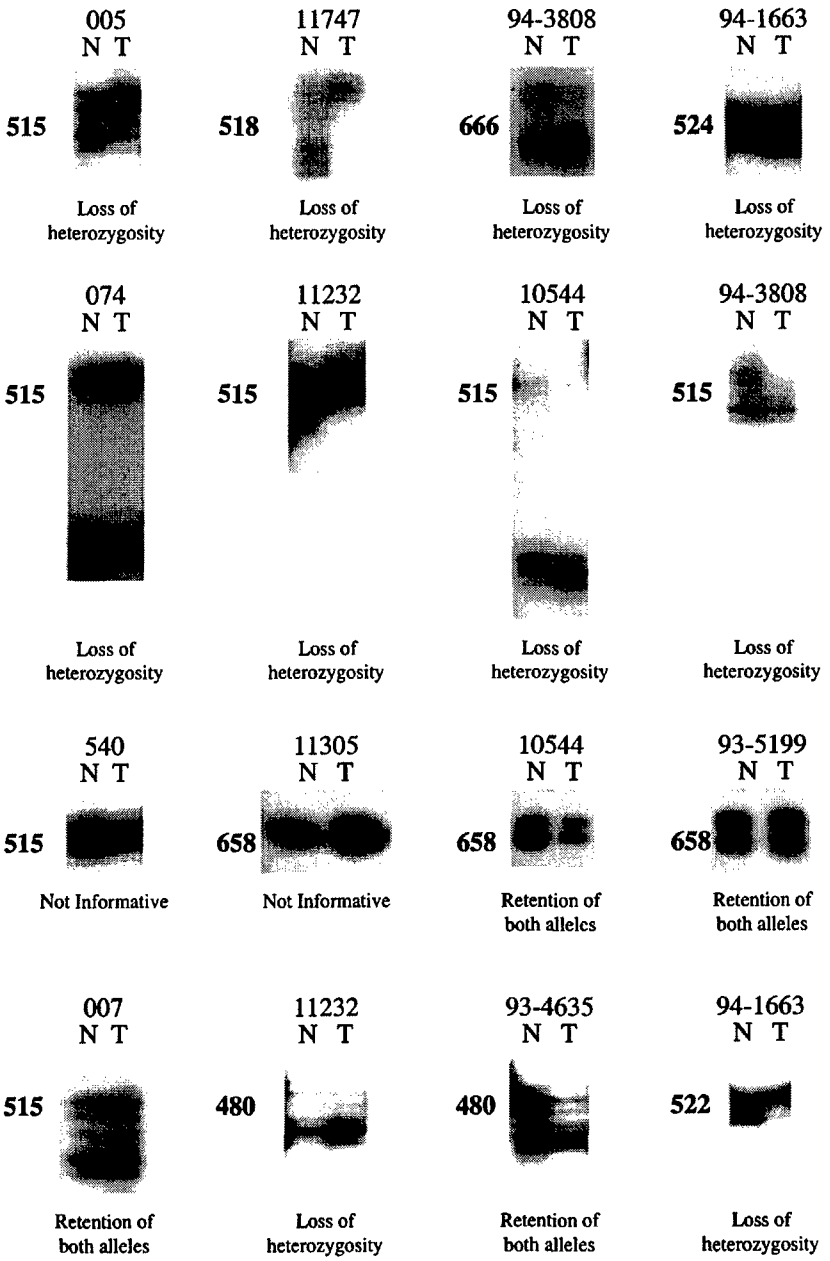


Figure 1 Representative PCR amplifications of (CA)n microsatellite repeats. DNAs from tumors (T) and matched normal breast tissues (N) were analysed from 66 patients with breast cancer. Representative PCR amplifications of (CA)n microsatellite repeats D7S518, D7S666, D7S524, D7S515, D7S522, D7S480, D7S650, D7S658 are shown. Oligonucleotide primers were used to PCR amplify the regions of DNA containing these markers, in the presence of radiolabeled dCTP. Products were denatured and separated on a standard 5% sequencing gel. A patient is considered to be informative if there are two major bands (corresponding to two alleles) in the normal DNA lane. A patient shows LOH if, in the tumor DNA lane, one of the alleles is absent or shows diminished intensity. For example, patient 10544 is informative for the marker D7S515, and shows LOH at that locus, but shows no LOH at locus D7S658 and patient 11305 is uninformative for marker D7S658

proposed rules for the definition of a tumor suppressor gene, a firm statement about the identity of the tumor suppressor gene at 7q22 will have to await the demonstration either that the remaining allele is mutated or that the function of the CUTL1 gene product is inactivated or altered in some breast tumors (Haber and Harlow, 1997). At the minimum, the fact that a polymorphic marker within CUTL1 is consistently deleted in breast tumors with LOH of 7q22 will provide a useful start point for positional cloning approaches to identify the critical tumor suppressor gene at 7q22.

Materials and methods

Specimen collection

Sixty-six pairs of samples from NCIC-Manitoba Breast Tumor Bank have been analysed in this study. All tumor samples selected from the Manitoba Breast Tumor Bank were high quality tissues selected to ensure >30% epithelial tumor component with minimal contaminating normal tissues, and matching normal tissues were also histologically verified. All tissue histological assessment was performed uniformly by a single pathologist (PW) and conducted on high quality paraffin sections from the face of the tissue block, thus ensuring consistency in assessing

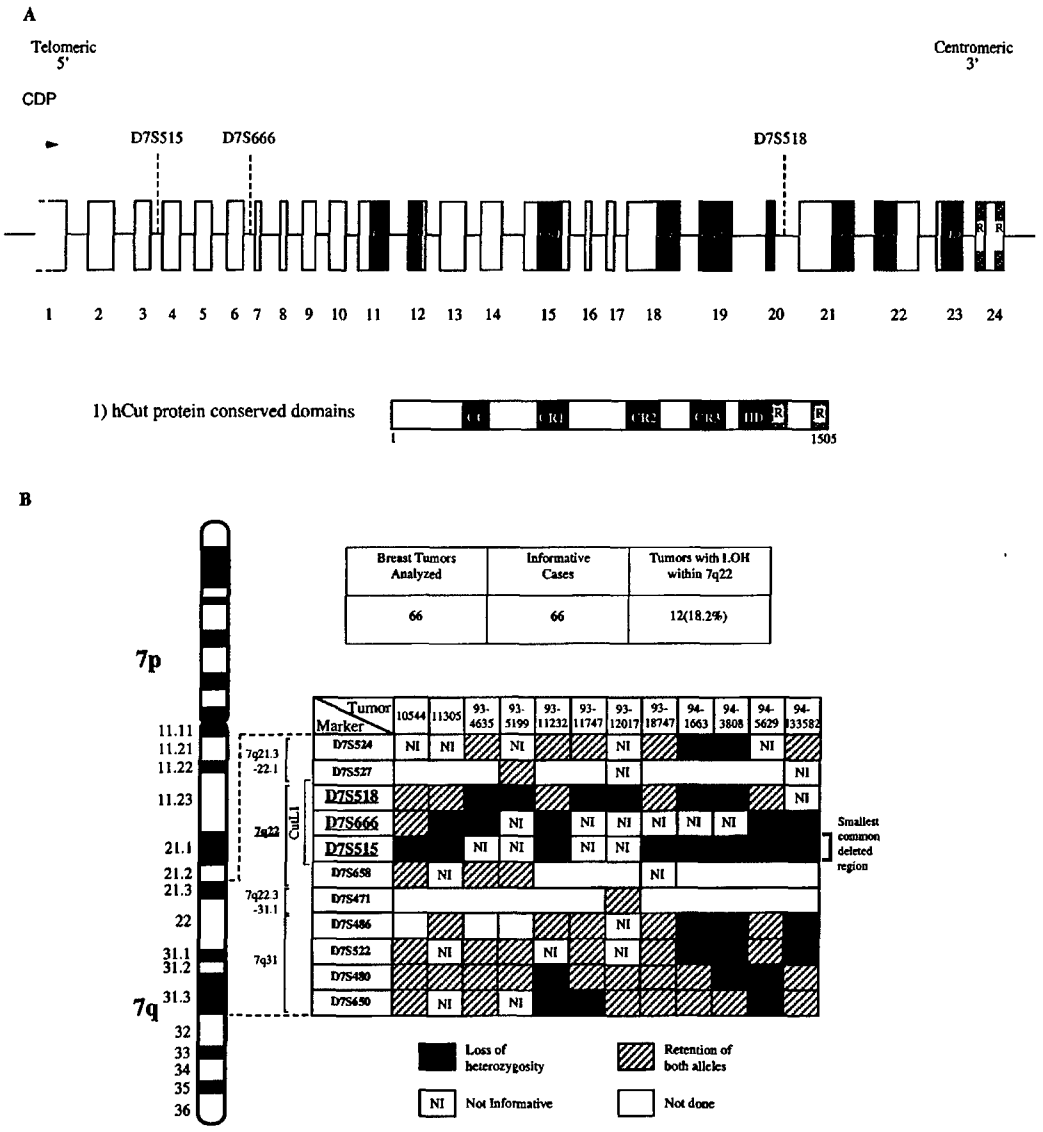


Figure 2 Mapping of LOH within 7q. (Left) Representation of human chromosome 7 including band assignments. (Right) Names of polymorphic markers used are given, along with chromosome band assignments and the position of the CUTL1 gene. Twelve columns representing 12 patients with a loss of heterozygosity on chromosome 7q22 are shown. Explanation of the symbols used is given. The smallest common deletion in these seven patients is indicated by the vertical line to the right. All 54 tumors not shown in this figure were informative for at least one marker within the critical region but failed to show LOH

Table 3 Relationship between allelic loss for loci at 7q22 and 7q31 and clinicopathological characteristics of the 66 breast tumors

		Allelic loss for loci at 7q					
		LOH 7q22			*LOH 7q31*		
Clinical-pathological characteristics		+	−	P *value*	+	−	P *value*
Type	Ductal	10	46		9	47	
	Lobular	1	4	ns	1	4	ns
	Other	sa	co, pa, is, is		sa	co, pa, is, is	
Grade	well	0	10		0	10	
	mod	6	22		7	21	
	poor	5	20	ns	3	22	ns
	not graded	1	2		1	2	
ER	positive	7	36	ns	8	36	ns
	negative	4	17		3	18	
PR	positive	5	24		4	23	
	negative	6	29	ns	7	31	ns
	not assayed	1	4		1	1	
LN status	positive	4	17		2	17	
	negative	8	32	ns	9	34	ns
	unknown	–	4		–	4	
Size	mean (cms)	4.99	2.93	0.0082	2.77	3.45	ns
Age	mean (yrs)	53	59	ns	55	59	ns

The statistical analysis was performed using the Mann-Whitney test. Grade = Nottingham grade, Sa = sarcoma, co = colloid carcinoma, pa = papillary carcinoma, is = predominant ductal carcinoma *in situ* with small ductal invasive foci, ns = not significant, ER = estrogen receptor, PR = progesterone receptor, LN = lymph node

histological features (Watson *et al.*, 1996). Samples from the Lin *et al.* (1996) study were obtained from two different sources: 43 from the Surgical Pathology Laboratory of the Royal Victoria Hospital, Montreal, Canada and 20 from the Mount Sinai Hospital, Toronto, Canada. Samples from the Bièche study were obtained from the Centre René Huguenin, Saint-Cloud, France (Bieche *et al.*, 1997).

Genomic DNA extraction from paraffin blocks

The basic protocol has been reported before (Watson *et al.*, 1993). Twenty μm sections were cut and placed in 1.5 ml Eppendorf tubes. This material was then subjected to two xylene washes and two ethanol washes (100 and 95%), incubated for 1–3 days in lysis buffer (800 μg/ml proteinase K, 0.1 M Tris (pH 8) and 0.004 M EDTA), followed by heat inactivation (boiling) and ethanol precipitation. The quantity or quality of DNA of 66 pairs of samples were good enough for the following microsatellite repeat analysis of all markers.

Microsatellite repeat analysis

DNA from tumors and matched control tissues were characterized by polymerase chain reaction (PCR) analysis using a pair of oligonucleotide primers for each polymorphic marker. Ten $(CA)_n$ repeat microsatellite markers on chromosome 7 were used to identify the region of loss of chromosome 7 in breast cancers: D7S524, D7S527, D7S515, D7S666, D7S518, D7S471, D7S486, D7S522, D7S480, D7S650. Information of the primer sequences and allele lengths are available in the Genome Database (GDB). Some of these markers, D7S527, D7S658 and D7S471, have been used only with a subset of samples in order to define more precisely the deleted region. To enhance the sensitivity of the assay and to limit the number of amplification cycles, radioactively labeled deoxyribonucleotides have been included in the reaction. The PCR amplification conditions were similar to that reported before (Zeng *et al.*, 1997). The $(CA)_n$ repeats were amplified by PCR in a final volume of 50 μl, containing 50 ng DNA, 1.5 mM $MgCl_2$, 5 μl standard 10 × PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 0.4 μM of each primer, 0.125 mM dNTPs, 1 μCi [α-^{32}P]dCTP and 1 unit of Taq polymerase (Gibco/BRL, Burlington, Ontario). An initial step of 5 min at 95°C was followed by 35 cycles of 1 min denaturation at 94°C 30 s of annealing at 59°C and 1.5 min of extension at 72°C, followed by a final extension step of 7 min at 72°C. Pilot experiments with different PCR cycle numbers indicated that 35 cycles was still within the linear amplification range. PCR products were separated in a 6% sequencing gel containing formamide or a conventional 6% sequencing gel.

Determination of LOH

Allelic loss was scored only on informative patients whose normal DNA samples were polymorphic at a given locus. Patients who were uninformative were not considered. LOH was identified, visually or following Phosphoimager densitometric analysis, as a loss in intensity (>50%) or complete loss of one allele in the tumor DNA when compared with the normal DNA from the same patient. All cases of LOH were confirmed by three separate experiments with two different reviewers. It is worth noting that we have been very conservative in our appreciation of LOH. In particular, we did not count as LOH any sample where both alleles in the normal control were not clear. This bias possibly led us to underestimate LOH frequencies. In certain cases, additional markers have been used in order to define more precisely the deleted region.

Acknowledgements

We are grateful to Drs Irene Andrulis and Ivan Bièche for allowing us to use samples collected respectively at the Mount Sinai Hospital in Toronto, Canada and the Centre René Huguenin, Saint-Cloud, France. Thanks are also due to Jo-Ann Bader for advice on deparaffinization, and Lam Leduy for the mapping of polymorphic markers and processing the tumors used in this study. This research was funded by an operating grant to AN from the US Army Medical Research and Materiel Command (USAMRMC) as part of the Breast Cancer Research Program (BCRP). AN is the recipient of a scholarship from Le Fonds de la Recherche en Santé du Québec. WRZ is the recipient of studentships from the Alexander McFee Studentship, the McGill University Department of Medicine Internal Studentship, and the Research Institute of the Royal Victoria Hospital Studentship.

References

Andres V, Nadal-Ginard B and Mahdavi V. (1992). *Dev.*, **116,** 321–334.

Berkerkarauzum S, Luleci G, Ozbilim G, Erdogan A, Kuzucu A *et al.* (1998). *Cancer Gene. Cytogen.*, **100,** 114–123.

Bièche I, Champeme MH, Matifas F, Hacene K, Callahan R *et al.* (1992). *Lancet,* **339,** 139–143.

Bièche I, Khodja A, Driouch K and Lidereau R. (1997). *Clin. Cancer Res.*, **3,** 1009–1016.

Bièche I and Lidereau R. (1995). *Genes, Chrom. Cancer,* **14,** 227–251.

Callahan R, Cropp C, Merlo GR, Diella F, Venesio T *et al.* (1993). *Clinica Chimica Acta,* **217,** 63–73.

Champeme MH, Bieche I, Beuzelin M and Lidereau R. (1995). *Genes Chrom. Cancer,* **12,** 304–306.

Coqueret O, Bérubé G and Nepveu A. (1998). *EMBO J.*, **17,** 4680–4694.

Dave BJ, Hopwood VL, King TM, Jiang H, Spitz MR *et al.* (1995). *Cancer Epidem. Biomarkers Prev.* **4,** 743–749.

Deng G, Chen LC, Schott DR, Thor A, Bhargava V *et al.* (1994). *Cancer Res.*, **54,** 499–505.

Devilee P, Hermans J, Eyfjord J, Boorresen AL, Liderau R *et al.* (1997). *Genes Chrom. Cancer,* **18,** 193–199.

Dufort D and Nepveu A. (1994). *Mol. Cell. Biol.*, **14,** 4251–4257.

Fischer K, Brown J, Scherer SW, Schramm P, Stewart J *et al.* (1998). *Recent Res. Cancer Res.*, **144,** 46–52.

Haber D and Harlow E. (1997). *Nature Genet.*, **16,** 320–322.

Harada R, Dufort D, Denis-Larose C and Nepveu A. (1994). *J. Biol. Chem.*, **269,** 2062–2067.

Harris JR, Lippman ME, Veronesi U and Willett W. (1992). *N. Engl. J. Med.*, **327,** 319–328.

Ishwad CS, Ferrell RE, Hanley K, Davare J, Meloni AM *et al.* (1997). *Genes Chrom. Cancer,* **19,** 156–160.

Isobe M, Emanuel BS, Givol D, Oren M and Croce CM. (1986). *Nature,* **320,** 84–85.

Kerangueven F, Noguchi T, Coulier F, Allione F, Wargniez V *et al.* (1997). *Cancer Res.*, **57,** 5469–5474.

Kristjansson AK, Eiriksdottir G, Ragnarsson G, Sigurdsson A, Gudmundsson J *et al.* (1997). *Anticancer Res.*, **17,** 93–98.

Larsson C, Bystrom C, Skoog L, Rotstein S and Nordenskjold M. (1990). *Genes Chrom. Cancer,* **2,** 191–197.

Lin JC, Scherer SW, Tougas L, Traverso G, Tsui LC *et al.* (1996). *Oncogene,* **13,** 2001–2008.

Lundgren R. (1991). *Scand. J. Urol. Nephrol.*, **136,** 1–35.

Lundgren R, Mandahl N, Heim S, Limon J, Henrikson H *et al.* (1992). *Genes Chrom. Cancer,* **4,** 16–24.

Mailly F, Berube G, Harada R, Mao PL, Phillips S *et al.* (1996). *Mol. Cell. Biol.*, **16,** 5346–5357.

Mann HB and Whitney DR. (1947). *Ann. Mathemat. Stat.*, **18,** 50–60.

McBride OW, Merry D and Givol D. (1986). *Proc. Natl. Acad. Sci. USA,* **83,** 130–134.

Mertens F, Johansson B, Hoglund M and Mitelman F. (1997). *Cancer Res.*, **57,** 2765–2780.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K *et al.* (1994). *Science,* **266,** 66–71.

Miller C, Mohandas T, Wolf D, Prokocimer M, Rotter V *et al.* (1986). *Nature,* **319,** 783–784.

Neufeld EJ, Skalnik DG, Lievens PM and Orkin SH. (1992). *Nature Genet.*, **1,** 50–55.

Ozïsik YY, Meloni AM,Surti U and Sandberg AA. (1993). *Cancer Genetics & Cytogenetics* **71,** 1–6.

Quaggin SE, Vandenheuvel GB, Golden K, Bodmer R and Igarashi P. (1996). *J. Biol. Chem.*, **271,** 22624–22634.

Sargent MS, Weremowicz S, Rein MS and Morton CC. (1994). *Cancer Genetics & Cytogenetics* **77,** 65–68.

Sato T, Tanigami A, Yamakawa K, Akiyama F, Kasumi F *et al.* (1990). *Cancer Res.*, **50,** 7184–7189.

Skalnik DG, Strauss EC and Orkin SH. (1991). *J. Biol. Chem.*, **266,** 16736–16744.

Solinas-Toldo S, Wallrapp C, Muller-Pillasch F, Bentz M, Gress T *et al.* (1996). *Cancer Res.*, **56,** 3803–3807.

Storto PD, Saidman SL, Demetris AJ, Letessier E, Whiteside TL *et al.* (1990). *Genes Chrom. Cancer,* **2,** 300–310.

Tougas L, Halwani F, Tremblay G, Sampalis J, Lin J *et al.* (1996). *Clin. Investig. Med. – Medecine Clinique et Experimentale,* **19,** 222–230.

Valarche I, Tissier-Seta JP, Hirsch MR, Martinez S, Goridis C *et al.* (1993). *Development,* **119,** 881–896.

Watson PH, Safneck JR, Le K, Dubik D and Shiu RP. (1993). *J. Natl. Cancer Inst.*, **85,** 902–907.

Watson PH, Snell L and Parisien M. (1996). *Cmaj,* **155,** 281–283.

Webster MA, Martinsoudant N, Nepveu A, Cardiff RD and Muller WJ. (1998). *Oncogene,* **16,** 1963–1972.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S *et al.* (1995). *Nature,* **378,** 789–792.

Xing YP, Powell WL and Morton CC. (1997). *Cancer Genetics & Cytogenetics* **98,** 69–74.

Yoon SO and Chikaraishi DM. (1994). *J. Biol. Chem.*, **269,** 18453–18462.

Zeng WR, Scherer SW, Koutsilieris M, Huizenga JJ, Filteau F *et al.* (1997). *Oncogene,* **14,** 2355–2365.

Zenklusen JC, Bieche I, Lidereau R and Conti CJ. (1994). *Proc. Natl. Acad. Sci. USA,* **91,** 12155–12158.

Appendix 2

Exon/intron structure and alternative transcripts of the CUTL1 gene

Key Words :

Human Cut homeobox gene, CCAAT displacement protein, alternative polyadenylation and splicing.

Wendy Rong Zeng[a], Erinn Soucie[b], Nam Sung Moon[a], Nathalie Martin-Soudant[a], Ginette Bérubé[a], Lam Leduy[a], and Alain Nepveu[a,c,*]

[a]Molecular Oncology Group, Departments of Medicine, Oncology, and [c]Biochemistry (A.N.) , McGill University, Montreal, Quebec, *Canada*. [b]Present address: Department of Microbiology, Toronto, Ontario, *Canada* M5G 1X8.

Abbreviations: CASP, CDP alternatively spliced product; CDP, CCAAT displacement protein; CIP1, cdk inhibitor protein 1; cpm , count per minute; CR, Cut repeat; CUTL1, Cut-like 1; HiNF-D, histone nuclear factor D; LT, large T; PyV, polyomavirus; WAF1, wild type 53 activated factor; SD1, senescence derived 1.

* To whom correspondence should be addressed, at Molecular Oncology Group, Royal Victoria Hospital, H5-08, 687 Pine Ave. West, Montreal, Quebec, Canada, H3A 1A1. Tel.: (514) 843-1479. FAX: (514) 843-1478; E-mail: alain@lan1.molonc.mcgill.ca

Abstract

The human CUTL1 gene (Cut-like 1) is a candidate tumor suppressor gene located on chromosome 7 at band 22, a region that is frequently deleted in several human cancers. The gene spans at least 340 Kbp and contains 33 exons. Synthesis of five different transcripts involves 2 promoter regions, 2 polyadenylation sites and 7 alternative splicing events. The two polyadenylation sites are located at the ends of exons 24 and 33 and are separated by approximately 40 Kbp. Transcription is initiated in two genomic regions, giving rise to alternate first exons which are spliced to a common exon 2. All transcripts contain exons 2 to 14, but differ in their 3' regions. Exon 14 can be spliced alternatively to the beginning or the middle of exon 15, or to exon 25, generating transcripts with exons 15 to 24 or exons 25 to 33. Moreover, exon 16 can be spliced out from the mature transcripts that contain exons 15 to 24. Overall, 5 distinct transcripts are generated as a result of alternative transcription initiation, splicing and polyadenylation. We discuss potential mechanisms by which alternate polyadenylation site usage may affect alternative splicing events and vice versa.

1. Introduction

The cDNAs encoding for homologues of the *Drosophila* Cut homeodomain protein have recently been isolated from several mammalian species including human, dog, mouse and rat and were respectively termed CDP/Cut (CCAAT displacement protein/human Cut), Clox (Cut-like homeobox), Cux (Cut homeobox) and CDP-2 (Andres et al., 1992; Neufeld et al., 1992; Valarche et al., 1993; Dufort and Nepveu, 1994; Yoon and Chikaraishi, 1994). More recently a second Cux gene, called Cux-2, was identified in mouse. In contrast to Cux-1 which is expressed in most tissues, Cux-2 was found to be expressed exclusively in nervous tissues (Quaggin et al., 1996). To simplify the terminology, the term "Cut" will be used in this paper to designate the protein.

Sequence homology between *Drosophila melanogaster* and mammalian Cut proteins is limited to five evolutionarily conserved domains: a region predicted to form a coiled-coil structure (CC), three regions of ~70 amino-acids, the Cut repeats (CR1, CR2 and CR3) and a Cut-type homeodomain (HD) (Blochlinger et al., 1988). The Cut repeats were shown to function as specific DNA binding domains and, at least for Cut repeat 3, to cooperate with the Cut homeodomain to bind to DNA (CR3HD) (Andres et al., 1994; Aufiero et al., 1994; Harada et al., 1994; Harada et al., 1995). In addition, two active repression domains have been identified in the carboxy-terminal domain of the protein (Mailly et al., 1996). In several vertebrate species, the gene was found to encode for a second protein, called CASP (cut alternatively spliced product), which contains the coiled-coil region but none of the DNA binding domains (Lievens et al., 1997).

Genetic studies in *Drosophila melanogaster* indicated that *cut* plays an important role in the determination and maintenance of cell-type specificity in several tissues. Mutations in *cut* cause defects in several tissues including the wings, legs, external sense organs, Malpighian tubules, tracheal system and some structures in the central nervous systems (Bodmer et al., 1987; Jack et al., 1991; Liu et al., 1991; Jack and DeLotto, 1992; Liu and Jack, 1992). In most instances, defects appear to result from the fact that some cells have enrolled in the wrong developmental program (Bodmer et al., 1987; Blochlinger et al., 1991; Liu et al., 1991; Liu and Jack, 1992). Ectopic expression of Cut proteins from either *Drosophila melanogaster*, man or mouse, was shown to have a similar effect on embryonic sensory organ development in *Drosophila melanogaster* (Ludlow et al., 1996). Moreover, the "cut wing" phenotype was rescued by either proteins

(Ludlow et al., 1996). These findings suggest that mammalian Cut proteins also play an important role in determining cell type specificity.

In mammalian cells, Cut proteins were found to play a role in cell cycle progression. Cut DNA binding activity was minimal in G0 and early G1, but raised at the end of G1 and during S phase as a result of increased expression as well as dephosphorylation by the Cdc25A phosphatase. Increased Cut activity was associated with the transcriptional repression of $p21^{WAF1/CIP1/SD1}$ (Coqueret et al., 1998). Moreover, the human Cut protein was found to be a component of the promoter complex HiNF-D, which is believed to participate in the induction of transcription of several histone genes at the G1/S phase transition of the cell cycle (Vanwijnen et al., 1996). Thus, Cut proteins appear to play at least two biological roles, one in proliferating cells and one in some differentiated cells. The molecular basis for this dual role of Cut proteins remains to be determined. In particular, it is not known whether the same gene products are expressed in different types of cells, or whether they interact with different partners. In this regard, it was suggested that Cut may alternatively interact with either pRb or p107 on different promoters (Vanwijnen et al., 1996).

Accumulated evidence indicates that CUTL1 is a candidate tumor suppressor gene. The human CDP/Cut protein was shown to repress c-Myc , a proto-oncogene that is frequently activated in several cancers (Dufort and Nepveu, 1994). Cytogenetic studies revealed that the 7q22 chromosomal region is frequently deleted in some human cancers, including leiomyomas, acute myeloid leukemia (AML), myelodysplastic syndrome (MDS). The loss of heterozygosity (LOH) at 7q22 has been reported in a subset of uterine leiomyomas and breast tumors (Ishwad et al., 1997; Zeng et al., 1997; Zeng et al., 1999). In every case the smallest deleted region included one or more of three polymorphic markers that are located within or close to the CUTL1 gene. Moreover, CUTL1 mRNA levels were reduced in eight tumors out of 13 (Zeng et al., 1999). Interestingly, in both leiomyomas and mammary tumors induced in transgenic mice expressing the Polyomavirus (PyV) Large T (LT) antigen, immunocomplexes of CUTL1 and PyV LT antigen were detected (Webster et al., 1998). The existence of such complexes suggested that alterations in the function of the Cut protein may be an important event in the etiology of some cancers.

To obtain a complete list of transcript variants and to facilitate a search for mutations within CUTL1 in human cancer, we have defined the exon-intron structure of the human CUTL1 gene and characterized its multiple transcripts. This study has revealed a very complex gene organization that allows the production of 5 alternative CUTL1 transcripts. These findings raise interesting questions regarding the regulation of alternative splicing and polyadenylation.

2. Materials and methods

2.1 Sequences of CUTL1 cDNAs and genomic DNA clones.

The cDNA sequences of HS CDP/Cut, HS CASP and MM CASP can be found at GenBank, accession numbers M74099, L12579, U68542 and U66249. The accession numbers for the CUTL1 genomic DNA clones are as follows: 46f6 (AF006752), Pac 50h2 (AF047825); cosmid 186D2 (AF024534), cosmid 123E15 (AF024533) and Pac 76h2 (AF030453). Sequence data for contigs on 7q22 can be obtained at the Washington University Human Genome Sequencing Project's World Wide Web site (http://genome.wustl.edu/gsc/gschmpg.html). The appropriate contigs are: 7/H RG339C12 (AC005096).

2.2 Isolation of genomic clones.

A human genomic placental DNA library cloned in the bacteriophage vector λDASH 1 (Stratagene) was screened with a CDP/Cut cDNA probe containing nt 18 to 4881. Positive clones were selected, purified and amplified in LE392 bacteria using standard procedures (Maniatis et al., 1992). The Lawrence Livermore National Laboratory chromosome 7-specific cosmid library (obtained from Dr. S. W. Scherer) was screened with the CDP/Cut cDNA probe to identify overlapping cosmid clones containing genomic DNA sequences not isolated from the phage library. DNA from each phage or cosmid clone was isolated and used in cycle sequencing reactions to determine intron-exon boundaries of CDP/Cut (Maniatis et al., 1992).

2.3 Plasmid construction.

The plasmid containing genomic DNA on either side of exon 1B, pCR2.1-11, was obtained in the following manner. A fragment of 637 bp was generated by PCR-amplification of human genomic DNA using as primers the oligonucleotides GGTGCGGCGCGGGAGGA and GAGCATTCTAAGCCGCGGGA, and was subcloned into the pCR2.1 vector (InVitrogen). The plasmid containing genomic DNA on either side of exon 2, Clone 57, was obtained by subcloning a 1.4 kb BamH1 fragment from pac68m21 into the corresponding site of the vector bluescribe pKS (Stratagene).

2.4 Cycle sequencing of phage and cosmid DNA.

The appropriate purified genomic DNA from phage or cosmid clones (500 ng) was used in cycle sequencing reactions with CDP/Cut-specific primers to determine the intron exon junctions and flanking intron sequences. These reactions were carried out using the GibcoBRL dsDNA Cycle Sequencing System (GibcoBRL, Burlington, Ontario) according to the instructions included in the kit. The sequences of the primers used for this purpose are available from the authors.

2.5 PCR to estimate introns size.

To estimate the sizes of the introns of the CUTL1 gene, appropriate CDP/Cut-specific oligonucleotide primers were generated and used in PCR reactions to amplify each individual intron. PCR for introns up to 4 kb in size was carried out in a final volume of 50 µl, containing 100 ng genomic DNA prepared from human placenta or the human cell lines HeLa and 293, 1.5 mM $MgCl_2$, 5 µl standard 10x PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 50 pmol of each primer, 1.25 mM dNTPs, and 2.5 units of Taq polymerase (GibcoBRL, Burlington, Ontario). An initial step of 3 minutes at 85°C was followed by 35 cycles of 40s of denaturation at 94°C, 30s of annealing at 55°C, and 40s of extension at 72°C, followed by a final extension step of 2 minutes at 72°C. Introns between 2 kb and 6 kb were amplified in conditions as described above, with the addition of 2.5 units of Taq extender PCR additive (Stratagene) and substituting the Taq extender buffer for the regular PCR buffer. Introns larger than 6 kb were amplified using the Expand Long Template PCR System (Boehringer Mannheim) according to the instructions included in the kit. PCR products were analyzed by electrophoresis of 1/10th of the reaction on a 1% agarose gel. The size of each intron was estimated by comparison of the PCR product with molecular size markers. Each intron was amplified at least twice from two different cell lines (HeLa and 293). The sizes of introns 5, 6, 7, 10, 12-19 (incl.), 21, 22 and 23 were estimated in this manner. Intron sizes obtained experimentally were also confirmed using genomic DNA sequence information (see below). The sizes of introns 1, 2, 3, 4, 8, 9, 11 and 20 were obtained from analysis of genomic DNA sequences of the appropriate portion of chromosome 7 (see Fig. 1A), available at the National Center for Biotechnology Information (NCBI) using BLAST search tools (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=0). The appropriate clones are: Pac50h2, cosmid 186D2, cosmid 123E15 and PAC 76h2.

2.6 RNA isolation.

Total RNA was isolated from various human cell lines by the method of Chomczynski and Sacchi (Chomczynski and Sacchi, 1987).

2.7 RNase mapping

The riboprobes were prepared by combining 1 µg of template DNA, transcription buffer (200 mM PIPES, 2 M NaCl, 5 mM EDTA), 10 mM DTT, RNasin (40 units), 500 mM ATP, CTP, GTP, 12 mM UTP, 50 µCi α-^{32}P-UTP and T7 RNA polymerase (69 units) (Pharmacia) and then incubating for 1 h at 37°C. After 1 h, 500 mM UTP was added and further incubated for 5 min. Next the riboprobes were treated with RNase free DNase at 37°C for 15 min and then extracted with chloropane and run through a Sephadex G50 spun column. Forty µg of total RNA was annealed to $8x10^5$ cpm of labeled riboprobe at 54°C for 16 h in 80% formamide -0.4 M NaCl -0.4 M piperazine-N,N-bis (2-ethanesulfonic acid) (PIPES) (pH 6.4)- 1 mM EDTA. RNA-RNA hybrids were digested with 30 U of RNase T2 (Gibco) per ml at 30°C for 1 h. After digestion hybrids were precipitated with 20 µg of tRNA, 295 µl 4 M guanidine thiocyanate and 590 µl of isopropanol. Pellets were resuspended in 80% formamide, 1x TBE and 0.1% XC+BPB, denatured and electrophoresed on 4% acrylamide-8M urea gel. Gels were dried and exposed to X-ray film with intensifying screens at -80°C for appropriate time.

2.8 Reverse transcriptase-PCR.

Reverse transcriptase-polymerase chain reaction (RT-PCR) was performed on Human Multiple Tissue cDNA (MTC™) of normalized, first-strand cDNA preparations derived from different adult human tissues, purchased directly from Clonetech. MTC cDNA ranges in size from 0.1 to at least 6 Kbp and is virtually free of genomic DNA (Supplier information). The cDNA preparations within each MTC Panel and between human panels have been carefully normalized to permit comparisons of relative abundance of specific transcripts within panels and between human panels. cDNA was amplified by oligos flanking the two deletions at nt 1266-1570 and nt 1938-2003 (HS CDP/Cut, GenBank, accession numbers M74099). The sequences of the oligos are: F1(nt 1186-1207): GGAGGTGCTGTTGCTGGAGAAG and B1(nt 2146-2165): AGCGGATGGGCTCATCAGAG. The amplified fragments are 981 bp for full length, 676 bp for deletion of nt 1266-1570 (deletion 5' of CR1), 915 bp for deletion of 1938-2003, 610 bp for deletion of both. The plasmid PMX/Cut containing full length CDP/Cut served as positive control, and the negative control had no DNA template. The specificity of the PCR product was confirmed by BamH1 digestion to produce expected sizes of fragments as well as Southern blot hybridization (not shown). PCR was performed in a final volume of 50 µl, containing 1 ng cDNA, 1.5 mM $MgCl_2$, 5 µl

standard 10x PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 0.45 μM of each primer, 0.12 mM dNTPs, 5% DMSO, 1 μCi [α-^{32}P] dCTP, and 1 units of Taq polymerase (Gibco BRL, Burlington, Ontario). An initial step of 4 minutes at 95°C was followed by 30 cycles of 1 of denaturation at 95°C, 30s of annealing at 59°C, and 2 minutes of extension at 72°C, followed by a final extension step of 7 minutes at 72°C. Pilot tests have been done to make sure the PCR reaction does not reach its plateau (not shown). PCR products were separated on a 4% non-denaturing polyacrylamide gel and visualized by autoradiography.

3. Results

3.1 CUTL1 exon/intron structure

One of our long term goals is to determine, in tumors with CUTL1 LOH, whether the remaining allele is mutated. For most tumors it will not be possible to obtain RNA that would serve to clone CUTL1 cDNAs. Thus, the search for mutation will have to be performed with genomic DNA. For this purpose, it is necessary first to determine the CUTL1 exon/intron structure, including the sequence at the 5' and 3' end of introns. The sequence information obtained in this manner will then serve to design oligonucleotides that can be used as primers to PCR-amplify individual exons as well as some of the flanking intronic sequences.

We have now determined the complete exon/intron structure of the CUTL1 gene (Fig. 1A). Starting from a panel of cosmid clones encompassing the entirety of the gene, we performed DNA sequencing using as primers oligonucleotides derived from the cDNA sequence. In addition, DNA sequencing of the long arm of chromosome 7, as part of the human genome project, has generated a lot of sequence information (see Fig. 1A, genomic DNA clones). The chromosome 7 DNA sequence was analyzed periodically to identify stretches that would contain CUTL1 cDNA sequence. In this manner, we have confirmed some of the results obtained through our own DNA sequencing project and found some exon/intron junctions not yet identified. The exon structure of the full length CDP/Cut mRNA is shown in Fig. 1B, and is aligned with the functional domains of the protein (fig. 1C). DNA sequence at the exon/intron junction is presented in Table 1. Alternative splicing events were also detected by the cloning of various cDNAs (Fig. 1D) and were confirmed by reverse-transcriptase polymerase chain amplification (RT-PCR) and RNase mapping analyses (Fig. 3 and 4; see section 3.5 and 3.6). The structure of the alternatively spliced mRNAs of the CUTL1 gene is shown diagrammatically (Fig. 1D). All transcripts include exons 2 to 14, but differ 5' and 3' of this common core. There are two alternative exons 1, 1A and 1B, and two alternative last exons, 24 and 33. Splicing downstream of intron 14 can take place at three different positions: at the start of exon 15 (nt 1266 in HS CDP), within exon 15 (nt 1570 in HS CDP) and at the start of exon 25 (nt 1275 in HS CASP). Moreover, exon 16 can be spliced out in those transcripts that contain exon 15. Thus, transcripts that end with exon 33 contain, downstream of exon 14, exons 25 to 33, whereas transcripts that end with exon 24 contain exons 15 to 24 (with the exception of exon 16 which may or may not be included). As a result of these alternative initiation, splicing and polyadenylation events, up to 5

distinct transcripts can be expressed (Fig. 1D). Which process, splicing or polyadenylation, dictates the course of events will be discussed in section 4.3.

3.2 Mapping of polymorphic markers relative to CUTL1

We had previously identified three polymorphic markers within or close to the CUTL1 gene: D7S515, D7S666 and D7S518 (Zeng et al., 1997). The cloning of CUTL1 genomic DNA allowed us to precisely map these markers by Southern blot analysis (data not shown). The positions of these three markers was confirmed by sequence analysis of genomic clones PAC50h2 and cos 123E15 (see Fig. 1A). D7S515 is located within intron 3, D7S666 in intron 6 and D7S518 in intron 20. Thus, all three markers are positioned within the CUTL1 gene.

3.3 Mapping of CUTL1 transcription start sites

A genomic DNA fragment corresponding to the 5' end of the CUTL1 cDNA was subcloned into the pCR2.1 vector. No TATA box was found within approximately 1 Kbp of DNA sequence. A riboprobe was made and used in RNase mapping analysis to map the 5' end of CDP/Cut transcripts and, by inference, the transcription start site(s) of the gene (Fig. 2A). Several protected fragments were observed, suggesting that transcription starts at various positions, a finding that is consistent with the absence of a TATA box in this region. The positions of transcription start sites are shown on the sequence (Fig. 2B). Importantly, many mRNAs appear to be initiated upstream of the position predicted from the published CDP/Cut cDNA sequence (Fig. 2B; CDP/Cut cDNA: GenBank accession No. M74099).

3.4 Mapping of potential CUTL1 translation start sites

The translation start site in the CDP/Cut cDNA sequence has been arbitrarily positioned at nt 44 where the first methionine codon is found (position 550 in the sequence of Fig. 2B). However, there is no termination codon preceding it in the same reading frame. Interestingly, when the upstream genomic sequence is abutted to the known sequence of CDP/Cut, the open reading can be extended towards the 5' end, and additional methionine codons can be found (positions 250 and 337, Fig. 2B). This finding raised the possibility that mRNAs which are longer at their 5' end could encode proteins with extended amino-terminal domains. To investigate this, we have carefully mapped the multiple transcription start sites and positioned them on the sequence shown in Fig. 2B. We found that most mRNAs were initiated downstream of the ATG codon at position 337, and therefore did not include those upstream ATG codons. However, one major transcript was initiated at position 334 in HeLa and 293 cells. It is thus possible that this mRNA encodes for a protein whose translation would start at nt 337. Importantly, in RNase mapping no protected fragment was observed at the position expected for a transcript that would initiate upstream of the NcoI site. Thus, we conclude that in no case can translation start at this NcoI site, and for the vast majority of transcripts translation must initiate at the ATG codon previously identified (position 550 in Fig. 2B; nt 44 in HS CDP sequence, M74099).

3.5 Ubiquitous expression of the alternatively spliced CDP/Cut transcripts

We have previously obtained multiples CDP/Cut cDNAs with internal deletions at either of two positions. Some cDNAs lacked sequences from nt 1266 to 1569, at the beginning of exon 15. Others missed the entire exon 16, from nt 1938 to 2105. To confirm whether alternatively spliced transcripts are expressed and to verify whether expression of these transcripts follows a tissue-specific pattern, we performed RT-PCR analysis using total RNAs from various human tissues. Two oligonucleotides, located in exons 14 and 18 respectively, were designed to amplify the region of alternative splicing (Fig. 3). Four products, corresponding to all possible isoforms, were observed in most tissues. The full length isoform was consistently the most abundant species, however, collectively the variant species represented a substantial fraction of all transcripts. The smallest isoform, lacking nt 1266-1569 and 1938-2105, was the least abundant but could be visualized in all tissues following a longer exposure (data not shown). Overall, the relative abundance of all isoforms did not vary much from one tissue from another. Similar results were obtained using a panel of cell lines from diverse origins. We conclude that all isoforms are expressed in all tissues.

3.6 CDP/Cut and CASP transcripts diverge downstream of exon 14

As shown in Fig. 1D, splicing downstream of exon 14 can involve three alternative acceptor sites one of which, generating the CASP transcript, is located approximately 76 kb downstream (Lievens et al., 1997). To confirm the existence of this transcript, and to compare its expression relative to the others, we performed RNase mapping analysis using riboprobes derived from the human CASP cDNA. The advantage of a CASP riboprobe in this experiment is that, in addition to the CASP-specific protected fragment, only one additional fragment should be generated by all CDP/Cut mRNA isoforms. As expected, using the human CASP-specific riboprobe, two protected fragments were detected in human placenta (Fig. 4). The ratio between these two fragments did not vary in diverse human cell lines (data not shown).

4. Discussion

4.1 Polymorphic markers within CUTL1

Analysis of the organization of CUTL1 indicates that it spans a minimum of 340 Kbp and comprises 33 exons ranging in size from 47 to 1445 bp. Three polymorphic markers previously found to be in close proximity to the CUTL1 gene, D7S515, D7S666 and D7S518, were mapped to introns 3, 6 and 20 respectively. Some of these markers have been shown in other studies to be included in the smallest commonly deleted region in some uterine leiomyomas and breast tumors (Zeng et al., 1997; Zeng et al., 1999). Interestingly, among 12 breast tumors with loss-of-heterozygosity (LOH) of 7q22, 5 tumors exhibited LOH of D7S515 while retaining D7S518 (Zeng et al., 1999). In one these tumors, the marker D7S666 within intron 6 was also retained. Thus, the centromere-proximal border of the smallest commonly deleted region can be precisely mapped to the region in between introns 6 and 20, further implicating the CUTL1 as a candidate tumor suppressor gene. It should be stressed, however, that the large size of some introns within CUTL1 makes it possible that they contain other genes. Indeed, analysis of their DNA sequence reveal several open reading frames (Glockner et al., 1998).

4.2 Transcription and translation start sites

We have cloned the genomic fragment containing exon 1B and generated a riboprobe to perform RNase mapping with total RNA from various sources. Multiple protected fragments were observed. Two conclusions can be drawn from this result. First, exon 1B is truly the first exon for some of the CDP/Cut mRNAs. If exon 1B was in fact the second exon, we would have observed only 1 protected fragment whose boundary would correspond to the junction between intron 1 and exon 2. Secondly, transcription upstream of exon 1B is initiated at multiple sites over a region of approximately 200 bp. This findings is consistent with the fact that this region does not include a TATA box and is GC-rich.

The sequence in the 5' region of the CDP/Cut cDNA does not include an in-frame termination codon upstream of the putative ATG initiation codon. As a result, the open reading frame can be extended to the 5' end of the cDNA. This raised the possibility that the cDNA was not full length and that another translation initiation site might exist further upstream. Indeed, analysis of the genomic DNA sequence at the 5' end of the gene indicated that not only the open reading frame could be extended in the 5' direction, but also two additional in frame ATG codons were present. It should be stressed that the riboprobe used in the RNase protection assay was designed in such a way as to easily detect any transcript that would protect the entire genomic sequence of the probe (see Fig. 2A). Since no such transcript was detected, we can be fully confident that no mRNA is initiated upstream of this region. Moreover, the vast majority of CDP/Cut mRNAs were found to start downstream of these additional ATG codons. In light of these results, we are compelled to conclude that the ATG codon at position 44 of the CDP/Cut cDNA sequence (nt 550 in Fig. 2B) is most likely the correct translation initiation site.

4.3 Alternative transcription initiation, splicing and polyadenylation

In the course of this study, several cDNA clones were isolated whose DNA sequence suggested the existence of alternative splicing events. The production of 5 alternatively spliced mRNAs was confirmed by RT-PCR analysis (Fig. 3) and RNase mapping analysis (Fig. 4). Interestingly, most of the alternative splicing events are concentrated in the same region of the gene, downstream of exons 14 and 15. Splicing downstream of exon 15 can lead to linkage with exon 16 or 17, while splicing downstream of exon 14 can involve 3 different acceptor sites, two at the beginning or in the middle of exon 15, and one at the 5' end of exon 25, approximately 76 Kb downstream. This last event is responsible for the production of CASP mRNAs which lack exons 15 to 24 but include exons 25 to 33. In contrast, mRNAs encoding for CDP/Cut proteins are spliced from exon 14 to 15 and are cleaved at the end of exon 24. At first sight, it appears that there might be competition between several splice acceptor sites downstream of exon 14. However, we favor the view that the primary decision entails the 3' end cleavage and polyadenylation at the end of exon 24. Indeed, since exon 25 is located approximately 24 Kb downstream of exon 24, there should be enough of time to cleave the primary transcript at the end of exon 24 while RNA pol II is progressing through that region. Cleavage at this position, in turn, would signal the RNA polymerase to terminate transcription before it has reached exon 25. This would effectively preclude the possibility of exon 14 being spliced to exon 25. In contrast, synthesis of the CASP mRNA necessitates that the primary transcript is not cleaved downstream of exon 24 and that transcription proceeds unabated up to and beyond exon 25. In this respect, we note that the polyadenylation signal at the end of exon 24 is probably inefficient because the AAUAAA motif is embedded within the sequence AAAAUAAAA. The presence of an excess of A residues is reminiscent of the polyadenylation signal at the end of the late transcription unit of Polyomavirus. It was shown that

processing of primary transcripts downstream of such signals is inefficient (Lanoix and Acheson, 1988; Levitt et al., 1989; Sheets et al., 1990). As a result, transcription was found to continue several times around the circular viral genome, generating primary transcripts with multiple genome equivalents (Acheson, 1978; Acheson, 1984). Within the CUTL1 locus, primary transcripts that are elongated up to exon 33 may invariably be spliced between exon 14 and 25, but the possibility of cleavage downstream of exon 24 would still exist. It ensues that the decision to make CDP/Cut or CASP mRNA is based on the competition between cleavage at the end of exon 24 and splicing between exon 14 and 25 (resulting in the skipping of exon 24). This type of post-transcriptional regulation has previously been described, notably in the case of the gene encoding for calcitonin and calcitonin gene-related peptide ((Amara et al., 1984; Leff et al., 1987); reviewed in (Edwalds-Gilbert et al., 1997)).

Acknowledgments

This research was funded by an operating grant to A.N. from the U.S. Army Medical Research and Materiel Command (USAMRMC) as part of the Breast Cancer Research Program (BCRP). A.N. is the recipient of a scholarship from Le Fonds de la Recherche en Santé du Québec. W.R.Z. is the recipient of studentships from the Alexander McFee Studentship, the McGill University Department of Medicine Internal Studentship, and the Research Institute of the Royal Victoria Hospital Studentship.

References

Acheson, N.H.: Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proceedings of the National Academy of Sciences of the United States of America 75 (1978) 4754-8.

Acheson, N.H.: Kinetics and efficiency of polyadenylation of late polyomavirus nuclear RNA: generation of oligomeric polyadenylated RNAs and their processing into mRNA. Molecular & Cellular Biology 4 (1984) 722-9.

Amara, S.G., Evans, R.M. and Rosenfeld, M.G.: Calcitonin/calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. Molecular & Cellular Biology 4 (1984) 2151-60.

Andres, V., Chiara, M.D. and Mahdavi, V.: A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins. Genes & Development 8 (1994) 245-57.

Andres, V., Nadal-Ginard, B. and Mahdavi, V.: Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding regulatory proteins differentially expressed during development. Development 116 (1992) 321-34.

Aufiero, B., Neufeld, E.J. and Orkin, S.H.: Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein. Proc. Natl. Acad. Sci. USA 91 (1994) 7757-7761.

Blochlinger, K., Bodmer, R., Jack, J., Jan, L.Y. and Jan, Y.N.: Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila. Nature 333 (1988) 629-35.

Blochlinger, K., Jan, L.Y. and Jan, Y.N.: Transformation of sensory organ identity by ectopic expression of Cut in Drosophila. Genes & Development 5 (1991) 1124-35.

Bodmer, R., Barbel, S., Shepherd, S., Jack, J.W., Jan, L.Y. and Jan, Y.N.: Transformation of sensory organs by mutations of the cut locus of D. melanogaster. Cell 51 (1987) 293-307.

Chomczynski, P. and Sacchi, N.: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analyt Biochem 162 (1987) 156-159.

Coqueret, O., Berube, G. and Nepveu, A.: The Mammalian Cut Homeodomain Protein Functions As a Cell-Cycle-Dependent Transcriptional Repressor Which Downmodulates P21(Waf1/Cip1/Sdi1) In S Phase. EMBO Journal 17 (1998) 4680-4694.

Dufort, D. and Nepveu, A.: The human cut homeodomain protein represses transcription from the c-myc promoter. Molecular & Cellular Biology 14 (1994) 4251-7.

Edwalds-Gilbert, G., Veraldi, K.L. and Milcarek, C.: Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Research 25 (1997) 2547-61.

Glockner, G., Scherer, S., Schattevoy, R., Boright, A., Weber, J., Tsui, L.C. and Rosenthal, A.: Large-scale sequencing of two regions in human chromosome 7q22: analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci reveals 17 genes. Genome Res 8 (1998) 1060-73.

Grant, C.E., Kurz, E.U., Cole, S.P. and Deeley, R.G.: Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA. Genomics 45 (1997) 368-78.

Harada, R., Berube, G., Tamplin, O.J., Denis-Larose, C. and Nepveu, A.: DNA-binding specificity of the cut repeats from the human cut-like protein. Molecular & Cellular Biology 15 (1995) 129-40.

Harada, R., Dufort, D., Denis-Larose, C. and Nepveu, A.: Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains. Journal of Biological Chemistry 269 (1994) 2062-7.

Ishwad, C.S., Ferrell, R.E., Hanley, K., Davare, J., Meloni, A.M., Sandberg, A.A. and Surti, U.: Two discrete regions of deletion at 7q in uterine leiomyomas. Genes, Chromosomes & Cancer 19 (1997) 156-60.

Jack, J. and DeLotto, Y.: Effect of wing scalloping mutations on cut expression and sense organ differentiation in the Drosophila wing margin. Genetics 131 (1992) 353-63.

Jack, J., Dorsett, D., Delotto, Y. and Liu, S.: Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development 113 (1991) 735-47.

Lanoix, J. and Acheson, N.H.: A rabbit beta-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. EMBO Journal 7 (1988) 2515-22.

Leff, S.E., Evans, R.M. and Rosenfeld, M.G.: Splice commitment dictates neuron-specific alternative RNA processing in calcitonin/CGRP gene expression. Cell 48 (1987) 517-24.

Levitt, N., Briggs, D., Gil, A. and Proudfoot, N.J.: Definition of an efficient synthetic poly(A) site. Genes & Development 3 (1989) 1019-25.

Lievens, P.M., Tufarelli, C., Donady, J.J., Stagg, A. and Neufeld, E.J.: CASP, a novel, highly conserved alternative-splicing product of the CDP/cut/cux gene, lacks cut-repeat and homeo DNA-binding domains, and interacts with full-length CDP in vitro. Gene 197 (1997) 73-81.

Liu, S. and Jack, J.: Regulatory interactions and role in cell type specification of the Malpighian tubules by the cut, Kruppel, and caudal genes of Drosophila. Developmental Biology 150 (1992) 133-43.

Liu, S., McLeod, E. and Jack, J.: Four distinct regulatory regions of the cut locus and their effect on cell type specification in Drosophila. Genetics 127 (1991) 151-9.

Ludlow, C., Choy, R. and Blochlinger, K.: Functional Analysis Of Drosophila and Mammalian Cut Proteins In Flies. Developmental Biology 178 (1996) 149-159.

Mailly, F., Berube, G., Harada, R., Mao, P.L., Phillips, S. and Nepveu, A.: The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy. Molecular & Cellular Biology 16 (1996) 5346-57.

Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbour, New York, 1992.

Neufeld, E.J., Skalnik, D.G., Lievens, P.M. and Orkin, S.H.: Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut. Nature Genetics 1 (1992) 50-5.

Quaggin, S.E., Vandenheuvel, G.B., Golden, K., Bodmer, R. and Igarashi, P.: Primary Structure, Neural-Specific Expression, and Chromosomal Localization Of Cux-2, a Second Murine Homeobox Gene Related to Drosophila Cut. Journal of Biological Chemistry 271 (1996) 22624-22634.

Sheets, M.D., Ogg, S.C. and Wickens, M.P.: Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucleic Acids Research 18 (1990) 5799-805.

Valarche, I., Tissier-Seta, J.P., Hirsch, M.R., Martinez, S., Goridis, C. and Brunet, J.F.: The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. Development 119 (1993) 881-96.

Vanwijnen, A.J., Vangurp, M.F., Deridder, M.C., Tufarelli, C., Last, T.J., Birnbaum, M., Vaughan, P.S., Giordano, A., Krek, W., Neufeld, E.J., Stein, J.L. and Stein, G.S.: Cdp/Cut Is the Dna-Binding Subunit Of Histone Gene Transcription Factor Hinf-D - a Mechanism For Gene Regulation At the G(1)/S Phase Cell Cycle Transition Point Independent Of Transcription Factor E2f. Proc. Natl. Acad. Sci. USA 93 (1996) 11516-11521.

Webster, M.A., Martinsoudant, N., Nepveu, A., Cardiff, R.D. and Muller, W.J.: The Induction Of Uterine Leiomyomas and Mammary Tumors In Transgenic Mice Expressing Polyomavirus (Pyv) Large T (Lt) Antigen Is Associated With the Ability Of Pyv Lt Antigen to Form Specific Complexes With Retinoblastoma and Cutl1 Family Members. Oncogene 16 (1998) 1963-1972.

Yoon, S.O. and Chikaraishi, D.M.: Isolation of two E-box binding factors that interact with the rat tyrosine hydroxylase enhancer. Journal of Biological Chemistry 269 (1994) 18453-62.

Zeng, W.R., Scherer, S.W., Koutsilieris, M., Huizenga, J.J., Filteau, F., Tsui, L.C. and Nepveu, A.: Loss Of Heterozygosity and Reduced Expression Of the Cutl1 Gene In Uterine Leiomyomas. Oncogene 14 (1997) 2355-2365.

Zeng, W.R., Watson, P., Lin, J., Jothy, S., Lidereau, R., Park, M. and Nepveu, A.: Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene. Oncogene 18 (1999) 2015-2021.

Figure Legends

Fig. 1 Structure of the CUTL1 gene, mRNAs and proteins

A The exon/intron structure of the CUTL1 gene is presented. Vertical lines represent individual exons. Above the map and in boxes are shown three polymorphic markers previously used in LOH studies. Under the map are shown various genomic clones that were used to determine the exon/intron structure of the gene. The distances between exons 1A (CASP-specific) and 1B (CDP/Cut-specific), and between 1B and 2, are not yet known. Similarly, the sizes of introns 2 and 8 are not known, but must be at least 84 Kb and 38 Kb, respectively.

B The full length CDP/Cut mRNA is made of 24 exons. Shown in gray are two regions that are included or not following alternative splicing events. Note that exon 15 has two splice acceptor sites. The exons can be aligned with the evolutionarily conserved domains of the protein, as shown in C.

C The full length CDP/Cut protein. Black rectangles represent the five evolutionarily conserved regions of the proteins: coiled-coil (CC), cut repeats 1 (CR1), 2 (CR2) and 3 (CR3) and the homeodomain (HD). Shown as hatched boxes are two active repression domains (R) as previously defined (Mailly et al., 1996).

D Alternative splicing events and their products. Shown as dashed lines are 7 alternative splicing events. Note that splicing event #1 is always followed by splicing event #5, while #2 is linked to #3 or #4. Three and two alternative splice acceptor sites can be used downstream of exon 14 (3, 4 and 5) and 15 (6 and 7), respectively. The five alternatively spliced mRNAs are shown below.

Fig. 2 RNase mapping analysis of CUTL1 mRNA 5' ends

A A genomic fragment corresponding to the 5' end of the CDP/Cut cDNA was subcloned into the pCR2.1 vector (InVitrogen). A riboprobe was prepared and used in RNase mapping analysis using 30 µg of total RNA from human placenta as well as Hel, 293 and HeLa cell lines. The same amount of tRNA was used as a control. The undigested riboprobe as well as protected fragments are indicated by arrows. The sizes of protected fragments are approximate (+/- 5 nt).

B Genomic DNA sequence at the 5' end of exon 1B of CDP/Cut. The published sequence of CDP/Cut starts at nt 513, and exon 1B sequence ends at nt 579. The amino acid sequence is shown under the nucleotide sequence. In frame with the long open reading frame, are three ATG codons at positions 250, 337 and 550. The region that is underlined corresponds to the suspected amino-terminus of CDP/Cut. The approximate positions of transcripts 5' ends are indicated by the arrows. Note that only one transcript, present in low abundance in 293, would start upstream of the second ATG codon at 337.

Fig. 3 Expression of CDP/Cut alternative mRNAs in various tissues

CDP/Cut alternative mRNAs were analyzed by reverse-transcriptase polymerase chain amplification (RT-PCR) using Human Multiple Tissue cDNA (MTC™) of normalized, first-strand cDNA preparations derived from different adult human tissues (Clonetech). Using as primers oligonucleotides corresponding to nt 1186-1207 and nt 2146-2165, the amplified fragments are 981 bp for full length, 676 bp for deletion of nt 1266-1570 (deletion 5' of CR1), 915 bp for deletion of 1938-2003, 610 bp for deletion of both.

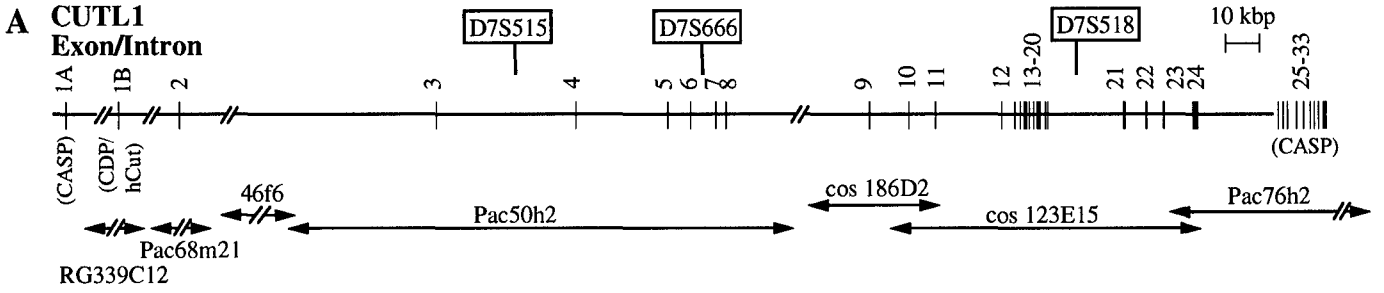
Fig. 4. RNase mapping analysis of CASP and CDP/Cut in human placenta.

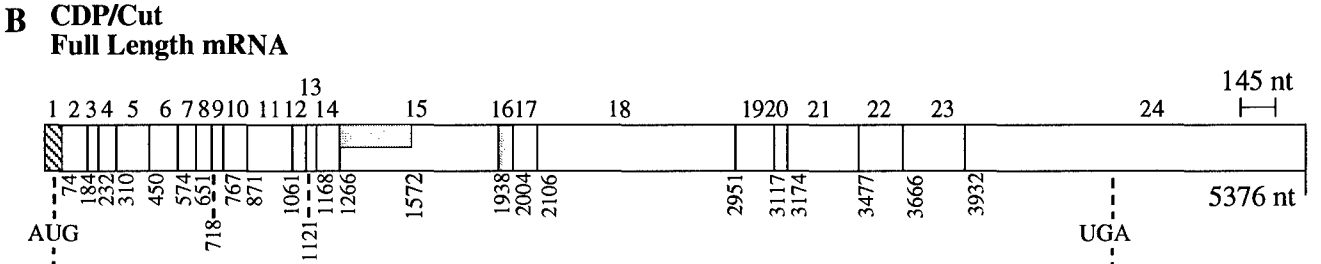
(Left) Total RNA (40 µg) from human placenta was submitted to RNase mapping analysis using a riboprobe derived from the CASP cDNA and encompassing sequences shared by the CDP/Cut and CASP mRNAs. (Right) Schematic representation of the human CDP/Cut and CASP cDNAs are displayed from exon 1 to 15 and exon 1 to 25, respectively. The probes and protected fragments are depicted below. The human CASP riboprobe spanned nucleotides 1020 to 1316 of the HSCASP sequence. Note that the fragment protected by the CASP mRNA is larger, since the riboprobe is derived from its cDNA.

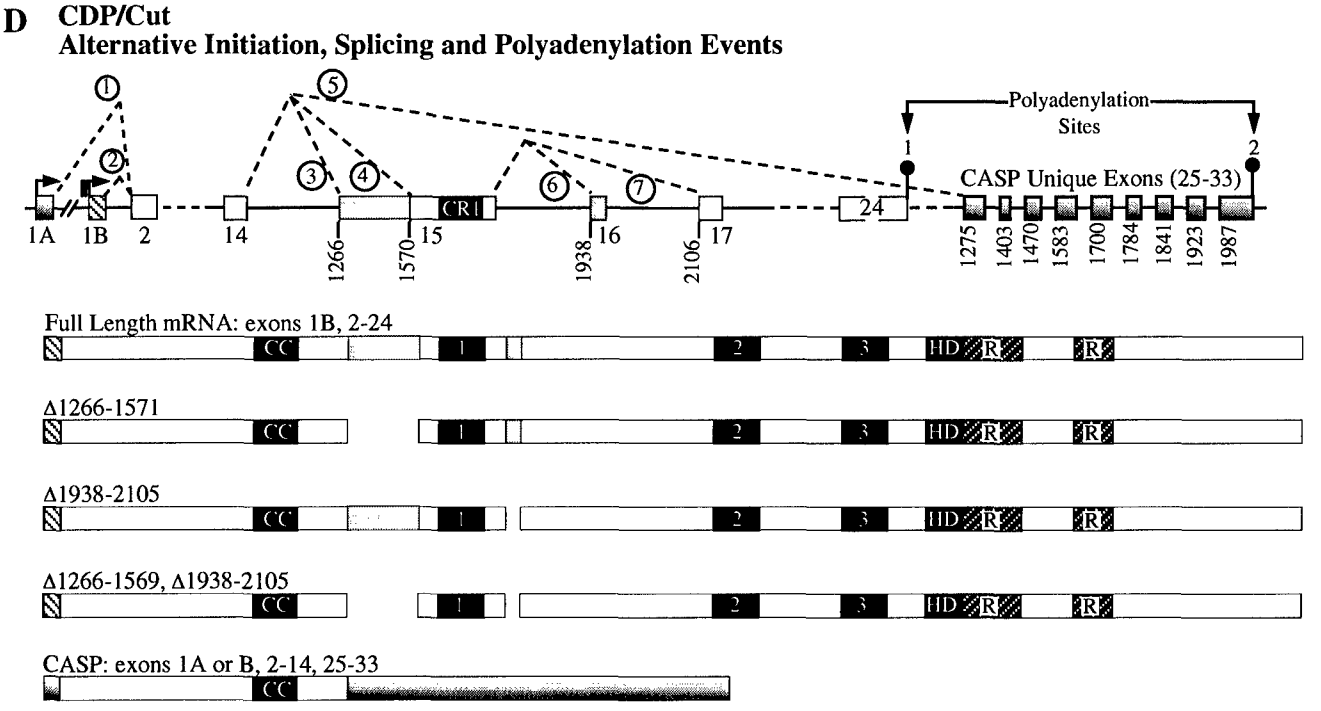
Table 1 Nucleotide Sequence of Exon/Intron Junctions of the CUTL1 Gene

CUTL1 exon #	5' end	3' end	5'end sequence of exon	3'end sequence of exon	Intron size (bp)
1B	1	73	multiple 5' ends	CAGGTTGAAG*gt*gagcggcgtgtgg	
2	74	184	*ttccccaacag*AGAGAACTCG	CACTCCAGAG*gt*gaggcgcgt	>84,334
3	185	232	*cgcctcctgctccag*GATTTGCGCA	CCAAGGAGAG*gt*aagcttttctatt	41,309
4	233	311	*ttccttccctttcag*ATTGATGCAC	GACGTCCCAG*gt*aagccccggcagt	27,029
5	312	449	*tttcctgttgtgcag*ATCCCGTACC	AAAAATCAAG*gt*tggtggaaaatgc	6,707
6	450	573	*ttttcttttctgcag*AGGTTACGAT	AAAAGGAGAG*gt*gagcatgacttcc	7,371
7	574	650	*cctcttcctttgcag*AAAGCTGCAG	CTACAAACAG*gt*tttgatactctcc	3,430
8	651	717	*ttctttgcctttcag*CCCTGGAAAA	CTACTGCAAA*gt*aagtctctctgct	>38,790
9	718	766	*ctctcaccctcctag*GGCCGACGAG	GGCAAACCAG*gt*aggaccctggacg	11,610
10	767	871	*tttcatttccttcag*AGGGCAGAGG	ACCAGACGTG*gt*gggtagccccggc	7,971
11	872	1060	*ttttctcctccccag*GAGCAGGCCA	CACACTCAAA*gt*aaggggctgcgg	11,155
12	1061	1119	*ttcttctcttttcag*CAACTGGAAG	AAGAGCTGAA*gt*aagtacggagagc	3,969
13	1120	1168	*gttgtgctcttgcag*CATTCTGAAG	TGGGACACAG*gt*acgtgtctcacct	1,614
14	1169	1265	*ctctgccccttctag*GATGCGGCCA	GACCTGAGCG*gt*aggttggccgggc	1,051
15	1266	1937	*tgcccttccttgtag*GGTCAGCCAG	AGACAAAGAG*gt*gagagactggcgt	1,495
16	1938	2003	*cttgtttttcttcag*AGAATCCAGG	GGGTCTGAAG*gt*atgttgcaggcag	1,202
17	2004	2105	*aacttctccccacag*GTAACATCAC	CAGAAAACTG*gt*acagcttccattt	1,186
18	2106	2950	*tgtttctccatgcag*CAGAGCCGGC	CGGGGAGAAG*gt*aagggatctgctc	2,185
19	2951	3116	*tcggtgccactccag*GTGCTGGGCC	CAAGGGCCAG*gt*aatgggggtcctg	556
20	3117	3173	*ctttttctactttag*TCCTCCACTC	GTGAGCTCAG*gt*aagcagcagtttc	22,184
21	3174	3476	*gtgctttaattacag*AAAGCACTCC	AACAACCTCG*gt*aggttctcctctc	6,374
22	3477	3665	*ctgttttctctctag*GCCAGCGCTT	GAGAAGAAAG*gt*aagtctccctgcc	5,077
23	3666	3930	*tctcttcctccgcag*CCTACATGAA	ACAACTACAG*gt*acgacggctggca	8,826
24	3931	5376	*cttgttgtcttgtag*GTCTCGGATC	TTACTCCACA*tattttttaacaaaa*	24,006
25	1275	1402	*ttcctttggccacag*GACGCTGTGC	CGATGCCGAG*gt*gagcccacccccc	752
26	1403	1469	*tcctgtcacctgcag*GGTGCCGCTG	CTATTCTACG*gt*aaggagaggcctg	937
27	1470	1582	*tcctctacctgctag*GACCTGCAGC	GCTTGAGGCC*gt*gagtcacatcctt	2,591
28	1583	1699	*cccctcccccccag*GAGAACCGCC	CCCTGGCCGG*gt*gagggccctcccc	2,061
29	1700	1783	*ctcctgtctgtgcag*GGCAGCCGGC	CAGCAAGCGG*gt*tcgtgagcccagc	682
30	1784	1840	*tccctccctgtgcag*GAGCGGCAGA	CCTCAGCATG*gt*gagtccctgcccc	970
31	1841	1922	*cctccttgctcccag*GGGCGTCTGG	GGTCTTCCTG*gt*gagtgtgcacacg	792
32	1923	1986	*tctccacctggccag*GTGCTCTACA	GCGCCAAGAA*gt*gaggacccccact	245
33	1987	2855	*tcccgtgtcccccag*GTTCGCTGAC	CACTGGCAGG*ccttcctgagttaga*	

Nucleotide numbers at the 5' and 3' ends of exons 1B to 24 corresponds to that of the CUTL1 cDNA sequence (GenBank, accession No. M74099). Nucleotide numbers at the 5' and 3' ends of exons 25 to 33 corresponds to that of the CDP/Cut cDNA sequence (GenBank, accession No. M74099). Intron and exon nucleotide sequences are shown in small and capital letters, respectively.

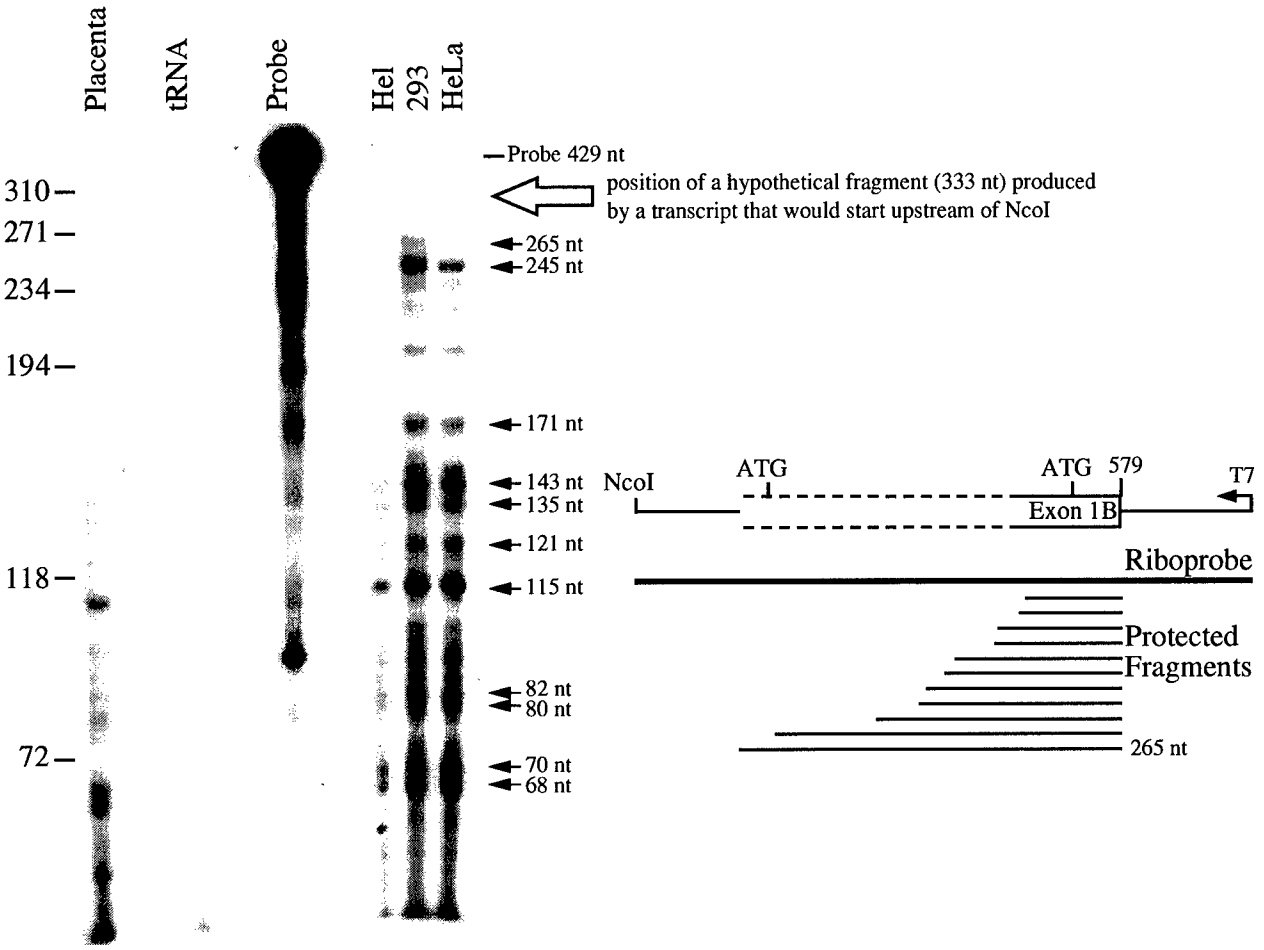






Zeng et al., Fig. 1

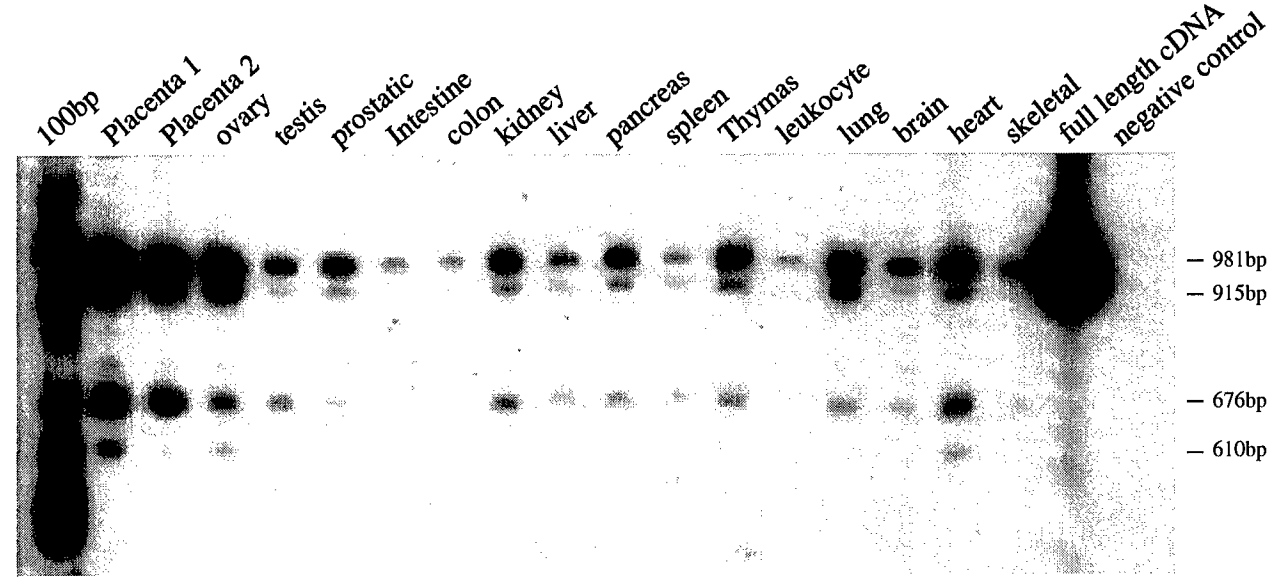
A

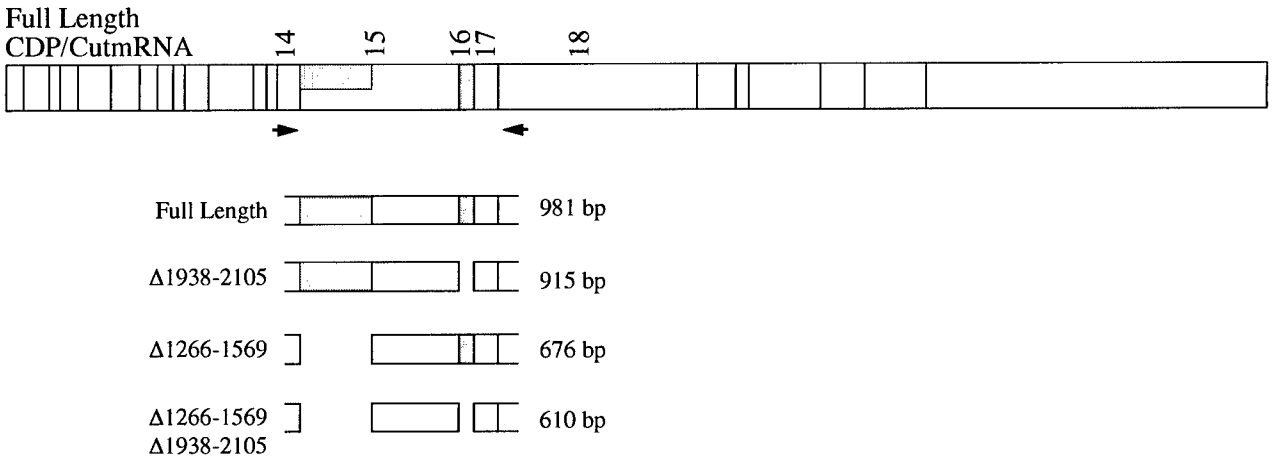


B

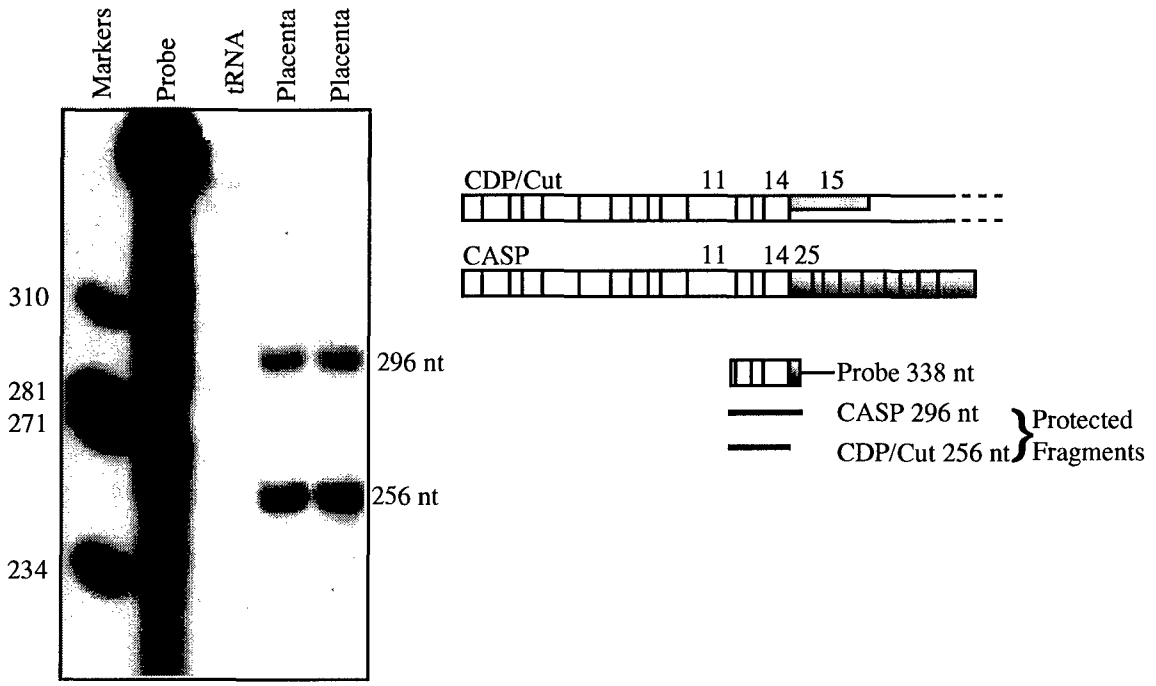
```
          10         20         30         40         50         60         70         80         90        100
GCGCATGTGTCGCTGGGTCCCGGTGTCCCGGACCGCCGGTGCGGCGCGGGAGGAGGAGGCTGCAACTTTCCCCTAGGCAGGCCCCAAGCTGTTCCCCCGG
         110        120        130        140        150        160        170        180        190        200
GTGCCCGGGCTGTGTACGCTGCCTCCCCGCCGCCGGTCCGAGCCGCGATCGCCGCGCGCCTGGGGGCTCCGGCGGCCCCCCGGTGACCTGCGGCGCCGGGT
         210        220        230        240        NcoI    260        270        280        290        300
CCCTTCCTTGCGTCCCCTCGGGGCTTCTGCCCTCTCTGCAGGGCCCGCCATGGAGCGCGCGGCGGGTGCGCTCGGAGATGCCTTCTCGGGCCGGGTTCTC
                                                  M  E  R  A  A  G  A  L  G  D  A  F  S  G  R  V  L
         310        320        330        340        350        360        370        380        390        400
GAAAGCAGAGCCCCGGGGGCCCGTTGGGACCGTGGCATGCCGGGCGGTGGTCCGCGCGCGGAGATCCCGGCGCCGCGGGGGGACCGTGCCGGGGCTCCCC
  E  S  R  A  P  G  A  R  W  D  R  G  M  P  G  G  G  P  R  A  E  I  P  A  P  R  G  D  R  A  G  A  P
         410        420        430        440        450        470        480        490        500
GGCCCCGCGCCCGAGTCGCCGCCGTGCCCAAGGGGCGAGTGCCTGCCTGCCACCCCCCGCCCGGAGGAGCGGCGCACCCTTAGGGTCCGCCCCCGCACGTGC
R  P  A  P  E  S  P  P  C  P  R  G  E  C  L  P  A  T  P  R  P  E  E  R  R  T  L  R  V  R  P  A  R  A
    508        520        530        540        550        560        570        579
CAGCTCGGCGCCCGCGGCGCCGGGACAGCCCCGGGACTCTGCCAGGTGGATGTTGTGCGTACGCGGAGCCAGGTTGAAG
  S  S  A  P  A  A  P  G  Q  P  R  D  S  A  R  W  M  L  C  V  R  G  A  R  L  K
```

Start of HS CDP cDNA sequence





Zeng et al., Fig. 3



Zeng et al., Figure 4

Prognostic Significance of Loss-of-Heterozygosity of the CUTL1 Putative Tumor Suppressor Gene in Breast Cancers

Original Statement of Work (July 96)

Task 8	Months 12-36	In breast tumors with CUTL1 LOH, determine whether the remaining allele is mutated.
	Months 12-24	For the tumors for which RNA is available, clone CUTL1 cDNAs by RT-PCR.
	Months 12-24	Determine the CUTL1 exon/intron structure, including the sequence at the 5' and 3' end of introns. Design and test primers for PCR-amplification of CUTL1 exons.
	Months 24-36	Sequence analysis of CUTL1 genomic sequences and cDNAs. Screening of mutations in CUTL1 genomic sequences and cDNAs, using dideoxy fingerprinting (ddF). Confirmation of mutations by direct sequencing.
Task 9	Months 12-36	Preparation of expression vectors to test the functionality of CUTL1 products expressed in breast tumors. Test the DNA binding activity of CUTL1 products in breast tumors: perform transfections and test cell extracts in electrophoretic mobility shift assays. Test repression potential of CUTL1 products in breast tumors: perform cotransfection studies using appropriate effector and reporter constructs.

As described in my 1999 annual report, CDP/Cut cDNAs have been cloned from 5 breast tumors with CUTL1 LOH. DNA sequencing of the regions encoding for functional domains of the proteins failed to reveal any mutation. In addition, the cDNAs were inserted into mammalian expression vectors and expressed in NIH3T3 cells to test whether the proteins encoded by these cDNAs would bind to DNA and repress transcription from a reporter construct containing CDP/Cut binding sites. No difference was noted in any of these assays. In light of these results as well as the novel findings detailed in the annual report, I propose to replace Tasks 8 and 9 of the original Statement of Work with other tasks (described in the next pages and numbered in consecutive numbers).

Prognostic Significance of Loss-of-Heterozygosity of the CUTL1 Putative Tumor Suppressor Gene in Breast Cancers

Revised Statement of Work

HYPOTHESES

1. I hypothesize that, in contrast to the full length CDP/Cut protein, the short CDP/Cut proteins stimulate cellular proliferation.
2. I hypothesize that a shorter CDP/Cut protein is responsible for the CDP/Cut DNA binding activity that is detected in breast tumor cell lines.
3. I hypothesize that elevated expression of a novel CDP/Cut mRNA initiated within intron 20 is associated with distinct biological properties.

OBJECTIVES

Using a tissue culture system, we will test the hypothesis that short CDP/Cut proteins stimulate, whereas full length CDP/Cut hinders, cellular proliferation.

To better characterize the CDP/Cut proteins expressed in breast tumor cell lines and breast tumors, we will develop novel antibodies raised against short portions of the protein.

We will use RT-PCR assays to analyze the expression of the intron 20 CDP/Cut mRNAs in breast tumors, and correlate its expression with various clinico-pathological features.

Task 10: Months 24-30 **Expression vectors for growth promoting/suppressing assays.** Prepare vectors expressing CDP/Cut proteins of varying lengths and with a Myc and an HA epitope tag at the amino- and carboxy-terminal ends, respectively.

Months 30-36 **Growth promoting/suppressing assays in tissue culture systems.** Perform transfections in immortalized cell lines, selecting for G418 resistance. Count stable transformants, expand 5 clones for each construct, verify stability of the integrated DNA, analyze protein expression and activity.

I hypothesize that, in contrast to the full length CDP/Cut protein, the short CDP/Cut proteins stimulate cellular proliferation. As a first assay, vectors expressing CDP/Cut proteins of various lengths, in particular the intron 20-mRNA derived cDNA, will be cotransfected in established cell lines together with a plasmid conferring G418 resistance, and the number of colonies will be counted. Using this assay, we have previously found that forced expression of the full length CDP/Cut protein greatly reduced the number of transformants. Moreover, when transformants were analyzed for the presence and expression of exogenous CDP/Cut, it was found that the CDP/Cut expressing vector was either not present or not expressed. In light of our recent findings, we now suspect that transformants that survived selection are those that could rapidly process the full length CDP/Cut protein into shorter forms. Therefore, CDP/Cut expression and DNA binding activity of G418 resistant clones will be analyzed to verify this hypothesis.

Task 11 Months 24-30 **GST-CDP/Cut fusion proteins for antibody production.** Prepare 5 fusion proteins containing different segments of the CDP/Cut protein. Raise antibodies in rabbits.

Months 30-36 **Characterize CDP/Cut antibodies** in immunoblotting, immunoprecipitation and electrophoretic mobility shift assays using extracts from our panel of breast tumor cell lines.

Our current polyclonal antibodies have been raised against large portions of the protein. Because of this, it is not possible to delineate the amino and carboxy-terminal boundaries of shorter CDP/Cut proteins. Furthermore, with these antibodies too many bands are detected and we cannot currently identify a shorter protein species whose presence and intensity would correlate with CDP/Cut DNA binding activity. We will

generate novel antibodies that recognize specific regions of the protein: the region upstream of CR1 but that is not contained within CASP (CDP-alternatively spliced product, see Fig. 1C); the linker region between CR1 and CR2; the linker region between CR2 and CR3; the first portion of the carboxy-terminal domain, immediately downstream of the homeodomain; and the last portion of the carboxy-terminal domain. These antibodies will help us define the boundaries of the shorter CDP/Cut proteins, compare CDP/Cut protein expression and activities in breast tumor cell lines, and eventually develop immunocytochemistry based assays to monitor CDP/Cut protein expression directly in breast tumors.

Task 12 Months 24-36 **Investigate expression of the intron 20-initiated CDP/Cut mRNA in breast tumors** to determine whether this could be used as a marker for lobular carcinoma.
Perform RT-PCR analysis on a panel of breast tumors representing various classes: lobular carcinoma, well differentiated ductal carcinoma, moderately differentiated ductal carcinoma, and poorly differentiated ductal carcinoma.

Our preliminary data indicated that the intron 20-initiated CDP/Cut mRNA was expressed at higher level in two breast tumors: an invasive lobular carcinoma (C8961B), and an invasive ductal carcinoma, but with a very diffuse growth pattern (A168). When a retroviral vector expressing this mRNA was introduced into MCF-12A cells, a large fraction of the G418-resistant colonies exhibited an altered morphology: cells did not form densely packed foci but instead tended to grow as isolated cells. Thus, the phenotype we observed in tissue culture correlated well with the higher expression of this mRNA in two carcinomas with a diffuse growth pattern. I propose to analyze the expression of the intron 20-initiated CDP/Cut mRNA in a panel of breast tumors with different clinico-pathological features. In particular, we will verify if indeed this mRNA is expressed preferentially in breast tumors with lobular characteristics. If this association was to be confirmed, expression of this mRNA could be used a marker to predict the pattern of growth and invasion of a given breast tumor. Results obtained in these experiments will determine the pertinence of performing further studies to investigate the value of this novel mRNA as a prognostic marker.

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y)

1 JUN 2001

MEMORANDUM FOR Administrator, Defense Technical Information Center (DTIC-OCA), 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports. Request the limited distribution statement for reports on the enclosed list be changed to "Approved for public release; distribution unlimited." These reports should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Judy Pawlus at DSN 343-7322 or by e-mail at judy.pawlus@det.amedd.army.mil.

FOR THE COMMANDER:

PHYLLIS M. RINEHART
Deputy Chief of Staff for
Information Management

Encl

DAMD17-94-J-4413	ADB261602
DAMD17-96-1-6112	ADB233138
DAMD17-96-1-6112	ADB241664
DAMD17-96-1-6112	ADB259038
DAMD17-97-1-7084	ADB238008
DAMD17-97-1-7084	ADB251635
DAMD17-97-1-7084	ADB258430
DAMD17-98-1-8069	ADB259879
DAMD17-98-1-8069	ADB259953
DAMD17-97-C-7066	ADB242427
DAMD17-97-C-7066	ADB260252
DAMD17-97-1-7165	ADB249668
DAMD17-97-1-7165	ADB258879
DAMD17-97-1-7153	ADB248345
DAMD17-97-1-7153	ADB258834
DAMD17-96-1-6102	ADB240188
DAMD17-96-1-6102	ADB257406
DAMD17-97-1-7080	ADB240660
DAMD17-97-1-7080	ADB252910
DAMD17-96-1-6295	ADB249407
DAMD17-96-1-6295	ADB259330
DAMD17-96-1-6284	ADB240578
DAMD17-96-1-6284	ADB259036
DAMD17-97-1-7140	ADB251634
DAMD17-97-1-7140	ADB259959
DAMD17-96-1-6066	ADB235510
DAMD17-96-1-6029	ADB259877
DAMD17-96-1-6020	ADB244256
DAMD17-96-1-6023	ADB231769
DAMD17-94-J-4475	ADB258846
DAMD17-99-1-9048	ADB258562
DAMD17-99-1-9035	ADB261532
DAMD17-98-C-8029	ADB261408
DAMD17-97-1-7299	ADB258750
DAMD17-97-1-7060	ADB257715
DAMD17-97-1-7009	ADB252283
DAMD17-96-1-6152	ADB228766
DAMD17-96-1-6146	ADB253635
DAMD17-96-1-6098	ADB239338
DAMD17-94-J-4370	ADB235501
DAMD17-94-J-4360	ADB220023
DAMD17-94-J-4317	ADB222726
DAMD17-94-J-4055	ADB220035
DAMD17-94-J-4112	ADB222127
DAMD17-94-J-4391	ADB219964
DAMD17-94-J-4391	ADB233754