Award Number: DAMD17-99-1-9267

TITLE: Long Interspersed Nuclear Element (LINE) Mediated Genomic Instability

PRINCIPAL INVESTIGATOR: Lalitha Nagarajan, Ph.D.

CONTRACTING ORGANIZATION: The University of Texas M.D. Anderson Cancer Center Houston, Texas 77030

REPORT DATE: July 2001

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release; Distribution Unlimited

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

Lalitha Nagarajan, Ph.D.

The University of Texas M.D. Anderson Cancer Center
Houston, Texas 77030
E-Mail: inagarajan@notes.mdacc.tmc.edu

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

In the second year, we have refined the chromosomal location as well as characterized the protein product of the novel homeo box gene MXL. In searching for a more global instability
Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

N/A Where copyrighted material is quoted, permission has been obtained to use such material.

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

N/A 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.

N/A 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).

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

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

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

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

[Signature]
Date 7/25/01

Signatures
# Table of Contents

Cover .......................................................................................................................... 1

SF 298 ......................................................................................................................... 2

Foreword ....................................................................................................................... 3

Table of Contents ......................................................................................................... 4

Introduction .................................................................................................................. 5

Body .............................................................................................................................. 6-8

Key Research Accomplishments .................................................................................. 9

Reportable Outcomes .................................................................................................. 10

Conclusions .................................................................................................................. 11

References .................................................................................................................... 12

Appendices ................................................................................................................... 13-18
INTRODUCTION

The overall goals of this proposal are to isolate target oncogenes, tumor suppressor genes and regions of chromosomal instability by the use of a long interspersed nuclear element based PCR assay. In the second year of this award we exploited the near completion of the human genome sequencing with bioinformatic tools to achieve our goals.

We targeted both known and novel genomic loci for potential candidate genes. The first region was the c-Met oncogene, because of its association with poor prognosis in node negative breast cancer (1). The second line of investigations on a novel sequence specific single stranded DNA binding protein (SSDPs) stemmed from our parallel studies on human leukemia with poor prognosis. Originally isolated in the chicken as a gene whose product binds pyrimidine rich mirror repeats elements (2), the existence of three members that map regions of chromosomal instability was unknown until our studies uncovered this novel family. In the last quarter of the second year, our investigations centered on characterization of the protein products for the human MIXL gene reported last year and the SSDP2 gene product in breast cancer cell lines. Our progress with these promising candidate genes is likely to yield biologically and clinically relevant reagents to diagnose breast cancer with enhanced sensitivity.
Body

The first part of Year 2 of the proposal was to focus on tasks v and vi of Specific Aim 1 which were to Develop PCR for minimal material and correlate data from tasks (i-v) of specific aim 1 with stage and grade to determine consistent quantitative or qualitative patterns can be discerned. The remaining part was to complete specific Aim 2 that was to localize PCR rescued fragments to chromosomal regions and then identify whether these loci are targets of allelic imbalance in breast cancer.

With the near completion of the human genome sequencing studies under this specific aim were re-focused and approached in a cost and time effective manner by utilizing bioinformatic tools. In addition, we also expanded the scope of the proposed studies by initiating experiments on the protein products of potentially important candidate genes. Thus the current report comprises of two parts: (1) Chromosomal localization of regions of instability and identification of candidate genes and (2) Development of antibodies and characterization of expression of potential tumor promoting and tumor suppressor genes in breast cancer cell lines.

1. Chromosomal localization of regions of instability and identification of candidate genes:

1.1 Chromosome 7q31.1 locus:

A search of the human genome database for novel LINE rich segments showed a high LINE content on chromosome 7q31.1 The Met oncogene encoding a tyrosine kinase receptor, and a marker for poor prognosis in breast cancer (1) localizes to this region. Therefore, we searched the genomic sequences in the public domain for putative polymorphic sequences. A Bacterial Artificial Chromosome 354 L07 had a > 60% repetitive sequences and an unusually high number of L1Hs elements. We developed three novel markers from these sequences designated ca1, ca2 and gata (fig.1)

In collaboration with Dr. Nour Snejie in the Dept. of Pathology, we dissected 25 breast tumors and isolated normal and tumor DNAs. We developed PCR conditions to detect allelic imbalance for ca1, ca2, and gata in minimal patient material. These markers along with 7 other highly polymorphic markers (designated D7S***) were used in allelo typing studies of 25-breast tumor and matched control DNAs. 13 out of 25 tumors examined showed allelic imbalance. As shown in Fig.1 we could identify two critical regions, one flanked by the C-MET protooncogene and second a more centromeric region between the loci D7S471 and D7S1817.
All of the 25 breast tumors used to screen for allelic imbalance of 7q31.1 showed lymph node involvement. We will interrogate the clinical correlation of both the previously identified chromosome 1q41-42 locus as well as additional loci characterized here once we have conditions optimized to detect the candidate gene expression by immuno staining.

1.2 A novel family of sequence specific single stranded DNA binding proteins localize to regions of allelic imbalance:

Another novel family of evolutionarily conserved genes, which localize to regions of genomic instability, were identified in our laboratory as part of our studies on human leukemia. The three members of this family encode a putative sequence specific single stranded DNA binding activity (SSDPs1-3). More importantly these map to regions of deletions in breast cancer identified by several studies (3-5). SSDP1 maps to chromosome 1p32, SSDP2 to chromosome 5q13.3 and SSDP3 to chromosome 19p12 (Fig.2).

In addition, we also examined the genomic organization of SSDP2, which localizes near a chromosomal break point in the breast cancer cell line SKBR3, as well as a human leukemia cell line ML3 (Fig.3). The gene is encoded by 17 exons and there are two large introns (>100kbp), 1 and 4 with several copies of L1 Hs elements. Interestingly SSDP1 and 3 also have identical genomic organizations with 17 exons although the intron sizes are smaller.

In order to determine whether the expression pattern of the SSDP genes in altered in cancer we initiated preliminary screens with the SSDP2 gene. Relative RT PCR in a variety of malignant cell lines including six mammary cell lines is shown in Fig.4. In contrast to normal tissues and some malignant cell lines (Weri), SKBR3, MDA- MB453, MDA- MB435, MDA-MB468, SKBr-3 and MCF 7) as well as the SV40 immortalized mammary epithelial cell line MCF10A appear to express low SSDP2 transcripts.

2. Development of antibodies, characterization in breast cancer cell line models

A corollary of our hypothesis is that altered expression, either enhanced expression in the case of oncogenes or loss of expression in the case of putative tumor suppressor genes at the site of chromosomal instability, contribute towards the phenotypic evolution of breast cancer cells, we developed highly specific antibodies against peptide epitopes of both MIXL and SSDP2 gene products. Two different
epitopes were identified and affinity purified antiserum were generated. The specificity of the antibodies was verified in cell lines induced to over express under transient transfection conditions. Thus the antibodies were evaluated stringently and then utilized in immunoblotting experiments with human breast cancer cell lines.

2.1 MIXL2 protein is expressed in immortalized and malignant mammary cell lines:

As stated in last year's report, expression of MIXL appears to be highly restricted by RTPCR and Northern blotting analyses. This experimental evidence is confirmed by Bioinformatic analyses of cDNA sequences in the public domain. To date the dBEST Database of more than 10000 entries from over 30 tissues, contains only 4 sequences for the MIXL2 gene. These are from a infiltrating ductal carcinoma in situ, germ cell tumor, normal pre B lymphoid cells and the highly metastatic fibrosarcoma cell line HT1080. Interestingly, all the breast cancer cell lines (SKBR3, MDA- MB435, MDA-MB453, MDA-MB468 and MCF 7) and the SV40 immortalized mammary epithelial cell line MCF10A express the MIXL protein as shown in Fig. 5. This is in contrast to other malignancies including hematopoietic cancer where we detect a highly restricted expression. Thus aberrant MIXL expression might confer a proliferative/survival advantage to breast cancer cells.

2.1 SSDP2 protein is not expressed in immortalized and malignant mammary cell lines:

The putative sequence specific single stranded DNA binding proteins in themselves may be involved in genomic stability as they are postulated to bind single stranded sequences that loop out when mirror repeat elements in the DNA assume a triple helical confirmation. Loss of expression of these elements that reside at regions of genomic instability and deletion may confer a survival/loss of ability to repair DNA damage advantage to malignant cells. As shown in Fig.6 none of the breast cancer cell lines as well as the immortalized breast epithelial expresses the SSDP2 protein. This is in contrast to certain T lymphoid cell lines, which express abundant SSDP2 protein. Our future studies will elucidate the mechanism by which loss of SSDP2 expression confers survival advantage in cancer.
Key Research Accomplishments:

- Characterization of the 7q31.1 loci flanking the MET protooncogene
- Identification of a novel family of sequence specific single stranded DNA binding protein (SSDPs1-3) and localization to human genome
- Development a PCR assay to detect expression of SSDPs 1-3
- Development of MIXL specific antibodies
- Detection of MIXL expression in breast cancer cells in culture
- Development of antibodies to SSDP2
- Lack of SSDP2 expression in breast cancer
(8) Reportable outcomes:

Manuscripts in preparation:

1. Hejlik DP and Nagarajan L. Retro transposition and recombination of LINE-1 Elements as a mechanism of chromosomal deletion and genomic instability (Manuscript being revised for resubmission)

2. Guo W., Chan A., Liang H., Etkin L.D., and Nagarajan L. A human Mix.1 like gene shows functional conservation with the Xenopus homolog. (Manuscript to be submitted by Aug. 01)


Development of FISH and immunohistochemical staining probes:

We have characterized a BAC probes from chromosomes 1p32, 5q13.3, 7q31.1 and 19p13.2 that can be used in Fluorescence in situ hybridization.

We have developed highly specific antibodies to MIXL2 and SSDP2, which can be used in immuno histochemical staining.

Informatics:

The full length MIXL 2 sequence has been submitted to Genbank.

Full length cDNA and exon specific genomic sequences for SSDPs 1-3 have been submitted to Genbank

Funding applied for based on work supported by this award:

Some of the results obtained in the present award period contributed to the development of an RO1 proposal to NIH entitled “SSDP2 gene pathway in Myeloid Neoplasm”, submitted Feb.1, 01.
Conclusions:

1) In breast cancer, there are two critical regions of allelic imbalance centromeric of the c-Met oncogene. Whether imbalances of these intervals are selected for or result from an overall genomic instability remains to be investigated.

2) The sequence specific single stranded DNA binding proteins are encoded by a novel gene family. All three members localize to regions of chromosomal deletions in breast cancer.

3) A novel paired type- homeodomain protein Mixl2 is expressed in immortalized as well as transformed mammary epithelial cell lines.

In summary, we have delineated 5 distinct regions of instability that harbor potential target genes.
REFERENCES


Fig. 1  Delineation of Critical Regions of Allelic Imbalances on Chromosome 7q

- LOH
- NI
- Retained both alleles
- Negative
- Not tested

CR - CRITICAL REGION

Sample Number

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
<th>7</th>
<th>8</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Markers:

- D7S1817
- D7S2418
- D7S523
- D7S471
- D7S687
- gata
- Ca2
- Ca1
- D7S486
- MET
- D7S672

CR regions
Fig. 2  Predicted open reading frames of SSDPs1-3.

Arrows denote intron- exon boundaries
Conserved amino acid residues are bolded
Fig. 3 Genomic Organization of SSDP2

Panel A shows the genomic BAC clones and panel B shows individual exons. The large, L1Hs rich introns of unknown sizes are denoted by slashes.
Fig 4. SSDP2 Transcript Levels are Low in Breast Cancer Cell Lines

Breast tumor cell lines
Fig. 5 MXL Protein is expressed in Breast Cancer Cell lines

- SKBR3
- MDA-MB435
- MDA-MB453
- MDA-MB468
- MCF7
- MCF10A
Tagged SSDP2 was used as control.
COS-1 cell line transfected with Flag epitope

Cancer Cell Lines

Fig. 6 Abundance of SSDP2 Protein Expression in Breast