Computer aided analysis of restriction landmark genomic scanning images from tumor and cell line models

James Lambert Patrick
Medical University of Ohio

Follow this and additional works at: http://utdr.utoledo.edu/theses-dissertations

Recommended Citation
http://utdr.utoledo.edu/theses-dissertations/1375

This Dissertation is brought to you for free and open access by The University of Toledo Digital Repository. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of The University of Toledo Digital Repository. For more information, please see the repository's About page.
FINAL APPROVAL OF DISSERTATION
Doctor of Philosophy in Biomedical Sciences

Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Tumor and Cell Line Models

Submitted by:
James Lambert Patrick

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biomedical Sciences

Examination Committee

Major Advisor: Ming You, M.D., Ph.D. ____________________________

Academic Advisory Committee:
Amira Gohara, M.D. ____________________________
William Davis, D.D.S., M.S. ____________________________
Peter F. Lemkin, Ph.D. ____________________________
Peter J. Goldblatt, M.D. ____________________________

Dean of the College of Graduate Studies
Keith K. Schlender, Ph.D. ____________________________

Date of Defense: May 30, 2001
Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Tumor and Cell Line Models

James Lambert Patrick
Medical College of Ohio
2006
Adviser: Ming You
I would like to thank my parents for not telling me what to think, but giving me the tools to explore my own world and make judgments based on my observations.

I would like to thank Dr. You for giving me an opportunity to investigate and discover the world of science at the doctorate level.

I would like to thank Dr. Davis for his support and guidance during the early years of my academic career at MCO.

I would like to thank Dr. Gohara for all her help and guidance during my stay at MCO.

I would like to thank Dr. Lemkin for his support and his insight into the world of computational biology.

I would like to thank Dr. Goldblatt for his help during his long process and his unique way of seeing the world.

I would like to thank Jeff Liu for the use of his manuscripts and data in this work.

I would also like to thank Jenny Holzer for putting up with me for so many years and her support and understanding and for also helping me to explore our world.

I would like to thank the Graduate School staff and faculty in guiding me throughout this whole process.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>LITERATURE</td>
<td>3</td>
</tr>
<tr>
<td>MANUSCRIPT I The Reduction of Genetic Alterations by 4-Hydroxyphenylretinamide in MethylNitrosourea Induced Rat Mammary Tumors</td>
<td></td>
</tr>
<tr>
<td>MANUSCRIPT II Detection of Genomic Alterations in Human Cervical Cancer by Two-Dimensional Gel Electrophoresis</td>
<td></td>
</tr>
<tr>
<td>MANUSCRIPT III Computer Assisted Analysis Of Restriction Landmark Genomic Scanning Images Of DNA From K-Ras Oncogene Transformed NIH-3T3 Cells</td>
<td></td>
</tr>
<tr>
<td>MANUSCRIPT IV Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Human Bronchial Epithelial Cells Irradiated with High-LET Alpha Particles</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>122</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>129</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>157</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>168</td>
</tr>
</tbody>
</table>
INTRODUCTION

Cancer is one of the leading killers in The United States and has been an area of intensive research and investigation for many years. Despite many advances in our knowledge base about the disease process and the production of effective treatments there is still much more that has to be done before we have a panacea for cancer. The past two decades has been a very exciting time for cancer research. With the advent of new technologies, particularly PCR and gel electrophoresis, and greater computational power we have been able to accumulate more information about cancer and its pathogenesis.

Evidence points to genetic alterations such as loss of heterozygosity (LOH), amplification, and deletion as being critical in the neoplastic transformation of cells. It is believed that a cell is transformed through a series of events. Previous techniques have focused on individual loci to evaluate genetic abnormalities. The analysis of these loci throughout a genome is time-consuming and laborious using these previous techniques. This problem has been overcome by the use of the new generation of genetic assays. One particular technique called restriction landmark genomic scanning (RLGS) is the focus of this
dissertation. RLGS can simultaneously illustrate a number of genetic abnormalities at once, which helps to provide an overall picture of the interplay of genetic events and their participation in the pathogenesis of cancer.
Cancer and Multistep Carcinogenesis

It is generally accepted that cancer is the result of a multistage process which leads to the loss of the regulation cellular growth and differentiation. Evidence indicates that it is an accumulation of genetic changes that alter the normal pattern of cell growth, proliferation, and expansion that leads to malignancy. Previous investigations have indicated that there are three steps, in this process: initiation, promotion, and progression (Liu, 2000).

Initiation is marked by a genetic change located within the genome of a cell. This mutation may be the result of exposure to exogenous agents such as chemicals, radiation, or viruses (Minamoto et al., 1999) or may also be the result of a random or inherited mistake that occurred during replication. If the mutation is not corrected before replication then it will be passed down and fixed into the genome of the progeny cells (Fearon, 1997).

Promotion or the clonal expansion of the initiated cell is the next phase in carcinogenesis and relates to the development of benign tumors and or focal proliferative
lesions. Promotion is believed to be epigenetic in which agents (chemical or physical) are able to selectively propagate the initiated cell’s progeny and produce expansion (Sugimura et al., 1992; Fearon, 1997).

The progression phase is the transformation of a benign tumor or a focal proliferative lesion into a malignant tumor. A number of attributes of the tumor change during this phase. These attributes include: Becoming locally invasive and capable of distal metastasis; The alteration histomorphology and gross morphology of the tumor; Changes in the rate of cellular proliferation, and differentiation (Liu, 2000).

**Oncogenes and Proto-Oncogenes**

Proto-Oncogenes code for cellular proteins responsible for signal transduction, stimulation of cell growth, and the activation of transcription factors that help the cell pass through the cell growth cycle. Proto-Oncogenes may become targets of mutation, thus becoming oncogenes. The end result of which may be: uncontrolled stimulation and production of growth factors and aberrant cell growth (Anderson et al., 1992). The transformation of a proto-oncogene into an oncogene can occur in a number of different ways. These include point mutations, chromosome
rearrangements, gene amplifications, and viral insertions (Rhim 1988).

The discovery of the first Proto-Oncogene began with Rous in 1911 who observed that sarcomas in chickens were caused by infectious particles later to be determined as viral in nature and named the Rous sarcoma virus (Varmus, 1989). Bishop and Varmus identified a gene in the Rous virus, v-src that was responsible for transforming normal chicken myocytes into neoplastic cells (Varmus, 1989). Later, it was determined that there was a normal cellular counterpart to the v-src oncogene know as c-src (Table I).

Ras Gene Family

Members of the ras family of proto-oncogenes are comprised of genes derived from the Harvey, and Kirsten rat sarcoma retroviruses and their normal cellular homologues H-ras, K-ras-2 and N-ras. (Minna, 1993) The ras gene family encodes for 21-kd proteins that are inner plasma membrane-associated GTPases that bind with GTP to cleave it to GDP; they are related to cellular signal transduction (Minna, 1993). Activated ras oncogenes have been found in large number of human tumors and are frequently detected using the NIH3T3 transfection assay. The involvement of ras proto-oncogenes in human tumors was first detected by the
<table>
<thead>
<tr>
<th>GENE</th>
<th>PROTEIN PROPERTIES</th>
<th>TUMOR TYPES</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-RAS</td>
<td>p21 GTPase</td>
<td>Pancreatic, colorectal, lung (adenocarcinoma), endometrial, other carcinomas</td>
</tr>
<tr>
<td>N-RAS</td>
<td>p21 GTPase</td>
<td>Myeloid leukemia, colon cancer</td>
</tr>
<tr>
<td>H-RAS</td>
<td>p21 GTPase</td>
<td>Bladder carcinoma</td>
</tr>
<tr>
<td></td>
<td>EGFR (ERBB)</td>
<td>Gliomas, squamous, other carcinomas</td>
</tr>
<tr>
<td>NEU(HER2/ERBB)</td>
<td>Growth factor (EGF) receptor</td>
<td>Breast, ovarian, gastric, other carcinomas</td>
</tr>
<tr>
<td>c-MYC</td>
<td>Transcription factor</td>
<td>Burkitt's lymphomas, Small cell lung carcinoma (SCLC), other carcinomas, glioblastoma</td>
</tr>
<tr>
<td>N-MYC</td>
<td>Transcription factor</td>
<td>Neuroblastoma, SCLC, glioblastoma</td>
</tr>
<tr>
<td>L-MYC</td>
<td>Transcription factor</td>
<td>Burkitt's lymphoma, SCLC</td>
</tr>
<tr>
<td>BCL-2</td>
<td>Anti-apoptosis protein</td>
<td>B-cell lymphoma (follicular type)</td>
</tr>
<tr>
<td>CYCD1</td>
<td>Cyclin D, cell cycle control</td>
<td>Breast and other carcinomas, B-cell lymphoma, parathyroid adenoma</td>
</tr>
<tr>
<td>BCR-ABL</td>
<td>Chimeric nonreceptor tyrosine kinase</td>
<td>CML, ALL (T cell)</td>
</tr>
<tr>
<td>RET</td>
<td>GDNF receptor tyrosine kinase</td>
<td>Thyroid cancer (papillary type), Thyroid cancer (medullary type)</td>
</tr>
<tr>
<td>CDK4</td>
<td>Cyclin-dependent kinase</td>
<td>Sarcoma, glioblastoma, Melanoma; colorectal, endometrial,</td>
</tr>
<tr>
<td>beta-CAT</td>
<td>Transcriptional co-activator,</td>
<td></td>
</tr>
<tr>
<td>Genes</td>
<td>Function/phenotype</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>(CTNNB1)</td>
<td>links E-cadherin to cytoskeleton. ovarian, hepatocellular, and other carcinomas, hepatoblastoma</td>
<td></td>
</tr>
<tr>
<td>HST</td>
<td>Growth factor (FGF-like) Gastric carcinoma</td>
<td></td>
</tr>
<tr>
<td>PML-RARalpha</td>
<td>Chimeric transcription factor APL</td>
<td></td>
</tr>
<tr>
<td>E2A-PBX1</td>
<td>Chimeric transcription factor Pre-B ALL</td>
<td></td>
</tr>
<tr>
<td>MDM-2</td>
<td>p53 binding protein Sarcoma</td>
<td></td>
</tr>
<tr>
<td>GLI</td>
<td>Transcription factor Sarcoma, glioma</td>
<td></td>
</tr>
<tr>
<td>TTG</td>
<td>Transcription factor T-cell ALL</td>
<td></td>
</tr>
<tr>
<td>AKT2</td>
<td>Serine/threonine kinase Pancreatic and ovarian carcinoma</td>
<td></td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Catalytic subunit of PI3-kinase Ovarian carcinoma</td>
<td></td>
</tr>
</tbody>
</table>

Adopted and modified form Abeloff: Clinical Oncology, 2nd Ed., Copyright © 2000 Churchill Livingstone, Inc.

Transfection of tumor DNA into NIH3T3 cells (Murray, 1981). The activated ras genes may contain point mutations at the 12th, 13th or 61st amino acid of the 189 amino acid, 21-kd protein (Barbacid, 1986).

**Myc Gene Family**

The myc family was discovered by the isolation of the v-myc transforming gene from the avian myelocytomatosis virus. Its cellular counterpart is called c-myc and two other similar genes were isolated called N-myc (isolated from neuroblastoma) and L-myc (isolated form SCLC) (Minna, 1993). The myc family of genes (c-myc, N-myc, and L-myc)
are DNA-binding phosphoproteins which are involved in transcriptional control and may play an important role in regulating cell growth. Unlike the ras family of oncogenes, no coding mutations have been found in the myc genes in human cancer; rather, changes in DNA amplification and RNA expression of myc genes is seen many types of cancer. C-myc amplification also is a negative prognostic indicator. Cancers with C-myc over expression have increased radioresistance and a lower mean survival time (Minna, 1993).

C-myc

C-myc proto-oncogene encodes for two phosphoproteins, p64c-,myc and p67c-myc, and is located on human chromosome 8 (Taub, 1982). DNA binding properties of these proteins suggest a role in transcriptional control and a role in regulating cell growth.

**Tumor suppressor genes**

Tumor suppressor genes require the loss of both alleles to become inactive (Minna, 1993). These genes are characterized by their recessive mode of action; i.e., their absence predisposes a cell to cancer. Knudson proposed the existence of tumor suppressor gene in human
tumors. This model states that two genetic lesions are required for tumor development and is refer to as the “two knockout theory.” In the hereditary form of the disease, one lesion is transmitted via the germline whereas the second lesion is caused by somatic mutation. In the sporadic form of the disease, both lesions are caused by somatic mutation, thus explaining relative rarity of the sporadic form of the disease.

Retinoblastoma Gene

The retinoblastoma gene was the first suppressor gene discovered and is found in childhood retinoblastoma. The retinoblastoma gene normally codes for a nuclear phosphoprotein and both copies of the gene are typically inactivated in retinoblastomas. In Small-cell lung cancer tumors and tumor cell lines studied, the retinoblastoma locus showed a deletion or rearrangement of the DNA in 20% of the cases, absent or barely detectable mRNA expression in 55% of the cases, and undetectable protein in 70% of the cases (Harbour, 1988; Hensel, 1990).

p53 Gene

The p53 gene normally codes for a nuclear phosphoprotein which is needed for the propagation of the
cell cycle, DNA synthesis and repair, cell differentiation, and programmed cell death (Minna, 1993). Mutations of the p53 tumor suppressor gene are currently the most common genetic alteration identified in human cancers (Harris, 1993). The abnormalities in the region of the p53 gene in human lung cancer were first identified on the short arm of chromosome 17. The mutations of the p53 gene occur throughout the length of the entire gene (Minna, 1993). The mutations can be missense mutations, nonsense mutations, DNA deletions, or insertions. Mutations of the p53 gene have been found in 100% of the tumors, and tumor cell lines from patients with small-cell lung cancer and 42% of the tumor and tumor cell lines from patients with non-small-cell lung cancer (Minna, 1992).

Genetic Alteration in Cancer

Point Mutations

As mentioned earlier, oncogene activation can occur by several different mechanisms. These include point mutations, chromosomal rearrangement, translocation gene amplification, and viral activation (Bishop, 1991). Point mutations are the result of base-pair changes within a single DNA codon. There are four different types of point mutations: (1) the insertion of a nucleotide, (2) the
deletion of one or more base pairs, (3) transitions, and (4) transverions. These events may produce an alteration in the coding sequence of a specific amino acid or may cause a premature termination of a coding sequence. The end results are altered proteins that may become oncogenic.

Chromosomal rearrangement or translocation

Oncogene activation may also occur as a result of translocation or chromosomal realignment or rearrangement. These changes may move a proto-oncogene to another area of the genome that causes it’s activation, or move a promoter element close to a proto-oncogene thus activating it (Croce 1987).

An example of chromosome rearrangement producing an oncogene activation is found in chronic myelogenous leukemia (CML) (Rabbits, 1994). This rearrangement involves a reciprocal translocation between chromosome 9 and chromosome 22. A proto-oncogene known as c-abl is moved into a region on chromosome 22 called the breakpoint cluster region (BCR) (Shitvelman et al., 1985). The oncogene c-abl is transcribed with the BCR producing an abnormal bcr-c-abl fusion protein which can cause neoplastic transformation. 90% of all chronic myelogenous leukemia cases have this reciprocal translocation between chromosome
9 and chromosome 22 which produces a chromosomal structure called a Ph1 or Philadelphia chromosome (Sawyers, 1992). Burkitt's lymphoma is endemic to Africa and other regions of the world. In 80% of the cases show a chromosomal translocation between chromosome 8 and 14 (Favera et al., 1982). Other translocations, including chromosomes 8 and 2 and chromosomes 8 and 22, have been seen in neoplastic transformation. All of these translocations have one thing in common; They all involve a region on the short arm of chromosome 8 which encodes for the cellular proto-oncogene c-myc (MYC) (Nowell et al. 1990). In the case of 8,14 translocation, MYC is moved close to a locus for a immunoglobulin heavy chain on chromosome 14. It is believed that this immunoglobulin heavy chain locus and its promoter lead to over expression of MYC which leads to the transformation process (Nowell et al. 1990). In the case of the 8,22 and 8,2 translocations, it is MYC’s proximity to immunoglobulin kappa and lambda promoters that leads to transformation (Table II).

Gene Amplification

Gene amplification refers to the increase of the copy number of a particular gene. This amplification serves to increase the expression of the gene product which may cause
transformation. The process of gene amplification occurs when genomic DNA is over replicated. In most cases, this gives rise to two different types of karyotyptic changes: Homogeneous staining regions and Double-Minute chromosomes (Cowell, 1982). Both karyotypic changes represent highly amplified geneomic DNA. A number of studies have shown that the amplification of proto-oncogenes may be an important step in cellular transformation for a number of the human neoplasias. Three particular proto-oncogene families, myc,

Table II. Amplification of myc, erb B and ras Gene Family in Human Tumors

<table>
<thead>
<tr>
<th>Gene</th>
<th>Tumor</th>
<th>Frequency of Amplification (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-myc</td>
<td>Breast cancer</td>
<td>15-23</td>
</tr>
<tr>
<td></td>
<td>Colon carcinoma</td>
<td>3-6</td>
</tr>
<tr>
<td></td>
<td>Lung-squamous cell Carcinoma</td>
<td>12-15</td>
</tr>
<tr>
<td>N-myc</td>
<td>Neuroblastoma</td>
<td>10-31</td>
</tr>
<tr>
<td>c-myc or L-myc</td>
<td>Lung-adenocarcinoma</td>
<td>2-11</td>
</tr>
<tr>
<td>c-myc, N-myc,</td>
<td>Lung-small cell Carcinoma</td>
<td>11-23</td>
</tr>
<tr>
<td>or L-myc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>erb B</td>
<td>Breast cancer</td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>Gastric and esophageal Carcinoma</td>
<td>4-8</td>
</tr>
<tr>
<td></td>
<td>Head and neck Squamous cell Carcinoma</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Glioblastoma</td>
<td>38-50</td>
</tr>
</tbody>
</table>
Adapted from Ericson O. Gene amplification and tumor progression. Biocham Biophys Acta 1155:26, 1993

erb B, and ras, are amplified in a variety of human tumors

Alitalo et al., 1986) (Table III).

Table III. Examples of Proto-oncogenes activated by chromosomal translocations.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Tumor type</th>
<th>Tumor showing gene involvement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-myc</td>
<td>Burkitts lymphoma</td>
<td>&gt;95</td>
</tr>
<tr>
<td>c-myc</td>
<td>Murine plasmacytoma</td>
<td>&gt;95</td>
</tr>
<tr>
<td>c-myc</td>
<td>Acute T-cell leukemia</td>
<td>10-20</td>
</tr>
<tr>
<td>bcl-1</td>
<td>Chronic lymphocytic leukemia</td>
<td>10-30</td>
</tr>
<tr>
<td>bcl-2</td>
<td>Follicular lymphoma</td>
<td>85-95</td>
</tr>
<tr>
<td>c-abl</td>
<td>Chronic myelogenous leukemia</td>
<td>&gt;90</td>
</tr>
<tr>
<td>c-abl</td>
<td>acute lymphocytic leukemia</td>
<td>10</td>
</tr>
</tbody>
</table>

Restriction Landmark Genomic Scanning (RLGS)

In the late 70's and early 80's a number of groups were investigating the separation of DNA by two-dimensional gel systems. In 1979 Fischer and Lerman used a mass
gradient in the first dimension and a denaturing gradient in the second dimension to look at DNA patterns from EcoRI digested E.Coli genomic DNA. Two dimensional separation gel systems using human DNA was performed by Uitterlinden et al. in 1989 using a mass-denaturing system with a hybridization step using minisatellite probes to evaluate tandem repeats (Uitterlinden et al., 1989).

A probe free method was perfected by Hatada and his group in 1991 (Hatada et al., 1991). He demonstrated that by using a rare base pair cutter and subsequent radioactive labeling of the genomic DNA at the cleavage sites, followed by a two-dimensional separation, produced a large number of restriction fragments (1500-3000). Hatada’s group refered to this system as Restriction Landmark Genomic Scanning or RLGS.

DNA from malignant cells generally contains numerous complex genomic aberrations which include loss of heterozygosity, deletions, and amplifications. Restriction Landmark Genomic Scanning has been employed to detect these genomic changes. This technique has a number of distinct advantages in this area of investigation which include: its ability to scan over the entire genome of a species in a short period of time; The assay may be used on any species do to the nonspecific nature of the end labeling system;
when using methylation-sensitive enzymes RLGS can screen and determine the methylation state of a genome; and its ability to scan many restriction landmarks and their subsequent loci simultaneously. In RLGS, genomic DNA is harvested and purified then digested using a rare base pair cutter such as NotI. The cleavage sites are end-labeled using a radioactive probe. The labeled DNA fragments are digested by a second restriction enzyme usually EcoRV. The fragments are run through a one-dimensional electrophorect agarose. After the first dimensional separation, the DNA fragments are “in-gel” digested using a third restriction enzyme usually HindIII. After the last digestion the DNA fragments are separated using polyacrylamide gel electrophoresis. The gel is autoradiographed to produce a RLGS electrophoretogram containing thousands of spots. Each spot represents a DNA fragment and the intensity of the spot correlates with the copy number of that DNA fragment (figure 1.) (Hirotsune et al., 1992). The RLGS electrophoretogram is then digitized for image processing and analysis.

RLGS electrophoretograms and the role of image progressing and analysis RLGS electrophoretograms produce a complex data set containing thousands of spots. This information is
Figure 1. RLGS System

A: Site for restriction enzyme A (Restriction landmark)
B: Site for restriction enzyme B
C: Site for restriction enzyme C

- Blocking (with analogues of nucleotide)
- Landmark cleavage (with restriction enzyme A)
- Labeling (at restriction landmark)

B(A) C A

1 2: Isotope label

(-) The fragmentation with restriction enzyme B)

The first fractionation (by 0.8-1% agarose gel electrophoresis)
The fragmentation of labeled DNA with restriction enzyme C
The second fractionation (by 5-6% polyacrylamide gel electrophoresis)

Autoradiography

of little value unless a reliable method for comparing spot data between control and experimental RLGS electrophoretograms can be determined. Despite the usefulness of RLGS there are a number of inherent problems when dealing with RLGS electrophoretograms. First and perhaps the greatest technical problem deals with "reproducability" of gel patterns. Despite careful attention to detail in attempting to reproduce the exact conditions between gel runs there are nonlinear spot migration patterns between experiments. Visual detection methods can overcome this problem but require a laborious matching process that is very time-consuming and difficult with any large dataset. Another problem arises from the dataset itself, an average RLGS gel will contain around 2000 fragment spots. If an average study contained ten experimental samples and ten control samples and two replicate runs of each sample the investigator would be required to keep track of 80,000 fragment spots. Another difficulty arises from the discrepancies in the end labelling signal. Due to variations in the end-labelling, RLGS autorad exposures between gels. This variation requires that RLGS electrophoretograms be normalized before any meaningful quantitative data with respect to spot intensity can be harvested and analyzed.
Image Processing and Analysis Methods

Warping Method

Warping is a method that requires the operator to identify landmarks in each gel that have undergone little or no migrational change (usually ribosomal DNA). Once these landmarks have been identified a "warp" spot coordinate system is constructed and the gel images are "stretched" to compensate for the nonlinear migration between gels (Figure 2).

Color Overlay System

A color hybridization Macro or program combines two images that have been properly aligned into a single image, which can be qualitatively scanned visually for differences in spot intensity between the two images. The procedure is as follows: "gel A's" grayscale look up table (LUT) is converted to a pure green LUT and "gel B's" grayscale LUT is converted to a pure red LUT. The gel images then are superimposed to produce a hybrid image for visual inspection. If corresponding spots between the two gels have a similar intensity the resultant spot will be yellow. If the spot intensity of "gel A" is greater than "gel B" the resulting spot will color shift towards the green end.
Figure 2. Image-warping or rubber-sheet transformations may be specified by a set of parametric equations that specify the output coordinate for a given input position.

\[
x' = \alpha(x, y) ; \quad y' = \beta(x, y) \rightarrow g(f[x, y]) = f[x', y'] = f[\alpha(x, y), \beta(x, y)]
\]

The input image function \( f(x, y) \) is transformed by the function \( Q{} \)

\[ Q(f(x, y)) \rightarrow g(x, y) = f(x', y') \]

The second-order polynomial warping function, \( Q{} \), can be defined as:

\[
\begin{align*}
  u &= a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 \\
  v &= b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2
\end{align*}
\]

Rewritten in matrix notation using 6 input \([x_{ij}, y_{ij}]\) and 6 output \([u_{ij}, v_{ij}]\)Control points:

\[
Ax = x' \\
\begin{bmatrix}
  1 & 1 & 1 & 1 & 1 & 1 \\
  x_0 & x_0 & x_1 & x_1 & x_2 & x_2 \\
  y_0 & y_0 & y_1 & y_1 & y_2 & y_2 \\
  x_0y_0 & x_0y_0 & x_1y_1 & x_1y_1 & x_2y_2 & x_2y_2 \\
  x_0y_0x_0 & x_0y_0x_0 & x_1y_1x_1 & x_1y_1x_1 & x_2y_2x_2 & x_2y_2x_2 \\
  y_0 & y_0 & y_0 & y_0 & y_1 & y_1 \\
  x_0y_0 & x_0y_0 & x_1y_1 & x_1y_1 & x_2y_2 & x_2y_2
\end{bmatrix}
\begin{bmatrix}
  a_0 \\
  a_1 \\
  a_2 \\
  a_3 \\
  a_4 \\
  a_5 \\
  b_0 \\
  b_1 \\
  b_2 \\
  b_3 \\
  b_4 \\
  b_5
\end{bmatrix} = \begin{bmatrix}
  u_0 \\
  u_1 \\
  u_2 \\
  u_3 \\
  u_4 \\
  u_5 \\
  v_0 \\
  v_1 \\
  v_2 \\
  v_3 \\
  v_4 \\
  v_5
\end{bmatrix}
\]

The coefficients \([a_{ij}, b_{ij}]\) are found by:\(A \times x-1 = A = x' \times -1\)

Once in matrix form, the Visual Numerics math Double Matrix class provides the functionality to find the least-squares solution for the coefficients.

of the spectrum. If the spot intensity of "gel B" is
greater then "gel A" the resulting spot will color shift
towards the red end of the spectrum (Figure 3.). This
relatively simple process has been incorporated in many
biological imaging systems, including comparative genomic
hybridization and microarray technology. A copy of the
source code of the Color Macro used by our group can be
found as Appendix A.

Spot Segmentation and Landmarking System

New computer software has been brought to bear to deal
with the inherent problems of RLGS spot patterns. Unlike
previous methods this type of software has the ability to
segment spots and pair them with corresponding spots in
separate gels. This type of software has four main
components. An image processor which is responsible for
"cleaning up" the image, and reducing background noise
produced during autoradiography. A segmentation algorithm
to separate out spots on a gel image. The algorithm employs
the digital analog of a spatial second derivative of a
Gaussian like function in which the 2nd maximum of the
second derivative determines the outer extent of the spots
outward propagation (Lemkin, 1981). A Landmarking and Spot
Pairing algorithm which is responsible for pairing spots between gels using a minimum Euclidean distance system to accomplish its task. The final component of the program produces a data set of the spots that have been collected from the gels in the study, so they may be analyzed by a relational database.

Figure 3. 2DE Color Overlay
THE PURPOSE OF THIS DISSERTATION

The purpose of this study is to evaluate the usefulness and efficacy of the RLGS technique and subsequent image analysis in detecting and analyzing genetic changes in a number of cancer models, including 1) genetic changes found in rat mammary methylnitrosourea (MNU)-induced tumors which were also treated with 4HPR and without 4HPR administration; 2) genetic aberrations noted in cervical human intraepithelial neoplasia, cervical squamous cell carcinomas, and cervical adenocarcinomas; 3) genomic changes found in K-Ras Oncogene Transformed NIH-3T3 Cells; and 4) genetic differences noted between human bronchial epithelial cells irradiated with high-LET alpha particles and a control group. Two of the studies employed a novel qualitative image processing system to prescreen the RLGS images for genetic anomalies. The other two studies in this dissertation employed a quantitative spot matching program donated by the Scanalytics Corp. and the NCI.

1) The Reduction of Genetic Alterations by 4-Hydroxyphenylretinamide in Methylnitrosourea Induced Rat Mammary Tumors
Although the rat mammary tumor model has been used extensively, the understanding of Methylnitrosourea chemical induction and its genetic affects are still poorly understood at a genomic level. The purpose of this study is to evaluate effects of Methylnitrosourea on genomic instability and to evaluate the chenopreventative effect of 4-Hydroxyphenylretinamide on Methylnitrosourea treated tissue using the RLGS image analysis enhanced technique.

2) Detection of Genomic Alterations in Human Cervical Cancer by Two-Dimensional Gel Electrophoresis

Cervical cancer is one of the leading killers in females and its genesis is a complicated concert of genetic alterations. The purpose of this study is to use RLGS quantitative image analysis to describe the genetic differences between CIN invasive cervical cancer and normal cervical tissue.

3) Computer Assisted Analysis Of Restriction Landmark Genomic Scanning Images Of DNA From K-Ras Oncogene Transformed NIH-3T3 Cells
The idea behind this study is to analyze the genetic alterations in K-ras transformed NIH3-3T3 cells lines by RLGS and to evaluate the efficacy of the GELLAB II+ and SQL relational database software in the processing and data extraction of RLGS spot patterns.

4) Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Human Bronchial Epithelial Cells radiated with High-LET Alpha Particles

Understanding the effects of high-LET radiation bronchial epithelial cells is important in elucidating how radon gas induces malignant transformation in that cell set. The purpose of this study is to investigate the genomic changes found in human bronchial epithelial cells irradiated with high-LET particles using RLGS. The subsequent analysis will utilize the stand-alone GELLAB-II+. Along with our standard analysis software and we will implement a novel Java base system called Webgel, which is accessible over the Internet.
MANUSCRIPT I

The Reduction of Genetic Alterations by 4-Hydroxyphenylretinamide in Methylnitrosourea Induced Rat Mammary Tumors

Jiafan Liu¹, Yian Wang¹, James Patrick¹, Keith A. Crist¹, Vernon E. Steele², Gary J. Kelloff², Charles W. Boone², Ronald A. Lubet², and Ming You¹,³

Department of Pathology, Medical College of Ohio
Toledo, Ohio 43699

Running Title: Genetic Alterations in Rat Mammary Tumors.

Key words: 2-DGE, RLGS, 4HPR, mammary tumor, chemoprevention, MNU, genetic alteration

This document appears in the dissertation of Jeff Liu, due to it’s relevance to this author’s work it has been included in this dissertation.

Gel runs preformed by Jeff Liu.
Image processing and analysis preformed by James Patrick.
Footnotes

1Medical College of Ohio, Toledo, Ohio 43614.
2National Cancer Institute, Bethesda, Maryland 20892
3To whom requests for information should be addressed.

Abbreviations: 2-DGE, two-dimensional gel electrophoresis; RLGS, restriction landmark genomic scanning; 4HPR, 4-hydroxyphenyl-retinamide; MNU, methylnitrosourea; LOH, loss of heterozygosity; RAR, retinoic acid receptor.
Abstract

Two-dimensional gel electrophoresis (2-DGE) was used to comprehensively scan the whole genome of normal rat mammary tissues and methylnitrosourea (MNU)-induced rat mammary tumors with and without 4HPR administration for various genetic alterations, such as DNA amplification, homozygous deletion, loss of heterozygosity (LOH), chromosome translocation and altered DNA methylation. Computer-assisted systems were applied for spot detection, quantitation, and comparison of 2-DGE DNA patterns. By comparing normal and tumor 2-DGE profiles, 12 novel spots were detected only in mammary tumors, 10 spots were absent from tumor gel profile, 8 spots in tumor lost about half of their intensity suggesting genomic deletions, and 5 spots in tumor were found amplified. Totally, 35 reproducible altered spots on rat mammary tumors 2-DGE profiles were detected. Among them, 8 spots had significantly suppressed alterations, 12 spots were lowered by 30-40% of their incidence because of 4HPR administration. These genetic alterations may represent novel changes in cancer-related genes that are central to mammary tumorigenesis. Characterizations of the 4HPR responsive alterations are not only significant to the understanding of 4HPR chemopreventive mechanisms but also to
the searching for surrogate endpoint biomarkers for chemoprevention trial evaluation.
Introduction

Retinoids, vitamin A analogs, have shown promising effects as both prevention and therapeutic agents against malignancy in experimental models (Bollag et al., 1992). As one of the structurally modified retinoids, 4HPR has been found to be effective and less toxic than Vitamin A itself and is being assessed as a chemopreventive agent in various experimental and clinical trials (Lewis et al., 1994). 4HPR is effective in preventing carcinogen-induced cancers in the breast, urinary bladder, skin, lung and prostate. (Naik et al., 1995).

Retinoids are the most commonly used chemoprevention agents but their mechanism of chemopreventive action is not fully understood. It is hypothesized that retinoids modulate gene expression via interaction with nuclear retinoic acid receptors, thereby affecting cellular differentiation and suppressing progression of preneoplastic cells to cause neoplastic lesions (Sheikh et al., 1995). In target cells, retinoic acid is the form of retinoids that exerts actions. The nuclear retinoic acid receptors (RAR) appear to be the most important mediators of retinoid actions. RARα, RARβ, and RARγ act as ligand-dependent transcription factors, which bind to specific DNA sequences. When retinoic acid binds to
the receptor, a conformational change in the RAR occurs followed by formation of transcriptional initiation complexes (McBurney et al., 1993). In most instances retinoic acid induces gene transcription, but inhibition of transcription can also occur. Feldman and Foster (1979) reported a selective inhibition of DNA synthesis in preneoplastic and neoplastic mammary cells by retinoid, without any significant effect on normal mammary epithelial DNA synthesis.

Early studies in rats after oral administration of 4HPR indicated selective accumulation of 4HPR and its metabolites in the mammary glands in a dose-dependant manner, without any detectable rise in hepatic retinoid levels (Moon et al., 1979). It is very important for the practical use of 4HPR because hepatic accumulation and subsequent toxicity precludes prolonged administration of many other retinoids. In addition, 4HPR exhibits a longer half-life than many other retinoids, and achieves higher tissue concentrations relative to plasma (Formell et al., 1993).

Oral 4HPR reduced the overall incidence of breast cancer and increased the latency to cancer development in carcinogen-treated rats (Moon et al., 1979). Although it was not as potent as retinyl acetate, the toxicity was significantly less. Histological examination demonstrated that there was decreased ductal branching and end bud
proliferation in 4HPR-treated rats compared to rats receiving no retinoid (Naik et al., 1995). 4HPR was also reported to suppress the development of recurrent mammary cancer after the removal of carcinogen-induced first palpable tumor. Dowlatshahi et al. (1989) observed complete regression of MNU-induced first mammary tumors in 22% of rats and partial regression in 19% of rats which received a therapeutic dose of 4HPR. Additionally, 4HPR showed synergism with ovariectomy and tamoxifen in inhibition of carcinogen-induced tumors.

In this study, MNU-induced rat mammary tumors with and without 4HPR treatment were provided by Dr. Keith A. Crist for collaboration. The tumor multiplicity was found reduced about 50%, while no significant reduction on tumor incidence was observed. Although rat mammary tumor model has been used extensively, there have been relatively few reports on the genetic alterations underlying the chemical induction of these tumors. G-->A transition mutation at codon 12 of H-ras protooncogene was found in a majority of the MNU-induced rat mammary tumors (Sukumar and Barbacid, 1990) and PRAD-1, IGF2, H-ras amplifications were observed in advanced mammary tumors (Young et al., 1996). Cytogenetic analyses of MNU-induced rat mammary tumors indicate that rat chromosome 1q22 is a breakpoint for deletions and translocations, and LOH was also
frequently found at the distally located INS-1 Locus (region 1q55). There is a possibility of putative tumor suppressor gene (Gollahon et al., 1993). There is also no report about molecular changes associated with 4HPR administration in rat mammary tumor model or any other models. Changes in genetic alterations caused by 4HPR underlying chemical modulation mechanisms in its chemopreventive actions are not known. In order to achieve better understanding of the mechanism of the mammary carcinogenesis process, and to better evaluate the 4HPR chemoprevention trial, it is necessary to investigate the molecular events involved in these actions. These molecular events may be used as intermediate endpoint biomarkers which is essential in the design and evaluation of chemoprevention trials.

2-DGE is a new method to scan for genetic alterations on a whole genome level (Hatada et al., 1991). 2-DGE allows simultaneous examination of the entire genome of a given cell line or tumor in a single experiment. In this study, we investigate the rat mammary tumor model with and without 4HPR administration for genetic alterations. The molecular events were detected by the 2-DGE system and analyzed statistically in the evaluation of the chemopreventive efficacy of 4HPR.

**Material and Methods**
Tumor Introduction:

Female Sprague-Dawley rats were received at 35 days of age, acclimated for 3 days then fed the 4% Teklad control diet. Rats were then randomized into treatment groups (20 rats for each group). MNU/4HPR group animals were fed the 4HPR (786 mg/kg diet) beginning at 41 days of age, 1 wk prior to MNU injection. All the rats in both MNU/4HPR and MNU only groups received an intrajugular injection of MNU (50 mg/kg, pH 5.0, 12.5 mg/ml) prepared immediately before use.

DNA Isolation:

About 50-75 mg tissue for each sample was pulverized and digested in proteinase K digestion buffer. After 3 times phenol extraction, DNA was precipitated with ethanol and dissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). High-molecular-weight DNA is essential for good quality of 2-DGE. DNA breaking was minimized by careful handling and manipulating of the DNA in this procedure.

2-DGE Procedure:

(1) NotI digestion:

About 10 μg DNA was digested in NotI digestion buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 100 mM NaCl) containing
100 units NotI for at least 2 hr in 37°C. We usually do it overnight.

(2) NotI cut site labeling:

The sticky ends (====GGCC) generated from NotI digestion were filled with [32P] α-dGTP and [32P] α-dCTP. The reaction buffer contained 40 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, 50 mM NaCl, 10 mM DTT, 2 μl each of [32P] α-dGTP (6000 Ci/mmol) and [32P] α-dCTP (6000 Ci/mmol), 20 units sequenase ver 2.0. This reaction took 20 min in 37°C, followed by 5 min 65°C incubation.

(3) EcoRV digestion:

DNA was digested with EcoRV for 2 hr in 37°C. The reaction buffer contained 100 units EcoRV, 50 mM Tris (pH 8.0), 10 mM MgCl₂, 50 mM NaCl.

(4) First dimensional (1st-D) electrophoresis:

The 1st-D electrophoresis was carried out in a vertical apparatus which had 5 mm inner diameter glass tubes connecting upper and lower buffer tanks. These tubes were 60 cm long and filled with Seakem Gold agarose gel (FMC Bio product), the top 2 cm of the 1st-D gel piece (stacking gel) was 0.5% agarose and 20% sucrose while the remainder of the gel (running gel) was 0.8% agarose and 5% sucrose. The 1st-D gels were prepared and run in 1x1st-D buffer. (10x1st-D
buffer preparation: dissolve 242 g Tris, 109 g sodium acetate trihydrate, 42 g NaCl, 23.4 g EDTANa$_2$ in 1.8 l of H$_2$O and adjust the pH to 8.15 with acetic acid. Adjust the volume to 2 l and send through a 0.2 μm filter. After EcoRV digestion, DNA was loaded on the top of 1st-D gel and electrophoresed in 100 volts for 2 hr followed by 230 volts for 22 hr.

(5) In situ digestion with HindIII:

1st-D gel was expelled from the glass tube and trimmed to a 30 cm long piece (so called “noodle”) which contained most of the DNA. After 30 min equilibration in HindIII reaction buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl$_2$, 1 mM DTT, pH 7.9), the “noodle” was aspirated into a 6 mm inner diameter plastic tube, then filled with HindIII reaction solution containing 2000 units HindIII. The tube was incubated in 37°C over night.

(6) Second dimensional (2nd-D) electrophoresis:

After HindIII digestion, the “noodle” was equilibrated in 1xTBE buffer and loaded on 0.8% horizontal ultra pure agarose gel. 2nd-D electrophoresis was carried out in 45 volts for 50 hr. Then the agarose gel was dried at 60°C under vacuum and autoradiographed.

Computer-assisted analysis of 2-DGE:
A Bio-Rad scanner converted the gel autoradiograph into a digitized image with a spatial resolution of 400 microns. Analysis was performed on a Macintosh 840av computer using the public domain NIH Image program (written by Wayne Rasband at the U.S. National Institutes of Health and available from the Internet by anonymous ftp from zippy.nimh.nih.gov). A color hybridization macro was used to register and align spot landmarks of two gels. One gel was assigned a green LUT (look up table) and the other a red LUT then the gel images were superimposed. The color addition of the two gels would produce a profile of yellow spot color if the gel were similar in spot intensity, green if the spot in gel “Two” was greater in intensity than gel “One” and red if the spot in gel “One” was greater in intensity than gel “Two”. This computer color overlay system produced a color shift base on matched spot intensities between two gels allowing a qualitative way to observe spot intensity differences.

A pair of images from tumor (red) and normal (green) respectively were matched each time. After carefully correcting the spot fitting of whole images, all spots were compared one with the other and showed yellow in color when they were unchanged. Unmatched spots stood out showing either green or red or displayed a various color shift in between depending on which genetic alterations they represented.
After the discovery of the altered spots, a densitometer was used to determine the radioactive intensity of the spots individually on their original autoradiographs (X-ray films). The intensities of 10 regular spots surrounding each altered spot were measured. The mean of these 10 intensities was taken as 1 unit and used as a reference in the determination of the intensity of their corresponding altered spot. The readings were analyzed by one way ANOVA and Bonferroni post hoc. test.

Results

Eighteen MNU-induced primary rat mammary tumors (10 from the MNU only group, 8 from the group with 4HPR preadministration) were analyzed by 2-DGE. Histological examinations of these mammary tumors revealed that all of the tumors were mammary adenocarcinomas. Seven normal rat mammary samples were also collected. Genomic DNAs were isolated from these tumor and normal samples, and their quality is evaluated by agarose gel electrophoresis. Figure 1. shows a typical 2-DGE profile.

After computer-assisted analysis of the genetic alterations by color overlay comparison of tumor and normal 2-DGE profile (Figure 2.), 12 novel spots were detected only in mammary tumors, 10 spots were absent from the tumor gel
profile, 8 spots in the tumor lost about half of their intensity suggesting genomic deletions, and 5 spots in the tumor were found amplified. By this approach, we detected about 35 reproducible altered spots on 2-DGE profiles of rat mammary tumors as compared with 2-DGE profile of normal mammary tissue.

In the tumor group with 4HPR preadministration, reduction of genetic alteration was observed in most of the altered spots in corresponding extent. After statistic analysis, 6 spots (I-8, I-9, I-14, II-1, II-5, III-5, all new spots) were found to have lost their significance in intensity difference between the normal group and the 4HPR/MNU group, while their intensities in the normal group were still significantly different from those in the MNU only group; the alterations in two spots (one new spot I-6, one missing spot III-9) were found significantly repressed in the 4HPR/MNU group even though the differences between the normal group and the 4HPR/MNU group were still significant. In the 4HPR/MNU group, spot I-6 possessed weaker intensity than that in the MNU group, spot III-9 remained less than half intensity and lost everything in the MNU group.

Other than intensity alteration reduction, the incidences of these altered spots in the 4HPR/MNU group were found reduced in an even wider spectrum. Twelve (spot...
I-1, I-6, I-10, I-11, I-12, II-5, II-7, III-3, III-6, III-7, III-8, III-10) out of a total of 35 spots reduced 30%-40% of their appearance rates in 4HPR/MNU group comparing those in MNU only group.

Discussion

MNU is a carcinogenic agent able to produce different types of tumors depending on the species and administration routes. In our experiment, mammary tumor is the major lesion caused by MNU. At least 35 altered spots were found on the tumor 2-DGE profiles compared to the counterpart of normal. It is the first time we detected so many genetic damages in the MNU-induced rat mammary tumor model. These alterations include spots with increased or reduced intensity, new spots (appeared in tumor only), and absent spots (totally lost in tumor). The amplified spots could be caused by the amplification of DNA fragment. The new spots could come from the demethylation of certain CpG island or chromosome translocation. CpG island methylation correlates positively with the repression of genes. New spots in tumor could be considered as the signal of certain genes involved in cell growth which are in active status. The spots which lost their intensity in the tumor are most likely representing tumor suppressor genes. The reduction of copy numbers of
tumor suppressor gene will loose the restriction of cell
growth in corresponding degree. Malignancy progression will
be promoted.

Carcinogenesis is now regarded as a complicated process
in which aberrant changes in oncogenes, tumor suppressor
genes and tumor susceptibility genes accumulate and finally
trigger the formation of tumor cells. The aberrant changes
could be amplification, homozygous deletion, loss of
heterozygosity, altered DNA methylation, or chromosome
translocation, etc. The accumulation of such genetic events
is involved in tumor cell uncontrolled growth and malignancy
progression. Amplification could be a tandem repeat which
results from several rounds of extra DNA synthesis.
Excessive expression of genes for growth factors, receptors
or intracellular signal transduction components are thought
to promote excessive growth of the malignant cells. While if
there is a loss of tumor suppressor genes, the restriction of
cell growth could be set free. Therefore, carcinogenesis and
malignancy progression all basically depend on the variation
of copy number and structure of each important growth control
gene.

In 4HPR chemopreventive treatment group, the repression
of genetic alterations are observed in most of the spots.
From Table 1, we can see that almost every intensity reading
in 4HPR group is somewhere between normal and MNU group. No matter which kind of genetic alteration it is, 4HPR could always turn it around in the orientation of normal. That is suggesting that 4HPR exerts its chemopreventive function in a relatively general and fundamental manner. In several spots (e.g. I-3, I-4, I-5), 4HPR dose not seem to bring about any change to their corresponding alterations. While in spots I-8, I-9, I-14, II-1, II-5, III-5, the genetic alterations in 4HPR/MNU group have been reduced to the level that are not significantly different from those of normal. In spots I-6, III-9, the changes brought by 4HPR become statistically significant from tumors without 4HPR treatment. Their differential sensitivities to 4HPR could suggest their roles in 4HPR chemoprevention. The several most sensitive spots which were induced by MNU and repressed by 4HPR could be used as surrogate endpoint biomarkers (SEB) in chemoprevention trials evaluation. By conventional approaches, to select valid molecular biomarkers is a very complicated and difficult job. 2-DGE has greatly simplified the search for genetic alterations at the whole genome level.

Several techniques have been developed to detect and identify the genetic alterations in tumor cells. Usually probes or primers are required for these techniques and methods. Mammalian genome is about as big as $3 \times 10^9$ base
pair. Analysis of many loci throughout the genome is too laborious and time consuming. Moreover, none of these methods are sensitive enough to demonstrate DNA modulations by chemicals at whole genome level. We used 2-DGE on the basis of the approach described by Hatada (1991). This approach provides a global analysis of gains and losses of genetic material in the genomic DNA of solid tumor as well as a cell line that would otherwise be unanalyzable by conventional cytogenetic techniques. Each spot on the profile represents the copy number of certain fragment on the gel. Combined with computer analysis, we can compare normal and tumor profiles and immediately pick up the altered spots in tumors. We can easily narrow our search on several candidate spots which represent recurrent alterations and never rely on any probes and primers. CpG island can be seen in the 5’ upstream region of nearly all housekeeping gene and some of the tissue-specific genes. In mammalian, 89% NotI sites lie in CpG islands. Therefore, most of the spots on the profile reflect the pieces of fragment in the vicinity of genes. The size of each spot is about 500 bp to 10 kbp, which is applicable for cloning and sequencing work later on.

Since several spots were sensitively affected by 4HPR treatment, these spots must represent the genetic alterations on which 4HPR exerts its impact. Cloning and characterization
of these spots not only can help us know the mechanism of 4HPR chemopreventive function, but also can improve our understanding of the mammary carcinogenesis process.
Reference


restriction landmark genomic scanning (RLGS) method.
Nucleic Acids Research, 21(24):5604-5608


Sheikh MS, Shao ZM, Li XS, Ordonez JV, Conley BA, Wu S, Dawson MI, Han QX, Chao WR, Quick T (1995): N-(4-
hydroxyphenyl)retinamide (4HPR)-mediated biological actions involve retinoid receptor-independent pathways in human breast carcinoma. Carcinogenesis. 16(10):2477-86. Oct


Legend

Figure 1. A typical 2-DGE profile.
Figure 2. Color Overlay System.
Figure 1. A typical 2-DGE profile.
Figure 2. Color Overlay System
MANUSCRIPT II

Detection of Genomic Alterations in Human Cervical Cancer
by Two-Dimensional Gel Electrophoresis

Jiafan Liu, MD1, Yian Wang, MD, PhD1, Ping Gu, MD1, James Patrick, BS/BA1, Keith A. Crist, PhD1, Carol Sabourin, PhD2, Gary D. Stoner, PhD2, Michele F. Mitchell, MD3, James D. Fanning, MD1, Kitai Kim, MD1, Peter J. Goldblatt, MD1, Gary J. Kelloff, MD4, Charles W. Boone, MD, PhD4, and Ming You, MD, PhD1

1Medical College of Ohio, Toledo, OH 43614
2Ohio State University, Columbus, OH 43210
3University of Texas M.D. Anderson Cancer Center, Houston, TX 77030
4National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

This manuscript was published in Journal of Cellular Biochemistry 25s:41-48, 1996

Running Title: Genomic alterations in human cervical cancer
Footnote:
Address correspondence to Ming You, MD, PhD, Department of Pathology, Medical College of Ohio, Health Education Building, Room 263, 3000 Arlington Avenue, Toledo, OH 43614.

Key Words: Two-Dimensional Gel Electrophoresis, Cervical Cancer, Genomic Alterations, Genomic Scanning, Chemoprevention

This document appears in the dissertation of Jeff Liu, due to its relevance to this author’s work it has been included in this dissertation.

Gel runs preformed by Jeff Liu
Image processing and analysis preformed by James Patrick
Abstract

Two-dimensional gel electrophoresis was used to comprehensively scan the whole genome of 6 cervical intraepithelial neoplasia (CIN) lesions, 7 cervical squamous cell carcinomas, 1 cervical adenosquamous cell carcinoma, and 2 cervical adenocarcinomas for multiple genetic alterations, such as DNA amplification, chromosome deletion, loss of heterozygosity, and chromosome translocation, as compared with the paired normal tissues. DNA spot analysis of the genomic 2-dimensional gels was performed by a computer color overlay system and by spot recognition software allowing for objective spot comparison and quantitation. Nine spots were found to be amplified in the cervical carcinomas while two amplified spots were detected in the CIN III lesions. Twelve DNA spots were either reduced in their intensity or absent in cervical carcinomas as compared to their normal paired tissues. Reduction of intensity in 6 spots was observed in the 5 CIN III lesions. These genetic alterations may represent changes in cancer genes that are associated with human cervical carcinogenesis. Further characterization of these alterations may be significant to the understanding of cervical tumorigenesis and to the development of biomarkers for clinical trials in cancer chemoprevention.
Introduction

Cervical cancer is one of the leading causes of cancer death in women (1,2). There are three major types of cervical cancer: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma (3). Preinvasive intraepithelial cervical lesions, also termed cervical intraepithelial neoplasia (CIN), are characterized by abnormal cellular maturation and proliferation, and atypical nuclei in the epithelium (3). The CIN lesions themselves and genetic alterations found in CIN lesions may provide a suitable target for intervention with chemopreventive agents (4).

Human papillomavirus (HPV) infection has been associated with the pathogenesis of cervical cancer but other genetic alterations are required for the development of malignancy (5, 6). For example, activation of proto-oncogenes (erb-B, myc, bcl-1, ras) (7-11) and inactivation of tumor suppressor genes (p53 and allelic loss at 3p, 11p, 17p) have been detected in human cervical carcinomas (12-19). Most high grade cervical squamous cell cancers contain HPV types 16, 18, 31, or 33 suggesting that HPV plays a causal role in the induction of cervical cancer (5, 6). The genetic basis for this stems from the observation that the tumor suppressor gene, p53, can be inactivated by HPV E6.
protein, and another tumor suppressor gene, Rb, can be inactivated by HPV E7 protein (20,21). HPV alone, however, is not sufficient for the induction of invasive cervical carcinoma since only a few of the HPV infected patients develop cancer (22). Increasing evidence indicates that other genetic alterations are required for the development of invasive cervical carcinoma (7-19).

Increased ras p21 expression was detected in 18% of CIN I, 29% of CIN II, and >50% of CIN III lesions and invasive cervical carcinomas (23,24), indicating that ras activation is an early step in cervical carcinogenesis. In addition, several oncogenes are amplified in cervical cancers. For example, erb B, myc, and PRAD 1 (cyclin D/Bcl-1) were frequently amplified in human cervical cancer and/or their cell lines (7-11). In some cervical tumors, Myc L1, H-ras, Sea, and Gli are amplified (7). Thus, activation of proto-oncogenes seems to play an important role in cervical tumorigenesis.

Inactivation of the tumor suppressor gene p53 is observed principally in HPV-negative cervical cancers since HPV E6 and E7 would inactivate p53 and Rb in HPV-positive tumors (12-15,20,21). In addition to the p53 gene, loss of heterozygosity (LOH) of chromosomes 1, 3p, 11p and 17q were also detected (16-19), suggesting that multiple putative
tumor suppressor genes are involved in the progression of cervical cancer. Identification of these novel tumor suppressor genes and detection of additional genomic alternations in cervical cancer is significant for understanding the genetic basis of malignant transformation in the cervix.

In this study, a newly developed two-dimensional gel electrophoresis (2-DGE) assay was used to scan the entire genome for DNA alterations in CIN lesions and invasive cervical cancers. This method involves cleaving of high molecular weight DNA, radioactive labeling and separating DNA fragments by 2-DGE (25,26). By comparing the 2-DGE profile of tumor tissues with their normal tissue counterparts, genetic alterations such as amplification, allelic loss, and hyper/hypomethylation can be detected.
Materials and Methods

Sample collection: Ten fresh cervical carcinomas (7 squamous cell carcinomas, 2 adenocarcinomas and 1 adenosquamous carcinoma) and their normal surrounding tissues, plus 5 CIN III lesions and 1 CIN I-II lesion were collected from the Medical College of Ohio, the University of Texas M.D. Anderson Cancer Center, and the Ohio State University. The histopathology of all lesions was assessed by board-certified pathologists in each institution.

DNA Isolation: High molecular weight DNA is essential for high quality 2-DGE. The general strategy for the 2-DGE has been described previously (25). DNA strand breakage is minimized by careful isolation, storage and manipulation of DNA in this procedure. Guided by histopathologic evaluation of thin tissue sections, the appropriate area of frozen thick tissue was removed from the slide and placed immediately in digestion buffer containing proteinase K. Following triple phenol extraction, DNA was precipitated with ethanol and dissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). The quantity and purity of the DNA was measured in a spectrophotometer at wavelengths of 260 nm/280 nm and the quality of the DNA was checked by electrophoresis in an 0.8% agarose gel.
Not I digestion: About 2-10 µg of DNA was digested in 100 µl Not I (New England Biolabs, Beverly, MA) digestion buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 100 mM NaCl) containing 100 U of Not I for a minimum of 2 hrs at 37°C.

Not I site labeling: The sticky ends generated from NotI restriction digests were filled with [³²P]α-dGTP and [³²P]α-dCTP. The reaction buffer contained 40 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, 50 mM NaCl, 10 mM DTT, 2 µl each of [³²P]α-dGTP (6000 Ci/mmol) and [³²P]α-dCTP (6000 Ci/mmol) and 20 units Sequenase (version 2.0, United States Biochemicals, Cleveland, OH). The reaction was allowed to proceed for 20 min at 37°C, and terminated by incubation at 65°C for 5 min.

EcoR V digestion: DNA was then digested with EcoR V (New England Biolabs, Beverly, MA) for 2 hr at 37°C. The reaction buffer contained 100 units of EcoR V, 50 mM Tris (pH 8.0), 10 mM MgCl₂ and 50 mM NaCl.

First-Dimensional (1st-D) gel electrophoresis: First-dimensional electrophoresis was carried out in a specially
designed vertical apparatus which contained 5 mm (inner diameter) glass tubes connecting the upper and lower buffer tanks. The tubes (60 cm long) were filled with Seakem Gold agarose gel (FMC Bio product). The top 2 cm of the 1st-D gel piece (stacking gel) consisted of 0.5% agarose and 20% sucrose while the remainder of the gel (running gel) consisted of 0.8% agarose and 5% sucrose as previously described (25). After EcoR V digestion, the DNA was loaded onto the top of the 1st-D gel and electrophoresed at 100 V for 2 hrs, followed by 230 V for 22 hr.

In situ digestion with Hind III: The 1st-D gel was expelled from the glass tube and trimmed to a 30 cm long piece (so called "noodle") which contained most of the DNA. After 30 min equilibration in Hind III reaction buffer, the "noodle" was aspirated into a 6 mm inner diameter plastic tube and filled with Hind III reaction solution containing 2000 units of Hind III. The tube was incubated at 37°C overnight.

Second-dimensional (2nd-D) gel electrophoresis: After Hind III digestion, the "noodle" was equilibrated in 1x TBE buffer and loaded onto a 0.8% horizontal ultra pure agarose gel, and electrophoresed at 45 V for 50 hr. The agarose gel was then dried at 60°C and autoradiographed.
Scanning and color overlay system: A Bio-Rad scanner was used to convert the gel autoradiographs into digitized images with a spatial resolution of 400 microns. Analyses were performed on a Macintosh 840av computer using the public domain NIH Image program (written by Dr. Wayne Rasband at the U.S. National Institutes of Health and available from the Internet by anonymous ftp from zippy.nimh.nih.gov). A color hybridization Macro was used to register and align spot landmarks of normal and tumor tissues. One gel, usually the reference normal, was assigned a green look up table (LUT) and the other, usually the tumor, a red LUT. Gel images were then superimposed. The color of the two gels produced a profile of yellow spots at regions of similar spot intensity, green if the corresponding spot on the tumor gel was lower in intensity or absent and red if a novel or amplified spot was present on the tumor gel. The color shift, based on relative intensity of matched spots between the two gels, allowed a qualitative means for evaluation of spot intensity differences.

Scanning and spot detection: The 2-dimensional gel analysis software from Applied Imaging, Inc. was used for computer
assisted analysis of relative spot density between reference normal tissue and tumor tissue gels.

Gel matching: After monitoring and correcting the spot detection of all images, reference normal and tumor DNA images were matched. Automatic spot matching makes use of landmark references, usually ribosomal DNA, identified by the operator to "warp" spot coordinates within defined limits and identify matches between the two gels. Warping was a necessary analysis feature as gels set to run under "identical" electrophoretic conditions never yielded exact duplication over the entire surface. By iterative matching and landmark identification steps, all spots were identified as matched or novel fragments. After matching, the image file contained density information on every DNA spot for both normal and tumor gels for quantitative comparison.

Results

A total of 7 squamous cell carcinomas, 2 adenocarcinomas, and 1 adenosquamous carcinoma were analyzed by 2-DGE for genomic alterations when compared to normal cervical tissues from the same patient. There are more than 1,500 spots in each of the 2-D profiles of either cervical carcinomas or their paired normal cervical
tissues. After color-overlay analysis of each pair of 2-D gels, altered spots were identified and quantified by using computer-assisted systems. Upon screening, 9 spots were found amplified in 10 cervical carcinomas. Spot 1 was amplified 2 to 5-fold (100% of the cases); spot 2, 2 to 4-fold (100%); spot 3, 2 to 8-fold (20%); spot 4, 8-fold (10%); spot 5, 4-fold (10%); spot 6, 20-fold (10%); spot 7, 5-fold (10%); spot 8, 2-fold (30%); and spot 9, 20-fold (10%). Amplification of spots 3-9 was seen only in squamous cell carcinomas (patients 1-7). Spot 1 was also amplified in 1 of the 6 CIN lesions analyzed and spot 2 in 3 of the 6 CIN lesions, suggesting that amplification of spots 1 and 2 may be early events in cervical carcinogenesis.

Five spots were found reduced in their intensity in cervical carcinomas. Spot 10 was reduced 50%-90% (90% of tumors), spot 11, 50%-90% (90%), spot 12, 50%-75% (80%), spot 13, 50%-80% (50%), and spot 14, 50%-60% (80%). The reduction in these spots may be attributed to either hypermethylation of DNA, or deletions and rearrangements of the genome in the tumor tissues (27-30).

Nine spots were completely lost in cervical cancer: spot 15 (60% of tumors), spot 16 (20%), spot 17 (13%), spot 18 (38%), spot 19 (75%), spot 20 (50%), spot 21 (50%), spot
22 (40%), and spot 23 (60%). Seven additional spots were lost in CIN lesions: spot 24 (50%), spot 25 (50%), spot 26 (83%), spot 27 (17%), spot 28 (60%), spot 29 (33%), and spot 30 (17%). Hypermethylation or homozygous deletion may be responsible for the observed spot loss (27-30).

**Discussion**

The results from this study, together with other recent reports (25-30), indicate that 2-DGE analysis of DNA is a very powerful and useful technique for detection of genomic alterations. This methodology can be used for high-speed survey for the presence or absence of restriction landmarks (or spots) throughout a genome and for measurement of their copy number in each locus, based on the principle that restriction endonuclease sites can be used as landmarks (25-30). Since the total length of human genomic DNA is approximately $3 \times 10^9$ bp and around 1500 spots can be revealed on one 2-DGE profile, 2-DGE can scan the restriction landmarks at intervals of less than 0.5 to 1 megabase (Mb). One spot on the 2-DGE profile corresponds to a single locus (26). By using 2-DGE assisted by color-overlay analysis, we have shown that 9 spots were amplified in the cervical carcinomas and 2 were amplified in CIN lesions. In contrast, 5 spots were reduced in intensity and nine were lost in tumors when compared to normal paired
tissues. The determination of the genes involved in these spots may provide insights into the mechanism(s) of cervical cancer development.

Some of the altered spots may represent genetic changes that have been observed previously. For example, gene amplifications previously detected in cervical cancer should be considered as primary candidates for the amplified spots found in the 2-DGE analysis. Amplification of HPV 16 and 18 were observed in some of the HPV-positive cervical carcinomas (5,6). In addition, amplified proto-oncogenes such as erb B, myc, PRAD 1, Myc L1, H-ras, Sea, and Gli were also detected in cervical carcinoma or their cell lines (7-11). The loss of spots probably results from: 1) homozygous deletions, since all lost spots were based upon a >50% reduction in intensity; 2) DNA rearrangements that could result in new spots and, simultaneously, the loss of spots in tumor tissues; and 3) hypermethylation, which could abolish the Not I site in tumor DNA when compared to paired normal tissue. Candidate genes for the lost spots could include the previously observed allelic loss of chromosomes 1, 3p, 11p and 17q (12-19), and those tumor suppressor genes known to be inactivated by hypermethylation, i.e. pl6/CDKN2/MTS1 and VHL genes (32-34).
The detection of DNA methylation that may inactivate a tumor suppressor gene(s) in tumor tissue is a major focus of cancer research (32-34). Cytosine moieties at CpG sites in the mammalian genome are heavily methylated at the 5'-position (35). Although the exact function of DNA methylation is unknown, current evidence suggests a role for DNA methylation in transcriptional control of gene expression (36). There are approximately 6,000 Not I sites (Not I is a methylation-sensitive enzyme) per mammalian haploid genome and 90% of these are located in CpG islands (37) associated with transcriptional units (35,38). In one 2-DGE gel, 3,000 spots can be seen that correspond to at least 1,500 different Not I sites (25% of total Not I sites). Thus, 2-DGE would appear to be the method of choice for analysis of methylation status in tumor tissues (39-42).

Some of the altered spots in cervical carcinomas and CIN lesions could be genetic markers that have not been previously identified. These novel changes can be further characterized by spot target cloning techniques (40,43-44). The cloned DNA fragments will be used either as probes for Southern blot analysis to confirm the alterations, or as probes for screening cDNA library in order to determine whole cDNA sequences.
A major effort in cancer chemoprevention is to develop surrogate endpoint biomarkers for clinical trials in cancer chemoprevention (45). 2-DGE analysis of preneoplastic and neoplastic lesions could provide new molecular biomarkers for the carcinogenesis process in specific organ sites. Furthermore, chemopreventive agent-specific biomarkers might be identified in tumor tissues treated with the agent. Recently, we used 2-DGE to scan genetic alterations in normal rat mammary tissues, and in methylnitrosourea (MNU)-, MNU/dehydroepiandrosterone-, and MNU/N-(4-hydroxyphenyl)retinamide-induced rat mammary tumors (31). More than 50 altered spots were seen in MNU-induced tumors while 20% to 30% fewer altered spots were observed in MNU/DHEA- and MNU/4-HPR-induced tumors (31). Those spots suppressed by the two chemopreventive agents may be chemopreventive agent-specific markers in the rat mammary tumor model. Similarly, genomic alterations detected by 2-DGE in cervical lesions, especially those found in CIN III lesions including the amplifications of spot 1 & 2 and the losses of spot 24 to 29, are strong candidates for surrogate endpoint biomarkers in human cervical cancer chemoprevention trials. Limitations of 2-DGE analysis of DNA include: 1) there is no absolute correlation between altered spot in tumors with altered expression of mRNA or
protein; and 2) this technique is more suited for the detection of genotypic changes rather than phenotypic changes.
Acknowledgments

We thank Dr. Herman A.J. Schut for his critical reading of the manuscript. We are indebted to Ms. Kathy Deanda for art work and Sophia Grecos for her assistance in collecting some of the cervical tissues. This work was supported by NCI CN-25486-02.
References


Computer Assisted Analysis Of Restriction Landmark Genomic Scanning Images Of DNA From K-Ras Oncogene Transformed NIH-3T3 Cells

Patrick, James L., Gu, Ping., Liu, Jiafan., Lemkin, Peter., You, Ming.

'Medical College of Ohio, Toledo, OH 43614  2 Laboratory of Experimental and Computational Biology, National Cancer Institute/FCRF, Frederick, MD

Running Title: Genomic Analysis of K-ras Transformed Cells using GELLAB II+

To whom requests for reprints should be sent. Address correspondence to Ming You, MD, Ph.D., Department of Pathology, Medical College of Ohio, Health Education Building Rm. 206, 3000 Arlington Avenue, Toledo, OH 43699.

Abbreviations: RLGS, restriction landmark genomic scanning; 2-DGE, Two-dimensional gel electrophoresis
Abstract

Previous studies have shown that restriction landmark genomic scanning (RLGS) is an effective tool in the detection of genetic alterations in tumors. Images derived from the RLGS analyses produced a complex spot pattern. The proper segmentation of these spot patterns, the extraction of useful data, and the management of these data can be problematic. The propose of this study was to analyze the genetic alterations in K-ras transformed NIH-3T3 cell lines by RLGS and to evaluate the efficacy of the GELLAB II+ software in the processing and data extraction of the RLGS spot patterns. The test groups consisted of 3 NIH-3T3 control cell lines and 8 K-ras transformed cell lines. After RLGS, the images were processed and segmented by the GELLAB II+ software, each image contained about 1400 spots. Images were then landmarked and that spots between gels were matched. The results indicate ten spots were amplified (2 of the spots were statistically significant p < 0.01), one spot was deleted and one new spot was detected in the K-ras transformed cell lines as compared to NIH-3T3 cells lines. Although designed for protein gel electrophoresis, GELLAB’s spot matching and database support provide important tools for the investigation of RLGS spot patterns.
Introduction

The pathogenic basis for cancer involves the activation of oncogenes and inactivation of tumor suppressor genes. Few techniques have been able to show us these changes on a genomic level. A newly developed two-dimensional gel electrophoresis assay, Restriction Landmark Genomic Scanning (RLGS) can simultaneously examine the entire genome of a given cell line or tumor for DNA abnormalities. Genomic DNA is cut with a rare cutter restriction enzyme and end-labeled with $^{32}$P-nucleotides. The labeled fragments are separated into two dimensions by size to generate thousands of individual spots on X-ray films (1, 2). The radioactive intensity of each spot represents the copy number of this fragment in the genome (1). By comparing normal and tumor DNA patterns, genetic alterations in tumor tissues or cell lines can be detected.

Images derived from the RLGS technique produce a complex spot pattern, containing thousands of data points per gel. The proper segmentation of these spot patterns, ascertaining useful data, and data management can be problematic. Before the advent of computer systems which could analyze gel electrophoretograms, gel images were compared by visual inspection and subjective observations.
were made about the changes in spot density from one gel to the next. This qualitative method is effective when observing novel or deleted spots between gel sets. However, the method is inappropriate for examining finite changes in spot density which represent some of the DNA copy number abnormalities. The problem is compounded by the “unwieldy” size of the data sets produced by RLGS, local inhomogenates found in the gels, and the differences in the exposure time between gels.

In this study, we investigated GALLAB-II+’s ability to analyze RLGS images of DNA from K-ras oncogene transformed NIH-3T3 cells. The software was evaluated on the ability to segment gel spots from the image background, and to pair spots between the reference (R-gel) and the other gel images. The software was also evaluated on it’s ability to extract heuristic data from the gel spots, that pertains to quantitative changes and trends between normal and transformed cell line gels.
Materials and Methods

Sample collection

High molecular weight DNA from mouse lung tumors were transfected into NIH-3T3 mouse fibroblasts by the calcium phosphate precipitation method. The cell lines were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum (Gaithersburg, MD). All cell were cultured at 37° C in humidified 7.5% CO₂/92.5% air.

DNA Isolation

High molecular weight DNA is essential for high quality 2-DGE. The procedure for DNA isolation has been described previously (1). DNA strand breakage is minimized by careful isolation, storage and manipulation of DNA in this procedure. After harvest, the cell were placed immediately in digestion buffer containing proteinase K. Following triple phenol extraction, DNA was precipitated with ethanol and dissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). The quantity and purity of the DNA was measured in a spectrophotometer at wavelengths of 260 nm/280 nm and the quality of the DNA was checked by electrophoresis in an 0.8% agarose gel.

Procedure for 2DGE

Not I digestion: About 2-10 µg of DNA was digested in 100 µl Not I (New England Biolabs, Beverly, MA) digestion
buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 100 mM NaCl) containing 100 U of Not I for a minimum of 2 hrs at 37°C.

Not I site labeling: The sticky ends generated from Not-I restriction digests were filled with [³²P]α-dGTP and [³²P]α-dCTP. The reaction buffer contained 40 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, 50 mM NaCl, 10 mM DTT, 2 µl each of [³²P]α-dGTP (6000 Ci/mmol) and [³²P]α-dCTP (6000 Ci/mmol) and 20 U Sequenase (version 2.0, United States Biochemicals, Cleveland, OH). The reaction was allowed to proceed for 20 min at 37°C, and terminated by incubation at 65°C for 5 min.

EcoR V digestion: DNA was then digested with EcoR V (New England Biolabs, Beverly, MA) for 2 hr at 37°C. The reaction buffer contained 100 U of EcoR V, 50 mM Tris (pH 8.0), 10 mM MgCl₂ and 50 mM NaCl.

First-Dimensional (1st-D) gel electrophoresis: First-dimensional electrophoresis was carried out in a specially designed vertical apparatus which contained 5 mm (inner diameter) glass tubes connecting the upper and lower buffer tanks. The tubes (60 cm long) were filled with Seakem Gold agarose gel (FMC Bio product). The top 2 cm of the 1st-D gel piece (stacking gel) consisted of 0.5% agarose and 20% sucrose while the remainder of the gel (running gel) consisted of 0.8% agarose and 5% sucrose as previously
described (1). After EcoR V digestion, the DNA was loaded onto the top of the 1st-D gel and electrophoresed at 100 V for 2 hrs, followed by 230 V for 22 hr.

In situ digestion with Hind III: The 1st-D gel was expelled from the glass tube and trimmed to a 30 cm long piece (so called "noodle") which contained most of the DNA. After 30 min equilibration in Hind III reaction buffer, the tube gel was aspirated into a 6 mm inner diameter plastic tube and filled with Hind III reaction solution containing 2000 U of Hind III. The tube was incubated at 37°C overnight.

Second-dimensional (2nd-D) gel electrophoresis: After Hind III digestion, the tube gel was equilibrated in 1x TBE buffer and loaded onto a 0.8% horizontal ultra pure agarose gel, and electrophoresed at 45 V for 50 hr. The agarose gel was then dried at 60°C and autoradiographed.

Image Processing and Analysis
Image Acquisition: The autoradiographs were converted to a TIFF-image format with a spatial resolution of 400 µm by Bio-Rad densitometer and a grayscale-to-OD calibration was conducted. The resulting images were 712 X 970 pixels at 8 bits in size. The images were then imported into the GELLAB II+ software running in a Windows environment.
Image Processing: After experiment files were created for each of the gels in the GELLAB II+ software and a reference gel was set. The images were smoothed using a lowpass averaging filter (3 X 3), and background subtracted.

Spot Segmentation: After the gel images have been processed by GELLAB II+ the images are then “batch job” segmented for spot detection based on the following parameter values:
Spot Area Limit = 12 - 2400 pixels, Spot Optical Density Limit = 3 - 9000 OD, Spot Core Threshold = 4

The gel images were then visually inspected for imperfections in the spot segmentation algorithm, and were manually corrected if necessary.

Landmarking: After segmentation, the images were interactively assigned a total of 116 landmarks. All spots used for landmarking were common in all the images and were of similar normalized optical density. The landmarks were automapred from the R-gel to all the other gels, where they were align to the R-gel’s landmarks. After completion of the landmarking, the gels were inspected for improper alignment and corrected if necessary.
Pairing: The gels were then paired to the reference gel based on the 116 aligned landmarks using GELLAB’s pairing operation. The correct pairing of sample gel spot to R-gel was confirmed by visual inspection and spot pairing was corrected manually if necessary.

Spot Normalization: Spots for each gel were then normalized as a percentage of the total density of a gel. This normalization allows for meaningful comparisons about spot density to be drawn between gels of different exposure times.

Data Export and Analysis: Spot data including Spot ID, Spot raw density, normalized density, mean density, X location, and Y location was exported as a .MDB file to Microsoft Access 2.0 database software. The data was queried to look for differences in spot density between the normal and transfected cells line images.

Statistical Methods: Spot densities were analyzed by F-test and standard Student t-test.
Results

The GELLAB-II+ software (Scanalytics, Billerica, MA) derived from the GELLAB programs by Peter F. Lemkin is a 2-D electrophoresis analysis program. Although GELLAB-II+ was originally created for 2-D gel protein analysis, it seems well suited for overcoming the aforementioned problems of RLGS analysis. This software has four main components:

Gel Image Processor: This operator allows the images to be enhanced to maximize the accuracy of the segmentation, and pairing algorithms. This is accomplished by filters that improve contrast, reduce low spatial frequency signals (background), and increase the signal to noise ratio.

Segmentation Algorithm: This algorithm is responsible for the proper separation of spots produced by the gel image. This algorithm is based on the digital analog of a spatial second derivative of a Gaussian like function (3). The 2nd maximum of the second derivative determines the outer extend of the spots outward propagation (3) (Figure 1). The GELLAB-II+ segmentation algorithm allows for the user to define the parameters of central core size, spot density, and spot area. This “filters out” objects that may be an unwanted or experimental artifact. The segmentation
algorithm also characterizes the spots based on their area, integrated optical density, normalized optical density, spot centroid, mean density, location on the gels Cartesian coordinate plane, and minimum and maximum density.

Landmarking and Spot Pairing: Despite being subjected to rigidly controlled conditions, replicate 2-D gels will still exhibit small changes in spot migration patterns from gel to gel. Landmarking is used to mathematically normalize these changes in spot migration between gels. The reference gel (R-gel) acts as a representative gel for all other gel images in the experimental set. Landmarks are interactively placed on the R-gel and the experimental gel sets. Then these landmarks are aligned by the user. The landmark represents an “anchor” between gels, which allows the computer to align all other spots in the vicinity with their corresponding spots in the other gels (4). The area around a landmark contains a half-radius of certainty Ri, representing landmark I, which is defined to be half the distance from the nearest landmark to landmark I (4). Spots within this radius have a higher probability of being aligned than the spots outside the radius. The pairing algorithm consists of two parts, the primary pairing algorithm and the secondary pairing algorithm. In the

89
primary pairing algorithm, spots surrounding a landmark are mapped on a Cartesian coordinate system, with the landmark spot being the origin (0,0). Each R-gel spot in the region is preliminary paired to its nearest neighbor by the minimum Euclidean distance in the other gels (Figure 2) (4). Based on the local image environment the spots around a landmark can fall in to one of four categories: A sure pair SP, A possible pair, PP, an ambiguous pair, AP, an unresolved spot, US. The second pairing algorithm is then used to re-pair the AP and US spots with neighboring landmark regions. The re-paired spots are assigned to a landmark regions with the smallest dL value (Figure 3).

Generation of Database and export: GELLAB II+ produces a database from the gel sets and gel spot set. Information about spot characteristics can be queried from within the software or may be exported to a relational database.

Spot Segmentation and Detection: Eleven two-dimensional electrophoresis gel images (3 normal cell line images and 8 K-ras cell line images) were analyzed using the GELLAB II+ software. The GELLAB-II+ system detected an average $1218 \pm 48.3$ spots per gel or a total of 13,397 spots. These numbers were consistent with the counts made by visual
inspection of the images, only minor interactive corrections in the spot segmentation had to be made.

Spot Pairing: The Gellab II+ paired an average 74.69% ± 1.1% spots per gel to the reference gel using a total 116 landmarks. Interactive spot pairing corrections were made before normalization and export of spot data to ensure proper spot pairing.

Quantitative Spot Analysis: After interactive filters were applied to the spot data in MS Access © a total of 10 spots showed differences in spot density between the normal cell lines and the K-ras cell line images. As summarized in Table 1, eight spots form K-ras cell lines were amplified, but were not statistically significant from the normal cell line densities (p > 0.05). Spot 107 was amplified (12.5% of the gels), spot 647 (12.5% of the gels), spot 890 (12.5% of the gels), spot 1084 (12.5% of the gels), spot 408 (12.5% of the gels), spot 256 (12.5% of the gels), spot 227 (12.5% of the gels), and spot 10 (12.5% of the gels). One spot, 187 was deleted (12.5% of the gels) in the K-ras gels and spot 137 was a new spot found in a K-ras gel. Two of the ten K-ras line gels showed spots with amplification that were statistically significant from the normal images.
Spot 1081 (87.5% of the gels) average expression was 58% greater in the K-ras cell lines than in normal cell lines, and was statistically significant (p < 0.01). Spot 257 (100% of the gels) also showed a statistically significant difference in expression in the K-ras gels from the normal gels (72% increase). All data was confirmed by visual inspection to ensure proper spot data alignment.

**Discussion**

Recent studies have show that RLGS is a powerful technique for investigation of genomic alterations involving carcinogenic events. This technique allows the examination of DNA abnormalities by scanning the whole genome in a single experiment. This technique, in concert with the use the GELLAB-II+ software, has allowed the investigator a more novel and effective way to analyze the aforementioned genomic alterations.

The RLGS’s complex spot patterns were segmented by the GELLAB-II+ segmentation algorithm with a very high level of accuracy. Only minor changes had to be made to properly segment all spots in each gel, based on the user defined segmentation algorithm parameters. The improperly segmented spots were generally found in the area of the high molecular weight DNA fragments. This area was at the top
left hand corner of the gel and contains the highest spot density per unit area. The large number of spots in this area were closely adjoined to other spots, which contributed to miss-segmentation. However, GELLAB-II’s Lapacian segmentation algorithm performed better than previous techniques we have employed, such as local thresholding and the color overlay technique.

The DNA fragment spots found among the gels in the RLGS study must be matched if any useful data about changes and trends between gel sets can be extracted. In this study we used 116 landmarks to guide the GELLAB-II+ pairing algorithm, which paired a mean 74.69% ± 1.1% of the spots between the sample gels to the R-gel. The pairing was confirmed by the SHOW PAIRING option which allows the operator to view the matching vectors (A line from the center of a spot that shows it’s direction and distance from it’s match pair on the R-Gel). This option shows the pairing geometry of the gel and points out pairs which are mismatched. The percent of correct pairs between gels is dependent on the geometric congruence and quality of the gels. Generally, most of the gel spot mispairings and unpairings in this study were attributed to differences in the exposure time of the gel autoradiographs. This is due to the reduced number of spots in low exposure gels in the
low molecular weight DNA area of the gel (Bottom right of the gel).

As 2-DGE gels have increased in size and complexity, it has become almost impossible to perform a thorough qualitative analysis of the gels by visual inspection. Quantitative analysis can be conducted only with the use of image analysis software. In this study the GELLAB-II+ software allowed use to query the spot data to make statistically meaningful conclusions about differences in spot density between the gel sets. Two spots (spot 1081, and spot 257) found by the GELLAB-II+ software, which showed amplification in the K-ras transformed cell line gels, were shown statistically significant (p < 0.01). These spots most likely represent genomic alterations that are related to the K-ras transformation process, because of their consistent representation throughout the transformed cell line gels. A more detailed characterization of these spots by PCR-SSCP and direct sequence analysis, and the ramifications of these analyses will be presented in a future paper. 8 spots were found to differ in spot density (change > 25% of normalized density) in the transformed cell line gels from the normal NIH-3T3 cell line gels. Each of these changes in spot density occurred only in one of the 8 transformed cell line gel. These changes implied that
these genomic alterations are sporadic, and are local secondary events, which are not part main mechanism of K-ras transformation.

The RLGS technique used with the GELLAB-II+ has provided a new way to heuristically screen the genome, looking at genetic alterations. Future enhancements in gel quality and an increased number of landmarks used, should allow for reduced operator interaction, and should improve the accuracy and efficacy of the assay.
References


Legend

Figure 1. Segmentation Algorithm
(a) 2D representation of a greyscale spot with a Gaussian-like shape \( g \). (b) The direction of the second derivative \( g'' \) or Lapacian of the Gaussian like function. (c) The limit of the propagated central core as determined by the second maximum in the \( g'' \) Lapacian function (3).

Figure 2. R-gel and Gel sets
Spots A, B, and D in the R-gel are R-spots. Spot E, though it is in the R-gel, is not an R spot. This is because it is unmatched in any of the other gels in the set. The list of R-spots in the gel is \( (A, B, D) \). R-spot set A consists of \( \{A_r, A_1, A_2, \ldots, A_{n-1}\} \). This set has one member from each gel. Some R-spot sets such as B and D have members only from some of the gels. R-gel and R-spots sets produce a 3D stacked array of all the gels in the experiment allowing for gel by gel comparisons to be made throughout the array. By looking at differences in the gel set density between the normal gel images and tumor gel images the investigator can make quantitative measurements about DNA copy number abnormalities (4).
Figure 3. Spot Pairing

The arrow is the half-radius of certainty for this landmark spot A. Each dL is the distance from the landmark spot to the midpoint of the line joining the positions of the potential paired spot in gels 1 and 2 Spot 1, an unresolved spot (dP = 0); spot pair 2, a sure pair (dL< sub>2 </sub> < R< sub>A </sub>, dP < dT< sub>1</sub>); spot pair 3, a probable pair (dL< sub>3 </sub> > R< sub>A </sub>, dP< sub>3</sub> < dP<sub>3</sub> < dT<sub>2</sub>); spot pair 4 a probable pair (dL<sub>4</sub> < R<sub>A</sub>, dT<sub>2</sub> > dP<sub>4</sub> > dT<sub>1</sub>); spot pair 5, a probable pair (dL<sub>5</sub> > R<sub>A</sub>, dP<sub>5</sub> < spot pair 6 a probable pair (dL<sub>6</sub> < R<sub>A</sub>, dP<sub>6</sub> < dT<sub>2</sub>): spot pair 6', ambiguous pair (dL<sub>6'</sub> > R<sub>A</sub>, dP<sub>6'</sub> < dT2, dP<sub>6</sub>' < dP<sub>6'</sub>); spot 7 an unresolved spot within the circle of certainty (4).

Table 1. Genetic alterations in normal and K-ras cell lines.
Figure 1. Segmentation Algorithm
(a) 2D representation of a greyscale spot with a Gaussian-like shape $g$. (b) The direction of the second derivative ($g''$) or Lapacian of the Gaussian like function. (c) The limit of the propagated central core as determined by the second maximum in the $g''$ Lapacian function (3).
Figure 2. R-gel and Gel sets
Spots A, B, and D in the R-gel are R-spots. Spot E, though it is in the R-gel, is not an R spot. This is because it is unmatched in any of the other gels in the set. The list of R-spots in the gel is \{A, B, D\}. R-spot set A consists of \{A_r, A_1, A_2, \ldots , A_n\}. This set has one member from each gel. Some R-spot sets such as B and D have members only from some of the gels. R-gel and R-spots sets produce a 3D stacked array of all the gels in the experiment allowing for gel by gel comparisons to be made throughout the array. By looking at differences in the gel set density between the normal gel images and tumor gel images the investigator can make quantitative measurements about DNA copy number abnormalities (4).
Figure 3. Spot Pairing
The arrow is the half-radius of certainty for this landmark spot A. Each dL is the distance from the landmark spot to the midpoint of the line joining the positions of the potential paired spot in gels 1 and 2. Spot 1, an unresolved spot (dP = 0); spot pair 2, a sure pair (dL₂ < Rₐ, dP < dT₁); spot pair 3, a probable pair (dL₃ > Rₐ, dP₁ < dP₃ < dT₂); spot pair 4 a probable pair (dL₄ < Rₐ, dT₂ > dP₂ > dT₁); spot pair 5, a probable pair (dL₅ > Rₐ, dP₅ < dT₁); spot pair 6 a probable pair (dL₆ < Rₐ, dP₆ < dT₂); spot pair 6', ambiguous pair (dL₆' > Rₐ, dP₆' < dT₂, dP₆' > dP₆'); spot 7 an unresolved spot within the circle of certainty (4).
Table 1.

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>10</th>
<th>137</th>
<th>107</th>
<th>187</th>
<th>227</th>
<th>256</th>
<th>*257</th>
<th>408</th>
<th>647</th>
<th>890</th>
<th>*1081</th>
<th>1084</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NIH3T3 2-R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NIH3T3 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NIH3T3 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Transformed</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>
<tr>
<td>K-RAS 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 3</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 5</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>K-RAS 6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>K-RAS 8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Genetic alterations in normal and K-ras cell lines.
0 = No change
+ = a new spot in the transformed cell line
- = a deletion in the transformed cell line
1 = increase in spot density greater than 25% the normal cell line normalized spot density.
2 = increase in spot density greater than 50% the normal cell line normalized spot density.
3 = increase in spot density greater than 100% the normal cell line normalized spot density.
-1 = decrease in spot density greater than 25% the normal cell line normalized spot density.
-2 = decrease in spot density greater than 50% the normal cell line normalized spot density.
-3 = decrease in spot density greater than 100% the normal cell line normalized spot density.
* Spot is statistically (p < 0.01
MANUSCRIPT IV

Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Human Bronchial Epithelial Cells Irradiated with High-LET Alpha Particles

Patrick, J.L.¹, Weaver, D.A.¹, Gu, P.¹, Lemkin, P.F.², Willey, J.C.¹ Hei, T.K.³ You, M.¹,⁴

¹Medical College of Ohio, 3120 Glendale Ave., Toledo, OH 43614
²Laboratory of Experimental and Computational Biology, National Cancer Institute/FCRF, Frederick, MD
³Center for Radiological Research, VC11-218, Columbia University, 630 West 168th St., New York, NY 10032
⁴To whom requests for reprints should be sent. Address correspondence to Ming You, MD, Ph.D., Department of Pathology, Medical College of Ohio, Health Education Building Rm. 206, 3000 Arlington Avenue, Toledo, OH 43699.

Abbreviations: RLGS, restriction landmark genomic scanning; 2-DGE, Two-dimensional gel electrophoresis; LET, Linear Energy Transfer
Abstract

How high-LET radiation affects the genome of human bronchial epithelial cells is important in the understanding of radon induced malignant transformation. In this study, we analyzed three groups of human papillomavirus immortalized human bronchial epithelial cell lines by restriction landmark genomic scanning (RLGS): 1) BEP2D control group which did not receive any radiation. 2) An R30 experimental group which received a single 30 cGy dose of 150 keV/µm 4He ions. 3. An H2BT experimental group which received a single 60 cGy dose of 150 keV/µm 4He ions. RLGS is a relatively new way to detect and analyze a number of genomic changes including DNA amplification, loss of heterozygosity, chromosomal deletion, and chromosomal translocations. RLGS spot patterns were processed by GELLAB II+ software and by a novel Java-base image-analysis program called Webgel.

The RLGS results in the three cell line groups were:
1) Eight spots had a decreased density greater than 100% compared to the BEP2D 54 reference gel. 2) Three deletions occurred in the H2BT, R30-TL pl1, and H1ATN p22 cell lines. 3) Two spots, 7 and 10, in the H2BT cell line were amplified to a density greater than 100% compared with the BEP2D 54 control group.
Introduction

Studies have shown a link between high radon exposure and bronchogenic carcinogenesis (1). Studies of uranium miners indicate that radon progeny is a risk factor in the development of the lung cancer. (1) Despite a good link between high radon exposure and lung carcinogenesis in uranium miners, extrapolation of these data to form a domestic exposure cancer model has been problematic in the past. Recent meta-analysis of data from different domestic radon retrospective studies has produced an adequate low-dose radon risk model (2).

Radon is a colorless gas that is a decay product of uranium-238. This radioactive gas produces a series of radionuclide progeny that emit high linear energy (LET) alpha particles (3).

Recent technology has allowed us to probe the effects of radon on the genome of respiratory epithelial cells in a single experiment. Restriction Landmark Genomic Scanning (RLGS) can scan the entire genome of a sample for DNA amplifications, loss of heterozygosity, chromosomal deletions, and chromosomal translocations (see ref. 4 for review). The analysis of the RLGS spot patterns is complex and requires the use of specialized software. In past studies our group used the stand-alone GELLAB-II+ software,
but in this experiment a novel Java based software called WebGel was implemented.

Many biological experiments utilize quantitative measurements in 2D gels – both protein and DNA gels. These gels may be scanned into image files and then spots or bands quantitated using stand-alone software such as GELLAB-II (5) or GELLAB-II+ (6). However, as more research collaborations are taking place over the Internet, it has become critical to share these quantitative data between researchers. This allows individual research group members to investigate their data and then share this work-in-progress with others in their group. Dr. Peter Lemkin developed a World Wide Web-accessible system, WebGel (7), for interactively exploring quantitative measurements and qualitative differences in electrophoretic gels across the Internet. In addition to presenting images of corresponding spots between gels and image comparison using Flicker-comparison, WebGel can perform exploratory data analysis. It may be used to compute and compare mean DNA concentration values between corresponding spots or bands for multiple samples run in separate gels, and then generate statistics on these comparisons. Because visualization is critical to understanding one’s data, WebGel helps organize labeled gel images in synthesized
graphic images, called montages, consisting of a corresponding spot as seen in different gels (7).

Other types of graphic views of gel images may be generated; sets of spots, segmented gels, etc., may be displayed simultaneously for multiple gels. These graphic displays are active and may be used for performing database operations directly by clicking on spots in the images.

The purpose of this study is to investigate the effects of high LET Radon radiation on the genome of human papillomavirus immortalized human bronchial epithelial cell lines by restriction landmark genomic scanning (RLGS), then analyze these RLGS spots using GELLAB II+ and use some aspects of the novel Java based WebGel program.

**Materials and Methods**

**Sample collection**

The nontumorigenic BEP2D HPV 18-immortalized human bronchial epithelial cell line (8-10,12,13) was used in the study as a control group. The tumor and BEP2D lines were cultured in serum-free LHC-8 medium (14) as previously described (15).

**DNA Isolation**
High molecular weight DNA is essential for high quality 2-DGE. The procedure for DNA isolation has been described previously (16). DNA strand breakage is minimized by careful isolation, storage, and manipulation of DNA in this procedure. After harvest, the cells were placed immediately in digestion buffer containing proteinase K. Following triple phenol extraction, DNA was precipitated with ethanol and dissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). The quantity and purity of the DNA was measured in a spectrophotometer at wavelengths of 260 nm/280 nm and the quality of the DNA was checked by electrophoresis in an 0.8% agarose gel.

Procedure for 2DGE

Not I digestion: About 2-10 µg of DNA was digested in 100 µl Not I (New England Biolabs, Beverly, MA) digestion buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl$_2$, 100 mM NaCl) containing 100 U of Not I for a minimum of 2 hrs at 37°C.

Not I site labeling: The sticky ends generated from Not-I restriction digests were filled with $[^{32}P]$$\alpha$-dGTP and $[^{32}P]$$\alpha$-dCTP. The reaction buffer contained 40 mM Tris-HCl (pH 7.5), 20 mM MgCl$_2$, 50 mM NaCl, 10 mM DTT, 2 µl each of
$[^{32}P]$$\alpha$-dGTP (6000 Ci/mmol) and $[^{32}P]$$\alpha$-dCTP (6000 Ci/mmol) and 20 U Sequenase (version 2.0, United States Biochemicals, Cleveland, OH). The reaction was allowed to proceed for 20 min at 37°C, and terminated by incubation at 65°C for 5 min.

EcoR V digestion: DNA was then digested with EcoR V (New England Biolabs, Beverly, MA) for 2 hr at 37°C. The reaction buffer contained 100 U of EcoR V, 50 mM Tris (pH 8.0), 10 mM MgCl$_2$ and 50 mM NaCl.

First-Dimensional (1st-D) gel electrophoresis: First-dimensional electrophoresis was carried out in a specially designed vertical apparatus which contained 5 mm (inner diameter) glass tubes connecting the upper and lower buffer tanks. The tubes (60 cm long) were filled with Seakem Gold agarose gel (FMC Bio product). The top 2 cm of the 1st-D gel piece (stacking gel) consisted of 0.5% agarose and 20% sucrose while the remainder of the gel (running gel) consisted of 0.8% agarose and 5% sucrose as previously described (8). After EcoR V digestion, the DNA was loaded onto the top of the 1st-D gel and electrophoresed at 100 V for 2 hrs, followed by 230 V for 22 hr.
In situ digestion with Hind III: The 1st-D gel was expelled from the glass tube and trimmed to a 30 cm long piece (so called "noodle") which contained most of the DNA. After 30 min equilibration in Hind III reaction buffer, the tube gel was aspirated into a 6 mm inner diameter plastic tube and filled with Hind III reaction solution containing 2000 U of Hind III. The tube was incubated at 37°C overnight.

Second-dimensional (2nd-D) gel electrophoresis: After Hind III digestion, the tube gel was equilibrated in 1x TBE buffer and loaded onto a 0.8% horizontal ultra pure agarose gel, and electrophoresed at 45 V for 50 hr. The agarose gel was then dried at 60°C and autoradiographed.

Image Processing and Analysis

Image Acquisition: The autoradiographs were converted to a TIFF-image format with a spatial resolution of 400 µm by Bio-Rad densitometer and a grayscale-to-OD calibration was conducted. The resulting images were 712 X 970 pixels at 8 bits in size. The images were then imported into the GELLAB II+ software running in a Windows environment.
Image Processing: After experiment files were created for each of the gels in the GELLAB II+ software and a reference gel was set. The images were smoothed using a lowpass averaging filter (3 X 3), and the background was subtracted.

Spot Segmentation: After the gel images had been processed by GELLAB II+ the images were segmented for spot detection based on the following parameter values:
Spot Area Limit = 12 - 2400 pixels, Spot Optical Density Limit = 3 - 9000 OD. The gel images were then visually inspected for imperfections in the spot segmentation algorithm, and were manually corrected if necessary.

Webgel: Gel images were placed on the WebGel server site at the NCI. The images were flickered to look for qualitative changes in gel spot density. Once candidate spots were identified, numeric density data between spots were analyzed GELLAB II+ to show quantitative changes in spot density.

Results
Spot Density Analysis: After RLGS images were analyzed using the GELLAB II+ software, a total of 11 spots showed
differences in spot density between the BEP2D cells and the experimental irradiated groups, R30 and H2BT. The following eight spots had a deceased density greater than 100% compared to the BEP2D 54 reference gel: Spot 1 had a deceased density in 87.5% of the gels, spot 2 (12% of the gels), spot 3 (25% of the gels), spot 4 (37.5% of the gels), spot 5 (12% of the gels), spot 6 (25% of the gels), spot 9 (37.5% of the gels), spot 11 (12.5% of the gels). Two spots, 7 and 10, in the H2BT cell line were amplified to a density greater than 100% compared to the BEP2D 54 reference gel. Spot 8 was deleted in 37.5% of the gels when compared to the BEP2D 54 reference gel. These deletions occurred in the H2BT, R30-TL p11, and H1ATN p22 cell lines. Spots 4 and 5 showed differences in density greater than 100% between the BEP2D p22 and BEP2D p54. These spots were removed from the study, for it is highly improbable that these changes are legitimate genomic abnormalities caused by radon exposure. These spot changes most likely represent polymorphs between the BEP2D control cell lines or they may represent experimental artifact. After discarding those spots, a total of 8 DNA copy number abnormalities remained. H2BT and H1ATN cell lines showed the greatest number of genomic events, with six each.
Discussion

The understanding of how high-LET radiation affects the genome bronchial epithelial cells has been an ongoing process for many years. RLGS has allowed us to examine some of the genetic events that occur when bronchial epithelial cells exposed to high-LET radiation. RLGS has demonstrated that a number of alterations occur during the interaction of high-LET radiation and DNA. Most of these alterations resulted in the deletion or the down regulation of particular DNA fragments. This seems consistent with the physical interaction of High-LET radiation and genomic DNA. Previous studies indicate that High-LET radiation has a profound effect of DNA. These large particles with their excessive momentum have the ability to literally remove atoms and molecules from the DNA structure. This deleterious effect on DNA is more likely to lead to DNA copy fragment down regulation or DNA copy fragment deletion than other genetic alterations. Further characterization using cloning and sequence technology on the down regulated spots along with the other genetic abnormalities may provide important answers to how high-LET radiation interacts and affects human DNA. The RLGS spot patterns in the study were analyzed by using GELLAB II+ software package which has been proven to be an effective tool for
this use. Another aspect of the study was to involve the Webgel software package to evaluate its potential in future studies as a method of image analysis.

Only the Flicker program of Webgel was used in the study because the other analysis and data processing tools were under construction. It is hoped that the system can be fully utilized in future studies allowing different groups of investigators to share data using RLGS Web based libraries.
References


121
DISCUSSION

This dissertation had two overall objectives: To investigate the applicability of RLGS in detecting genomic alterations in a number of different tumor and cell line models, and to investigate and develop new methods to process, analyze and harvest spot data from the complex RLGS gel patterns. The studies in this dissertation contribute to a growing body of data that suggests that RLGS is an effective tool in characterizing genetic changes found in a wide variety of tumor and cell line models. Information contained in this document also indicates that RLGS gel patterns require a substantial amount of computational analysis before useful data can be extracted from the RLGS patterns in a timely fashion.

In the first manuscript, "The Reduction of Genetic Alterations by 4-Hydroxyphenylretinamide in MethylNitrosourea Induced Rat Mammary Tumors," RLGS detected 35 altered spots in the MNU treated group compared to the untreated control group. Of the 35 altered spots 12 were detected as novel spots in the MNU treated group, 10 spots were absent in the MNU treated group, 8 spots were shown to have reduced intensity by 50%, and 5 spots in the tumor group were found to be amplified. The NMU-4HPR showed
a decrease in alterations compared to the only NMU group. This manuscript represents our first attempt at using image processing and analysis to characterize the genetic changes between experimental and control RLGS spot patterns. We employed two separate programs. The first was an in-house color hybridization macro based on a program written by Wyan Rasband at the NIH. This macro allowed the group to scan for differences between gels in a qualitative manner. Due to the nonlinear migration differences between the gels, the color overlay macro could only scan small regions at a time. The process was repeated until a complete mosaic of the overlay was completed. After the spots of interest were determined, a second program was used to quantify the differences in spot intensities between gels. Qualification was based on a normalization procedure in which intensities of 10 "regular" or unchanged spots were measured surrounding an altered spot of interest. The same procedure was performed on the corresponding altered spot on a different gel, thereby allowing us to make a meaningful measurement of the difference between the two. These two techniques were effective in scouting out gross changes in spot intensity between gels, but the process was laborious and required significant interaction by the operator. Another shortcoming of this project was that it
heavily relied on the observer’s subjective ability to scan gels. Despite the successes of this project, it was obvious that there was room for improvement as was demonstrated in the subsequent studies.

The second manuscript, "Detection of Genomic Alterations in Human Cervical Cancer by Two-Dimensional Gel Electrophoresis," compared the genetic differences between intraepithelial neoplasia (CIN) lesions, cervical squamous cell carcinomas, cervical adenosquamous cell carcinoma, cervical adenocarcinomas, and normal cervical tissue using computer assisted RLGS. The RLGS patterns showed that nine spots were amplified in the cervical carcinomas and two spots were amplified in the CIN lesions compared to normal cervical tissue. It was also shown that 12 spots were either absence or had a reduced intensity compared to the normal cervical tissue and 6 spots were shown to have reduced intensity the in CIN group. This study employed a similar image analysis strategy with some modifications. The color overlay analysis software from the previous study was used in concert with a commercial pointel based warping system. The warping system allows the operator to compensate for the nonlinear migration differences between gels by "stretching" the gel images thereby allowing a color overlay of an entire gel image. A second warping
system (Applied Imaging, Inc.) was used to warp spot coordinates to allow for a spot matching and quantification. The quantification method now relied on a global normalization (all spots in a gel were compared to the altered spot) as opposed to local normalization thereby reducing the risk that local variations in background noise of the gel images would affect spot intensity data. The first warping system was effective in reducing the time necessary to create color overlays between gels and allowed the operator to see color changes, which represent intensity changes, throughout an entire gel image. The second warping system and matching algorithm required exhaustive operator interaction due to the inappropriate matching of non-corresponding spots. Despite the improved quantification ability of the software, it was believed that another system with an improved matching algorithm would be more effective in delineating the data sets produced by RLGS patterns in a timely manner. That system became available during the start of our next study.

In the third study in this dissertation, a new image processing and analysis system donated by Scanalytics Corp. and derived from GELLAB programs by Peter F. Lemkin was implemented in the "Computer Assisted Analysis Of Restriction Landmark Genomic Scanning Images Of DNA From K-
Ras Oncogene Transformed NIH-3T3 Cells" project. Analysis of the RLGS spot patterns between K-ras transformed NIH3-3T3 and NIH-3T3 control cell lines indicated the 10 spots were amplified (2 of the spots were statistically significant $p < 0.01$), one spot was deleted, and one novel spot was detected in the K-ras transformed NIH3-3T3 gels. The two spots (1081, and 257) most likely represent genetic changes indicative of the K-ras transformation process. Future studies and characterization of these spots would be beneficial in the understanding of how K-ras or other gene products are involved in NIH-3T3 transformation. Despite its original intended use for two dimensional protein gel analysis, the GELLAB II+ software package proved to be invaluable in harvesting data from the RLGS spot patterns. The Lapacian segmentation algorithm of the software package was extremely accurate and only required minor changes by the operator. The pairing algorithms in the study required 116 user interactive landmarks, and correctly paired 74.69% $\pm$ 1.1% of the spots from the other experimental gel’s image to the reference gel image. This was confirmed by the operator using the show paring operation and the visual inspection of the matching vectors. The software package performed both Global and Local normalizations, effectively reducing any variation in gel image background noise from
affecting spot intensity data. The spot data were then transferred to an in-house relational database program written by the author in structured query language (SQL), called COG, for analysis. GELLAB II+ in concert with COG to be very effective in data acquisition and analysis, allowing the operator to scan the data set for meaningful differences in spot intensity.

The fourth manuscript, "Computer Aided Analysis of Restriction Landmark Genomic Scanning Images from Human Bronchial Epithelial Cells Irradiated with High-LET Alpha Particles," compared RLGS spot patterns from three groups of human papillomavirus immortalized human bronchial epithelial cell lines: The control group BEP2D did not receive any radiation; R30 which received a single 30 cGy dose of 150 keV/µm 4He ions; and the H2BT group which received a single 60 cGy dose of 150 keV/µm 4He ions. The RLGS analysis results indicated that eight spots in the experimental groups had a decrease in intensity greater than 100% compared to the control group and three deletions had occurred in the experimental group. Two spots were found to be amplified in the H2BT group, where their spot intensity was greater than 100 percent of the control group. One spot had a decreased density in 87.5% of the gels, and may represent an important genetic event relating
to the interaction of High-LET alpha particles in geneomic DNA. Further characterization of this DNA fragment copy number abnormality may provide some insight as to the effects of radiation on human bronchial epithelial cells and how it relates to transformation. The RLGS system was analyzed using GELLAB II+ as described previously, along with a novel java based program called Webgel. Only the Flicker program of Webgel was used in this study. Flicker provides a method by which two gels can be displayed simultaneously and produces a resultant image that oscillates between the two gels images allowing the user to observe differences in spot intensity between two gels. This is a qualitative method of image analysis and was not required in the study because of the presence of GELLAB II+, but it was utilized to explore new types of technology in web-based image processing.

RLGS has a number of advantages over more conventional means of genetic analysis in the field of cancer research, but it has been demonstrated that this method requires an effective image processing and analysis system to deal with the complex data that it produces. Future RLGS projects may benefit from advances in Web technology, allowing different investigative groups to collaborate and share data through the use of web based RLGS libraries.
Bibliography


Formell F.; Clerici M.; Campa T.; Di Mauro M.G.; Magni A.; Mascotti G.; Moglia D.; De Palo G.; Costa A.; and


Okuizumi, H.; Muramatsu, M.; Komatsubara, H.;
Sasaki, T.; Ueda, T.; Mise, N.; Takagi, N.; Plass, C.;
U2af binding protein related sequence on mouse
6:33-40.

59:1692-1696.

Imoto, H.; Hirotsune, S.; Muramatsu, M.; Okuda, K.; Sugimoto,
O.; Chapman, V.; Hayashizaki, Y. (1994) Direct
determination of NotI cleavage site in the genomic DNA
of adult mouse kidney and human trophoblast using whole-
range restriction landmark genomic scanning. DNA
Research, 1:239-243.

125-39.


Lemkin P.F.; JM Myrick; Y Lakshmanan; MJ Shue; JL Patrick; PV Hornbeck; G Thornwall; AW Partin (1999) Exploratory Groupware for Quantitative Electrophoretic Gel Analysis over the Internet - WebGel. In Prep.


Mitra, A.B.; Murty, V.V.V.S.; Pratap, M.; Sodhani, P.; Chaganti, R.S.K. (1994) ERB B2 (HER2/neu) oncogene is


Sheikh, M.S.; Shao, Z.M.; Li, X.S.; Ordonez, J.V.; Conley, B.A.; Wu, S.; Dawson, M.I.; Han, Q.X.; Chao, W.R.;


APPENDIX

Source code for Color Overlay Macro

Based on a program written by Wayne Rasband

```plaintext
var {global variables}
x1,x2,y1,y2,width,height,LineWidth,firstAngle,dx,dy,secondAngle,deltaAngle,ThisAngle,SinA:real;
x,y,ImageToRegister, RegistrationImage:integer;
OrigTop,OrigLeft,OrigWidth,OrigHeight,stage:integer;

Procedure GetSinA;
var
Hypot:real;
begin
Hypot:=sqrt(sqr(dx)+sqr(dy)); {gives the hypotenuse of the angle}
SinA:=dy/Hypot;
ThisAngle:=rAngle[rCount]; {grab the angle for this image}
if ((SinA=0) and (dx<0)) then ThisAngle:=0 else if
((SinA<0) or ((SinA=0) and (dx>0))) then
ThisAngle:=(180-rAngle[rCount]) else
ThisAngle:=(rAngle[rCount])*(-1));
end;

Procedure FindLineMidPoint;
begin
GetLine(x1,y1,x2,y2,LineWidth); {get x,y coords. of line ends}
if LineWidth = 0 then begin
putMessage('You must draw a line from left to right between
two landmarks.');
exit;
end;
measure;
dx:=x1-x2;
dy:=y1-y2;
GetSinA;
end;

macro 'Mark First Image [4]';
var
regpicname:string;
begin
SetOptions('Angle');
regpicname:=WindowTitle;
SelectAll;
Duplicate('Register to:' regpicname);
RestoreRoi;
if WindowTitle='Camera' then begin
```

157
PutMessage('The active window cannot be "Camera".'); exit;
end;
RegistrationImage:=pidnumber;
FindLineMidPoint;
firstAngle:=ThisAngle;
PutMessage('Now go to the image to be aligned to this one.');
end;

macro 'Mark and Rotate Second Image [5]'
begin
FindLineMidPoint;
secondAngle:=ThisAngle;
deltaAngle:=(firstAngle-secondAngle);
SelectAll;
ScaleAndRotate(1,1,deltaangle);
SetPicName('Image to Register');
SelectAll;
GetROI(OrigLeft,OrigTop,OrigWidth,OrigHeight);
if OrigWidth=0 then begin PutMessage('Use Select All or create an ROI on the image you want registered.'); exit;
end;
ImageToRegister := PidNumber;
ShowPasteControl;
Copy;
SelectPic(RegistrationImage);
Showmessage('Run the next macro.');</Showmessage>
RestoreRoi;
end;

macro 'Register the Images [6]'
begin
Showmessage('Move the ROI to the visual registration point.');
Paste;
SetOption;
DoXOr;
end;

macro 'Export LUT [E]'
{Copies the current look-up table to a text window.}
var
i:integer;
v:real;
tab:string;
begin
RequiresVersion(1.54);
NewTextWindow('LUT',200,400);
tab:=chr(9);
for i:=0 to 255 do
WriteIn(i:4,tab,RedLut[i]:4,tab,GreenLut[i]:4,tab,BlueLut[i]:4);
end;

macro 'Import Text LUT';
{
Imports a LUT stored as three column (red, green, blue) text file. If there are four columns then the first column is assumed to contain sequence numbers and is ignored.

```
var
  i,r,g,b, width, height, start, row:integer;
begin
  RequiresVersion(1.53);
  SetImport('Text');
  Import('');
  GetPicSize(width,height);
  if width=3 then begin
    r:=0;
    g:=1;
    b:=2
  end else if width=4 then begin
    r:=1;
    g:=2;
    b:=3
  end else begin
    PutMessage('The text file must have either 3 or 4 columns.');
    exit;
  end;
  if height=255 then
    start:=1
  else if height=256 then
    start:=0
  else begin
    PutMessage('The text file must have either 255 or 256 rows.');
    exit;
  end;
  i:=start;
  row:=0;
  repeat
    RedLut[i]:=GetPixel(r,row);
    GreenLut[i]:=GetPixel(g,row);
    BlueLut[i]:=GetPixel(b,row);
    if (i mod 10) = 0 then UpdateLUT;
    i:=i+1;
    row:=row+1;
  until row>=height;
  UpdateLUT;
end;

macro 'Log Tranform';
var
  i,v:integer;
  scale:real;
BEGIN
  scale := 255.0 / ln(255.0);
  for i:=1 to 254 DO begin
    v := 255-round(ln(i) * scale);
    RedLUT[i]:=v;
    GreenLUT[i]:=v;
    BlueLUT[i]:=v;
  end;
  UpdateLUT;
end;
```
macro 'Gamma Transform... [G]';
var
  i,v:integer;
  n,mode,min,max:integer
  gamma,mean:real;
begin
  gamma:=GetNumber('Gamma(0.1-3.0):',2);
  measure;
  GetResults(n,mean,mode,min,max);
  ShowMessage('min=',min:1,' max=',max:1);
  for i:=1 to 254 Do begin
    if (i>min) and (i<max) then v:=exp(gamma*ln((i-min)/(max-min)))*255
    else begin
      if i<=min then v:=0 else v:=255;
    end;
    RedLUT[i]:=255-v;
    GreenLUT[i]:=255-v;
    BlueLUT[i]:=255-v;
  end;
  UpdateLUT;
end;

macro 'Square Transform';
var
  i,v:integer;
  sqr255:real;
BEGIN
  sqr255:=sqr(255.0);
  for i:=1 to 255 Do begin
    v:=round(sqr(i)*255.0/sqr255);
    RedLUT[255-i]:=v;
    GreenLUT[255-i]:=v;
    BlueLUT[255-i]:=v;
  end;
  UpdateLUT;
END.

macro 'Parabolic Transform';
{ Generates a parabolic LUT}
var
  i,y:integer;
  scale:real;
begin
  scale:=1;
  for i:= 1 to 254 do begin
    y:= (i-127)*(i-127)*scale/64.25;
    if y > 255 then y:=255;
    RedLUT[i]:=y;
    GreenLUT[i]:= y;
    BlueLUT[i]:=y;
  end;
  UpdateLUT;
macro 'Square Root Tranform';
var
  i,v:integer;
  sqrt255:real;
BEGIN
  sqrt255:=sqrt(255.0);
  for i:=1 to 255 DO begin
    v:=round(sqrt(i)*255.0/sqrt255);
    RedLUT[255-i]:=v;
    GreenLUT[255-i]:=v;
    BlueLUT[255-i]:=v;
  end;
  UpdateLUT;
END;

macro 'Color Merge Two Images [C]';
{
  Merges a "red" image and a "green" image to create a composite color image. The macro does this by scaling both images to 0-15, multiplying the second by 16, creating a single 8-bit by ORing the two 4-bit images, and then generating a custom red and green LUT to display the composite image.
}
var
  i,w1,w2,h1,h2,merged:integer;
begin
  SaveState;
  if nPics<>2 then begin
    PutMessage('This macro operates on exactly two images.');
    exit;
  end;
  SelectPic(1);
  GetPicSize(w1,h1);
  SelectPic(2);
  GetPicSize(w2,h2);
  if (w1<>w2) or (h1<>h2) then begin
    PutMessage('The two images must have the same width and height.');
    exit;
  end;
  SetNewSize(w1,h2);
  MakeNewWindow('Merged');
  merged:=PicNumber;
  SelectPic(1);
  SelectAll;
  Copy;
  SelectPic(merged);
  Paste;
  SelectAll;
  MultiplyByConstant(1/16);
  ChangeValues(0,0,1);
  ChangeValues(16,16,15);
  SelectPic(2);
  SelectAll;
Duplicate('Temp');
MultiplyByConstant(1/16);
ChangeValues(16,16,15);
MultiplyByConstant(16);
ChangeValues(0,0,1);
SelectAll;
Copy;
SelectPic(merged);
Paste;
DoOr;
for i:=0 to 255 do begin
  RedLut[i]=(i mod 16)*16;
  GreenLut[i]=(i div 16)*16;
  BlueLut[i]:=0;
end;
UpdateLut;
SelectPic(nPics);
Dispose; {Temp}
RestoreState;
end;

Macro 'measure intensity[m]'; {include the ref}
var
  left, top, width, height, PicID, i: integer;
  name: string;
begin
  picID:=PicNumber;
  getRoi(left, top, width, height);
  if width=0 then
    begin
      putMessage('requires a selection');
      exit;
    end;
  for i:=1 to nPics do begin
    selectPic(i);
    restoreRoi;
    setOptions('Mean');
    {if the measurement starts from a phase contrast frame, two lines below marked with ** should be made active}
    {if PicNumber<>picID then begin} {**
      measure;
      markSelection;
      name:=windowTitle;
      selectWindow('My Results');
      writeln (rCount:1:0, chr(9), rMean[rCount]:1:3, chr(9), name);
    } {end;} {**}
  end;
  selectWindow('My Results');
  selectPic(PicID);
  restoreRoi;
end;

Macro 'Draw Arrow [A]' {Draws an arrow based on the current straight line selection.}
var
  size, angle, dx, dy, pi, theta: real;
begin
  size:=12;  \{pixels\}
  angle:=20;  \{degrees\}
  pi:=3.14159;
  GetLine(x1,y1,x2,y2,LineWidth);
  if x1<0 then begin
    PutMessage(\'Use the line tool(straight) to select a line first.'
    );
    exit;
  end;
  MoveTo(x1,y1);
  LineTo(x2,y2);
  KillRoi;
  GetPicSize(width,height);
  y1:=height-y1;
  y2:=height-y2;
  if LineWidth>1 then size:=size*LineWidth*0.5;
  angle:=(angle/180)*pi;
  dx:=x1-x2;
  dy:=y1-y2;
  if dx=0 then begin
    if dy>=0 then theta:=pi/2 else theta:=3/2*pi
  end else begin
    theta:=arctan(dy/dx);
    if dx<0 then theta:=theta+pi;
  end;
  moveto(x2,height-y2);
  lineto(x2+size*cos(theta+angle),height-(y2+size*sin(theta+angle)));
  moveto(x2,height-y2);
  lineto(x2+size*cos(theta-angle),height-(y2+size*sin(theta-angle)));
end;

\{Macros for performing simple visual translational image registration\}
\{Modified 3/1/93\}

var
  ImageToRegister, RegistrationImage:integer;
  OrigTop,OrigLeft,OrigWidth,OrigHeight,stage:integer;

macro 'Define Image to Register [1]'\;
begin
  RequiresVersion(1.48);
  If npics <> 2 then begin
    Putmessage(\'You must have exactly two images open. One to register to,\n    the other to register.\');
    exit;
  end;
  GetROI(OrigLeft,OrigTop,OrigWidth,OrigHeight); if OrigWidth=0 then begin
    PutMessage(\'Use Select All or create an ROI on the image you want\n    registered.\');
    exit;
  end;
  ImageToRegister := PicNumber;
  ShowPasteControl;
  Copy;

NextWindow;
RegistrationImage:=PicNumber;
Showmessage('Run the second macro.');
RestoreRoi;
Stage:=1;
end;

macro 'Register [2]';
begin
if stage<>1 then begin
PutMessage('Use the Define Image to Register macro first.'); exit;
end;
Showmessage('Move the ROI to the visual registration point and run the third macro.');
Paste;
SetOption; DoXOr;
stage:=2;
end;

macro 'Create registered image [3]';
var
MovedTop,MovedLeft,MovedWidth,MovedHeight:integer;
OrigMovedTop,OrigMovedLeft,OrigMovedWidth,OrigMovedHeight:integer;
NewTop,NewLeft,NewWidth,NewHeight:integer; width, height:integer;
BEGIN
if stage<>2 then begin
PutMessage('Use the Define and Register macros first.'); exit;
end;
Undo;
GetPicSize(width, height);
SelectPic(RegistrationImage);
RestoreRoi;
GetROI(Movedleft,Movedtop,Movedwidth,Movedheight); if MovedTop < 0 then begin
OrigMovedTop := OrigTop - Movedtop;
NewTop := 0;
OrigMovedHeight := OrigHeight + MovedTop; NewHeight := OrigHeight + MovedTop;
end else begin
OrigMovedTop := OrigTop;
NewTop := MovedTop;
OrigMovedHeight := MovedHeight;
NewHeight := MovedHeight;
end;
if Movedleft < 0 then begin
OrigMovedLeft :=OrigLeft-Movedleft;
NewLeft := 0;
end else begin
OrigMovedLeft := OrigLeft;
NewLeft := MovedLeft;
end;
if (Movedleft + MovedWidth) > width then begin
OrigMovedwidth := width-Movedleft;
NewWidth := width - Movedleft;
end;
else begin
OrigMovedWidth := MovedWidth;
NewWidth := MovedWidth;
end;
SelectPic(ImageToRegister);
KillROI;
MakeROI (OrigMovedLeft, OrigMovedTop, OrigMovedwidth, OrigMovedheight);
Copy;
SetNewSize(width, height);
MakeNewWindow('Registered');
MakeROI(Newleft, Newtop, Newwidth, Newheight); Paste;
Showmessage('The ROI portion of your original image has been copied to
the registered image.');
END;

macro 'Draw Grid...';
var
  x, y, xinc, yinc, width, height: integer;
begin
  GetPicSize(width, height);
  xinc := GetNumber('Horizontal Spacing:', 16);
  yinc := GetNumber('Vertical Spacing:', xinc);
  x := 0;
  y := 0;
  repeat
    x := x + xinc;
    y := y + yinc;
    moveto(0, y);
    lineto(width, y);
    moveto(x, 0);
    lineto(x, height);
  until (x > width) and (y > height);
end;
macro 'Position fixed size ROI 700x735 [Q]';
var width, height, x, y: integer;
begin
  width := 700; height := 735;
  repeat
    GetMouse(x, y);
    MakeRoi(x-width/2, y-height/2, width, height);
    DrawBoundary;
    Undo;
  until button;
end;
macro 'Position fixed size ROI 325x325 [I]';
var width, height, x, y: integer;
begin
  width := 325; height := 325;
  repeat
    GetMouse(x, y);
    MakeRoi(x-width/2, y-height/2, width, height);
    DrawBoundary;
    Undo;
  until button;
end;
macro 'Position fixed size ROI 50x50 [O]';
var width,height,x,y:integer;
begin
    width:=50; height:=50;
    repeat
        GetMouse(x,y);
        MakeRoi(x-width/2,y-height/2,width,height);
        DrawBoundary;
        Undo;
    until button;
end;
macro 'Position fixed size ROI 100X100 [X]';
var width,height,x,y:integer;
begin
    width:=100; height:=100;
    repeat
        GetMouse(x,y);
        MakeRoi(x-width/2,y-height/2,width,height);
        DrawBoundary;
        Undo;
    until button;
end;
macro 'Position fixed size ROI 125X125 [Z]';
var width,height,x,y:integer;
begin
    width:=125; height:=125;
    repeat
        GetMouse(x,y);
        MakeRoi(x-width/2,y-height/2,width,height);
        DrawBoundary;
        Undo;
    until button;
end;
macro 'Position fixed size ROI 238x238 [S]';
var width,height,x,y:integer;
begin
    width:=238; height:=238;
    repeat
        GetMouse(x,y);
        MakeRoi(x-width/2,y-height/2,width,height);
        DrawBoundary;
        Undo;
    until button;
end;
macro 'Position fixed size ROI 175x175 [D]';
var width,height,x,y:integer;
begin
    width:=175; height:=175;
    repeat
        GetMouse(x,y);
        MakeRoi(x-width/2,y-height/2,width,height);
        DrawBoundary;
        Undo;
    until button;
end;
macro 'Define Upper Left [7]';
var
  x1, y1, x2, y2, LineWidth: integer;
begin
  GetLine(x1, y1, x2, y2, LineWidth);
  if x1 < 0 then begin
    PutMessage('Click with line selection tool to define upper left corner of ROI.');
    exit;
  end;
  RoiLeft := x1 + (x2 - x1)/2;
  RoiTop := y1 + (y2 - y1)/2;
end;

macro 'Define Lower Right and Create ROI [8]';
var
  x1, y1, x2, y2, LineWidth: integer;
begin
  GetLine(x1, y1, x2, y2, LineWidth);
  if x1 < 0 then begin
    PutMessage('Click with line selection tool to define lower right corner of ROI.');
    exit;
  end;
  RoiRight := x1 + (x2 - x1)/2;
  RoiBottom := y1 + (y2 - y1)/2;
  if (RoiLeft = RoiRight) and (RoiTop = RoiBottom) then begin
    PutMessage('Upper left and bottom right are the same.');
    exit;
  end;
  MakeRoi(RoiLeft, RoiTop, RoiRight - RoiLeft, RoiBottom - RoiTop)
end;
Restriction Landmark Genomic Scanning (RLGS) was used in this dissertation to detect genetic changes such as loss of heterozygosity, amplification, and deletion, in a variety of tumor and cell line models. New image processing and analysis software packages and methods were also used to extract useful data from the complex RLGS spot patterns. The first study employed a novel color hybridization Macro to detect genetic alterations in Methylnitrosourea (MNU) induced rat mammary tumors that were also treated with a punitive chemopreventive agent 4-Hydroxyphenylretinamide (4HPR). Results indicated that the MNU group had 35 altered spots compared to the untreated group and the NMU-4HPR contained fewer genetic alterations when compared to the control group. In the cervical cancer project, the group employed the same color hybridization macro in concert with an image warping system to detect alterations between different types premalignant, cancerous, and normal tissue. Analysis indicated 20 alterations between the premalignant and cancerous groups and the cervical normal tissue control. The third study was responsible for mapping the RLGS alterations between K-ras transformed NIH3-3T3 and NIH-3T3 control cell lines. This study utilized a new
software package called GELLEB II+ to perform the image processing and analysis of the RLGS spot patterns. The software package and a relational database program, called COG, was extremely effective in the Lapacian segmentation, spot matching, and subsequent data mining of the RLGS gels. Results showed 12 alterations between the experimental and control groups. Two spots that were amplified in the K-ras transformed NIH3-3T3 group were statistically significant ($p < 0.01$) With this in mind, these alterations may represent key events in the transformation process of K-ras. In the human bronchial epithelial High-LET Alpha Particle study, RLGS and the GELLAB II+ system in concert with Java based Webgel showed a total of 13 alterations between the non-High-LET Alpha Particle group in the irradiated samples.

RLGS was shown to be an effective tool in delineating genetic alterations in a variety of models when used with specialized image processing and analysis software packages.