GENOMIC INSTABILITY-INDICATED MECHANISMS OF EARLY CARCINOGENESIS FOLLOWING HEPATOCARCINOGEN TREATMENT

A Dissertation in
Molecular Toxicology

by

Sarah Jean Byers Gramling

© 2009 Sarah Jean Byers Gramling

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

December 2009
The dissertation of Sarah Jean Byers Gramling was reviewed and approved* by the following:

Gary A. Clawson  
Professor of Pathology, and Biochemistry and Molecular Biology  
Dissertation Advisor  
Chair of Committee

Kristin A. Eckert  
Professor of Pathology, and Biochemistry and Molecular Biology

Harriet Isom  
Distinguished Professor of Microbiology and Immunology; Professor of Pathology

Teresa Wood  
Adjunct Professor of Neural and Behavioral Sciences

Jong Yun  
Associate Professor of Pharmacology  
Co-chair, Intercollege Graduate Degree Program in Molecular Toxicology

*Signatures are on file in the Graduate School
ABSTRACT

Carcinogenicity testing is perhaps one of the most challenging aspects in drug development. Carcinogenicity tests not only provide the basis for human risk assessment and regulatory policy on synthetic drugs, but also for all compounds, including “natural” compounds, that are carcinogenic. Current assays, both in vitro and in vivo, rely on the induction of tumors or the detection of DNA-based mutational events (such as the Ames test) to indicate the potential for the studied agent to cause cancer. These studies are often used in conjunction with one another to assess carcinogenicity of a particular compound. While providing a certain level of support for the carcinogenic potential of compounds, these studies prove inadequate for the assessment of non-mutagenic carcinogens and provide little to no data as to a compound’s mechanism(s) of action.

Regions of repetitive DNA sequence, known as microsatellites, and are particularly prone to genetic alteration, can serve as indicators of genomic instability, which has been implicated in the process of tumorigenesis. Recent studies of the sequenced genome have found many microsatellites in regions of DNA that do not directly code for proteins. Often they are found in untranslated regions (UTRs), namely the 5’ and 3’ UTRs, which surround protein-coding sequences and convey regulatory function, but microsatellites have also been found within introns. The presence of these repetitive sequences in non-coding DNA, has recently been supported by evidence indicating that a subset of non-coding regions are in fact transcribed, and that the resulting RNAs may function
as “riboregulators”, suggesting a functional importance of these regions. Assessment of the effects of compounds on these microsatellite regions may provide a better understanding about the underlying mechanistic process of carcinogenesis.

In this dissertation, I present the development of a novel dual-reporter model system aimed at detecting early genomic instability, using frameshift-prone microsatellites as indicators, to locate cells that are targets of non-mutagenic carcinogens. In addition, I focus this study on the liver, based on its recognition as the “first-pass” organ in the study of toxicology.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... x
LIST OF TABLES .............................................................................................................. xii
LIST OF ABBREVIATIONS ............................................................................................. xiii
ACKNOWLEDGEMENTS ................................................................................................. xviii

Chapter 1 Introduction and Rationale ............................................................................. 1
Chapter 2 Literature Review ........................................................................................... 4
  2.1. Toxicology and the Liver ....................................................................................... 4
    2.1.1. Hepatocyte ........................................................................................................ 6
    2.1.2. Metabolism ........................................................................................................ 6
      2.1.2.1. Phase I Metabolism-Cytochrome P450s ...................................................... 7
      2.1.2.2. Phase II Metabolism .................................................................................. 9
    2.1.3. Hepatic Foci ...................................................................................................... 10
  2.2. Microsatellites ........................................................................................................ 11
    2.2.1. Genomic Instability ........................................................................................ 13
  2.3. Detection Assays-Carcinogenicity Estimation .......................................................... 15
    2.3.1. Mouse Model Incorporating Plasmid-Based Microsatellite Sequence .............. 17
  2.4. Oxidative Stress ...................................................................................................... 18
  2.5. Epigenetic Regulation ............................................................................................ 22
    2.5.1. Methylation ..................................................................................................... 24
    2.5.2. Genomic Imprinting ....................................................................................... 25
    2.5.3. Effects of Methylation on RNA ..................................................................... 26
    2.5.4. Histone Modification ..................................................................................... 28
  2.6. Non-Coding Genes ............................................................................................... 28
2.6.1. Microsatellites in Non-Coding Genes ........................................ 29
2.6.2. RNA Interference (RNAi) ....................................................... 30
2.6.3. RNAi Expression in Hepatocarcinogenesis ................................ 33
2.6.4. Alternate Riboregulation ....................................................... 35
2.7. RNA Oxidative Damage ............................................................ 39
2.8. New Model of Non-Genotoxic Hepatocarcinogenesis-Altered Regulation ........................................................................ 42
2.9. Potential Hepatocarcinogens ...................................................... 45
  2.9.1. Thioacetamide ........................................................................ 45
    2.9.1.1. Metabolism ...................................................................... 45
    2.9.1.2. Carcinogenicity and Mutagenicity .................................... 46
  2.9.2. Acetaminophen ...................................................................... 46
    2.9.2.1. Metabolism ...................................................................... 47
    2.9.2.2. Carcinogenicity and Mutagenicity .................................... 47
  2.9.3. Aflatoxin B1 .......................................................................... 48
    2.9.3.1. Metabolism ...................................................................... 48
    2.9.3.2. Carcinogenicity and Mutagenicity .................................... 49
  2.9.4. Nodularin .............................................................................. 49
    2.9.4.1. Metabolism ...................................................................... 50
    2.9.4.2. Carcinogenicity and Mutagenicity .................................... 50

Chapter 3 Microsatellite-Containing Bicistronic Reporter Constructs: Generation and Validation ......................................................... 53
  3.1. Introduction ................................................................................ 53
  3.2. Cloning of pIβE Plasmid ............................................................ 54
    3.2.1. Method ................................................................................. 54
    3.2.2. Function ............................................................................... 55
Chapter 3

3.2.3. Mutation of pIβE ................................................................. 55

3.3. Inclusion of Microsatellite Repeats -pIβE-ntR Vectors ................. 57

3.3.1. Di-nucleotide Repeats ................................................................ 57

3.3.1.1. Method ........................................................................ 58

3.3.1.2. Sequencing .................................................................... 60

3.3.2. Mono-nucleotide Repeats ....................................................... 61

3.3.2.1. Method ........................................................................ 62

3.3.2.2. Sequencing .................................................................... 64

3.3.3. Function of pIβE-ntR Vectors ............................................... 65

3.4. Incorporation of Liver-specific Enhancer/Promoter -pαBE-ntR Vectors ........................................................................ 65

3.4.1. Design and Function of EALB/Pα1AT ...................................... 65

3.4.1.1. Activity of Chimeric Enhancer/Promoter in Cell Culture ........ 66

3.4.1.2. In Vivo-Mouse Tail Vein Injection Kinetic Activity Study ....... 67

3.5. EALB/Pα1AT-EGFP/pVAX1 (EPE-pVAX1) .................................. 68

3.6. CMV Replacement with EALB/Pα1AT in "ntR" Vectors ................. 70

3.6.1. Function ............................................................................ 71

3.7. Discussion .............................................................................. 72

Chapter 4

Utility of Reporter Constructs for Assessing Effects of Hepatocarcinogen Treatment in Murine Liver Cells .......... 73

4.1. Introduction ............................................................................ 73

4.2. Methods .................................................................................. 75

4.2.1. Mouse Metabolism Analysis .................................................. 75

4.2.2. Microsatellite-Based Frameshift Reporter Cell Line Design ...... 75
4.2.3. Analysis and Characterization of Morphology Changes in NMuLi Cell Line ................................................................. 76
4.2.4. Immunohistochemical Analysis ................................................................................................................................. 77
4.2.5. Establishment of LD_{50}s in NMuLi Cells .................................................................................................................. 77
4.2.6. Measurement of ROS Generation ............................................................................................................................ 78
4.2.7. Chemicals ................................................................................................................................................................... 78
4.2.8. Compound Treatments ............................................................................................................................................... 79
4.2.9. Flow Cytometry ....................................................................................................................................................... 79
4.3. Results and Discussion ................................................................................................................................................ 80
4.3.1. Gene Expression Profile of NMuLi Cells .................................................................................................................... 80
4.3.2. Establishment of LD_{50}s ................................................................. 82
4.3.3. Oxidative Stress Induction ........................................................................................................................................ 84
4.3.4. Morphology Changes in NMuLi Cell Line .................................................................................................................. 86
4.3.5. Frameshift Detection .................................................................................................................................................. 90
4.3.6. Frameshift Detection in pαβE-ntR Cell Lines ........................................................................................................... 93
4.3.7. Closer Examination of Flow Cytometry Results .................................................................................................... 99
4.4. Conclusion ..................................................................................................................................................................... 103

Chapter 5 Development of Transgenic Mouse Model for Determining In Vivo Effects of Hepatocarcinogen Treatment .......... 105
5.1. Introduction ................................................................................................................................................................. 105
5.2. Materials and Methods ............................................................................................................................................... 107
5.2.1. Transgene Purification ............................................................................................................................................ 107
5.2.2. Creation of EaβM_{OUT} Mouse Line ......................................................................................................................... 108
5.2.3. Mouse Characterization .......................................................................................................................................... 108
5.2.3.1. DNA Analysis ....................................................................................................................................................... 108
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.3.2</td>
<td>RNA Analysis</td>
<td>109</td>
</tr>
<tr>
<td>5.2.3.3</td>
<td>Protein Analysis</td>
<td>110</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Mutation Detection</td>
<td>111</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>112</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Design of Transgene</td>
<td>112</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Identification of Founder Mouse</td>
<td>113</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Characterization of EαβM&lt;sub&gt;OUT&lt;/sub&gt; Mouse Line</td>
<td>115</td>
</tr>
<tr>
<td>5.3.3.1</td>
<td>Presence of Transgene DNA</td>
<td>115</td>
</tr>
<tr>
<td>5.3.3.2</td>
<td>Presence of GFP Protein</td>
<td>116</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Mutation Detection</td>
<td>118</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Further Characterization of EαβM&lt;sub&gt;OUT&lt;/sub&gt; Mouse Line</td>
<td>119</td>
</tr>
<tr>
<td>5.3.5.1</td>
<td>RNA Analysis</td>
<td>119</td>
</tr>
<tr>
<td>5.3.5.2</td>
<td>Confirmation of Protein Analysis</td>
<td>122</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>124</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary</td>
<td>130</td>
</tr>
<tr>
<td>6.1.1</td>
<td>pαβE-ntR Cell Lines</td>
<td>131</td>
</tr>
<tr>
<td>6.1.2</td>
<td>EαβM&lt;sub&gt;OUT&lt;/sub&gt; Transgenic Mouse Line</td>
<td>133</td>
</tr>
<tr>
<td>6.2</td>
<td>Future Directions</td>
<td>135</td>
</tr>
<tr>
<td>6.2.1</td>
<td>pαβE-ntR Cell Lines</td>
<td>135</td>
</tr>
<tr>
<td>6.2.2</td>
<td>EαβM&lt;sub&gt;OUT&lt;/sub&gt; Transgenic Mouse Line</td>
<td>137</td>
</tr>
<tr>
<td>6.3</td>
<td>Conclusion</td>
<td>137</td>
</tr>
<tr>
<td>Appendix</td>
<td>Mouse Metabolism Gene Profile of NMuLi Cell Line</td>
<td>139</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>142</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2-1: General schematic of a lobule, the functional unit of the liver

Figure 2-2: ROS release from CYP catalysis

Figure 2-3: Structural similarity of Nodularin and Microcystin-LR

Figure 3-1: Vectors used in creation of bicistronic vector pIβE

Figure 3-2: Sequence of bicistronic vector pIβE-mutKpn

Figure 3-3: Primers designed to incorporate di-nucleotide repeat sequences into plasmid pIβE-mutKpn

Figure 3-4: Sequence of pIβE-DIN

Figure 3-5: Sequence of pIβE-DOUT

Figure 3-6: Primers designed to incorporate mono-nucleotide repeat sequences into plasmid pIβE-mutKpn

Figure 3-7: Sequence of pIβE-MIN

Figure 3-8: Sequence of pIβE-MOUT

Figure 3-9: Chimeric liver-specific enhancer/promoter E_ALB/P_α1AT

Figure 3-10: Activity of E_ALB/P_α1AT in liver and non-liver derived cell lines

Figure 3-11: Kinetic activity of E_ALB/P_α1AT in vivo

Figure 3-12: Construction of EPE-pVAX1

Figure 3-13: Promoter activity test in HepG2 and 293T cells

Figure 3-14: Microsatellite repeat sequences of pαβE-ntR plasmid sequences

Figure 4-1: Dose response curve for hydrogen peroxide

Figure 4-2: Induction of ROS in NMuLi cells
Figure 4-3: Distinct morphologies observed from stably-selected NMuLi cells transfected with pαβE-MIN

Figure 4-4: Hydrogen peroxide-induced frameshifts in 293T cells

Figure 4-5: Anti-β-gal antibody validation

Figure 4-6: Frameshift mutations in NMuLi cell lines expressing pαβE-ntROUT

Figure 4-7: Response of treatments in NMuLi cell lines expressing pαβE-ntROUT

Figure 4-8: Representative double-fluorescence peak

Figure 4-9: Frameshift mutations from EGFP-fluorescence-enhanced cells

Figure 4-10: Response of treatments in EGFP-fluorescence-enhanced NMuLi cell lines expressing pαβE-ntROUT

Figure 5-1: EαβMOUT transgene

Figure 5-2: PCR screening analysis for founder pup identification

Figure 5-3: PCR-based copy number analysis of EαβMOUT DNA

Figure 5-4: GFP protein expression in EαβMOUT tissues

Figure 5-5: Tissue Expression of GFP-containing cells

Figure 5-6: Frame-shifted β-gal positive cell in liver tissue

Figure 5-7: RT-QPCR analysis

Figure 5-8: Suspected GFP protein
LIST OF TABLES

Table 4-1: Phase I and II gene expression in NMuLi cells................................. 81
Table 4-2: LD<sub>50</sub> values for compounds in NMuLi cells .............................. 83
Table 4-3: Enzymatic substrates for β-galactosidase ........................................ 92
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>%</td>
<td>percent</td>
</tr>
<tr>
<td>λ</td>
<td>lambda</td>
</tr>
<tr>
<td>4-HNE</td>
<td>4-hydroxynononal</td>
</tr>
<tr>
<td>8OG</td>
<td>8 hydroxyguanine</td>
</tr>
<tr>
<td>8OdG</td>
<td>8 hydroxydeoxyguanine</td>
</tr>
<tr>
<td>AB&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Aflatoxin B&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>ADAR</td>
<td>Adenosine Deaminase Acting on RNA</td>
</tr>
<tr>
<td>AP</td>
<td>Acetaminophen</td>
</tr>
<tr>
<td>ARE</td>
<td>AU-rich Element</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>β-APP</td>
<td>Beta Amyloid Precursor Protein</td>
</tr>
<tr>
<td>β-gal</td>
<td>Beta-galactosidase</td>
</tr>
<tr>
<td>cabRNA</td>
<td>centromere-associated binding RNA</td>
</tr>
<tr>
<td>casRNA</td>
<td>centromere-associated silencing RNA</td>
</tr>
<tr>
<td>CCK-8</td>
<td>Cell Counting Kit-8</td>
</tr>
<tr>
<td>CENP-A</td>
<td>Centromere Protein A</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese Hamster Ovary</td>
</tr>
<tr>
<td>CIN</td>
<td>Chromosomal Instability</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CpG</td>
<td>Cytosine-phosphate-Guanine</td>
</tr>
<tr>
<td>CYP</td>
<td>Cytochrome P</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>δ</td>
<td>delta</td>
</tr>
<tr>
<td>DCF</td>
<td>Dichlorofluorescin</td>
</tr>
<tr>
<td>DCFH-DA</td>
<td>2',7'-Dichlorofluorescin-diacetate</td>
</tr>
<tr>
<td>DEN</td>
<td>Diethylnitrosamine</td>
</tr>
<tr>
<td>D&lt;sub&gt;IN&lt;/sub&gt;</td>
<td>in-frame di-nucleotide repeat</td>
</tr>
<tr>
<td>di-nt</td>
<td>di-nucleotide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>D&lt;sub&gt;OUT&lt;/sub&gt;</td>
<td>out-of-frame di-nucleotide repeat</td>
</tr>
<tr>
<td>dsRNA</td>
<td>double stranded RNA</td>
</tr>
<tr>
<td>ε</td>
<td>epsilon</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>E&lt;sub&gt;ALB&lt;/sub&gt;</td>
<td>Albumin enhancer</td>
</tr>
<tr>
<td>EGFP</td>
<td>Enhanced Green Fluorescent Protein</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>GSH</td>
<td>Glutathione</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione-S-transferase</td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Hydrogen peroxide</td>
</tr>
<tr>
<td>H3K9</td>
<td>Lysine 9 of Histone 3</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular Carcinoma</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HEK</td>
<td>Human Embryonic Kidney</td>
</tr>
<tr>
<td>HNPCC</td>
<td>Hereditary Non-polyposis Colon Cancer</td>
</tr>
<tr>
<td>IGF2</td>
<td>Insulin-like Growth Factor 2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosome Entry Site</td>
</tr>
<tr>
<td>ISB</td>
<td>Intracellular Staining Buffer</td>
</tr>
<tr>
<td>IV</td>
<td>Intravenous</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>KRH</td>
<td>Krebs-Ringer-HEPES</td>
</tr>
<tr>
<td>lacZ</td>
<td>β-galactosidase gene</td>
</tr>
<tr>
<td>LD₅₀</td>
<td>Lethal Dose 50</td>
</tr>
<tr>
<td>MIN</td>
<td>Microsatellite Instability</td>
</tr>
<tr>
<td>MᵢN</td>
<td>in-frame mono-nucleotide repeat</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>mono-nt</td>
<td>mono-nucleotide</td>
</tr>
<tr>
<td>MₒUₜ</td>
<td>out-of-frame mono-nucleotide repeat</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MTD</td>
<td>Maximum Tolerated Dose</td>
</tr>
<tr>
<td>NADP</td>
<td>Nicotinamide Adenine Dinucleotide Phosphate</td>
</tr>
<tr>
<td>NAPQI</td>
<td>N-acetyl-p-benzoquinone Imine</td>
</tr>
<tr>
<td>NCI</td>
<td>National Cancer Institute</td>
</tr>
<tr>
<td>ncRNA</td>
<td>non-coding RNA</td>
</tr>
<tr>
<td>NFkB</td>
<td>Nuclear Factor kappa B</td>
</tr>
<tr>
<td>NMuLi</td>
<td>normal mouse liver cell line</td>
</tr>
<tr>
<td>Nod</td>
<td>Nodularin</td>
</tr>
<tr>
<td>NTP</td>
<td>National Toxicology Program</td>
</tr>
<tr>
<td>ntR</td>
<td>nucleotide repeat</td>
</tr>
<tr>
<td>O₂</td>
<td>molecular oxygen</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>O$_2^*$</td>
<td>Superoxide anion</td>
</tr>
<tr>
<td>'OH</td>
<td>Hydroxylradical</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>OTA</td>
<td>Ochratoxin A</td>
</tr>
<tr>
<td>P$_{\alpha 1 AT}$</td>
<td>$\alpha$-1-antitrypsin promoter</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PLAP</td>
<td>Placental Alkaline Phosphatase</td>
</tr>
<tr>
<td>pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>pol III</td>
<td>RNA polymerase III</td>
</tr>
<tr>
<td>PnP</td>
<td>Polynucleotide Phosphorylase</td>
</tr>
<tr>
<td>PP1</td>
<td>Protein Phosphatase 1</td>
</tr>
<tr>
<td>PP2A</td>
<td>Protein Phosphatase 2A</td>
</tr>
<tr>
<td>QPCR</td>
<td>quantitative PCR</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-interfering Silencing Complex</td>
</tr>
<tr>
<td>RITS</td>
<td>RNA-interfering Transcriptional Silencing complex</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>ROS</td>
<td>Radical Oxygen Species</td>
</tr>
<tr>
<td>RPE</td>
<td>R-Phyceroerythin</td>
</tr>
<tr>
<td>SCE</td>
<td>Sister-chromatid Exchange</td>
</tr>
<tr>
<td>S. pombe</td>
<td>Schizosaccharomyces pombe</td>
</tr>
<tr>
<td>siRNA</td>
<td>small interfering RNA</td>
</tr>
<tr>
<td>SOD</td>
<td>Superoxide Dismutase</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>TAA</td>
<td>Thioacetamide</td>
</tr>
<tr>
<td>TASO</td>
<td>Thioacetamide S-Oxide</td>
</tr>
<tr>
<td>UbB</td>
<td>Ubiquitin B</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated Region</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside</td>
</tr>
<tr>
<td>YB-1</td>
<td>Y-box Binding protein</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

It goes without saying that this journey has been an extended one and there are many people for which I owe sincere thanks for the encouragement they have provided over these years.

I would first like to thank my committee for all their guidance and support. Despite all the pitfalls associated with this project, and there have been many, they have always been incredibly supportive and I will be forever grateful to them for that.

To all of my friends, I can’t possibly name you all, but I thank you.

I owe a special acknowledgement of thanks, to my parents. Throughout my life, you have always provided me with so much love and support. Not only have you always been a constant cheerleading team, you instilled in me a great sense of self-confidence and strength that I will carry for the rest of my life. I will always be thankful and love you so much.

There is one person left to thank, without whom, I would not have wanted to take this adventure. To my husband, Jim, words cannot express how deeply grateful I am to have had you by my side for all these years. You have been my strength when I have stumbled and were there to celebrate all the joys. Thank you so much for everything you have given me. I look forward to the many years we have to come and all that the future holds. I love you so very much, so from the bottom of my heart…Thank You!
Chapter 1

Introduction and Rationale

Doctor Thomas sat over his dinner,
Though his wife was waiting to ring,
Rolling his bread into pellets;
Said, 'Cancer’s a funny thing.'

'Nobody knows what the cause is,
Though some pretend they do;
It’s like some hidden assassin
Waiting to strike at you.'

'Childless women get it.
And men when they retire;
It’s as if there had to be some outlet
For their foiled creative fire'

-Miss Gee (65-76)

This quote by W.H. Auden from his poem entitled “Miss Gee”, written in 1938, illustrates not only the fact that cancer is a threat to all individuals, but also to the elusiveness that continues to plague cancer research today. By using the term ‘assassin’ (line 71), he alludes to the fact that much like an assassin must have a plan of attack, cancer is meticulous in its progression, that is, its growth follows a series of events, and is not the result of a haphazard accident. By incorporating 'hidden', he makes reference to the fact that while some of the events of cancer development have been deciphered, a root cause encompassing
all cancers has yet to be found and thus there is no way to predict when it will ‘strike at you’ (line 72).

Cancer is the number one killer of humans and in 2007 accounted for 13% of all deaths globally, a total of 7.9 million people. While it is encouraging that mortality from cancer has decreased, due in part to early detection and advances in treatment, it provides no impact on incidence rates and thus to the overall burden of cancer on our society.

The National Institutes of Health estimated the cost of cancer in the United States for the year 2007 at $219.2 billion. This total accounts not only for costs incurred from treatment, and funeral/burial arrangements, but also from loss of productivity caused by illness or premature death. This total is expected to approximately double by the year 2030, considering that cancer is primarily a disease of the elderly and the increasing age of the baby boomer population.

Early detection is critical and has been very successful for the diagnosis, and treatment, of breast, colon, and skin cancers; however, cancers can arise from any tissue in the body. In many cases, such as with liver and pancreatic cancers, they go undetected until another problem arises, either directly as a result of the growth of a tumor, or metastasis has occurred leading to problems in other organs/tissues. Only by truly understanding the underlying mechanisms of the cancer process will we one day, if not eradicate, relieve the emotional as well as financial burden of cancer on our society. Complicating this research task, to discover the mechanisms of cancer development and growth, is the interplay between genetics and environmental influences.
It is well studied that certain individuals may have a genetic predisposition toward the development of cancer, in that they are lacking, or have mutations in, genes that are involved in key cellular processes such as growth control\textsuperscript{81, 108, 307} or repair\textsuperscript{182, 221, 236}. Likewise, the external environment is filled with factors (physical,\textsuperscript{119, 292} biological,\textsuperscript{181, 207} and chemical\textsuperscript{130}) that are known to cause cancer. However, it is becoming increasingly evident that these interactions between environmental components and the genetic makeup of an individual that is key to understanding the true nature of diseases,\textsuperscript{91} especially cancer.

Considering chemical interactions, it is very apparent that substances isolated or synthesized for the purposes of medicine play a role in human health; however, there is another, possibly larger, class of substances that plays a role in human disease. It is the mechanisms of action of the latter that is intriguing to the science of toxicology.

In this dissertation, I aim to understand one such mechanism of interaction, namely how chemicals affect, either directly or indirectly, regions of nucleic acid sequence known as microsatellites, on the pathway to carcinogenesis.
Chapter 2

Literature Review

2.1 Toxicology and the Liver

The study of toxicology aims to identify the mechanisms by which chemical interactions affect body tissues/systems. These chemicals are classified as “xenobiotics”, which is a general term for a large group of foreign substances that enter the body. These substances are chemicals found in the body that are not inherent, and may be classified as nutrients, drugs, or toxins. A “toxin” refers to any chemical that has an adverse affect on the integrity of the body.

Each of these groups of chemicals are able to enter the bloodstream in a variety of ways, from direct injection into the bloodstream, as is the case with intravenous (IV) drugs, but also via inhalation, transdermally, and via digestion, whereby they are directed to the liver.

Most of these xenobiotics ultimately wind up in the liver, which is considered to be one of the most essential organs to organism welfare. Upon severe liver injury or removal, an organism will not survive. This is in part due to the lack of metabolic conversion of nutrients, but also to the lack of defense it provides against a constant influx of xenobiotics. As a first-pass organ, recognized for its ability to perform a great number of metabolic functions, the liver further emphasizes its functional importance to the body. That is, all xenobiotics (either nutrients or toxins) within the bloodstream, are initially
filtered through the liver to be metabolized and/or detoxified and, in turn, eliminated from the body as waste.

The human liver is the largest internal organ of the body weighing approximately 3 pounds. It is located in the right center of the body at the top of the abdomen, directly below the diaphragm. The liver consists of an array of histologically-defined units called lobules. Each lobule (Figure 2-1) is a hexagonal structure made up of: portal triads, small regions where the portal vein, hepatic artery, and bile duct triangulate with the central vein; and hepatocytes. Defined here is the classic lobule structure, other “lobules” of the liver, have been described such as the acinus or portal lobule; however, they are functionally-defined and will not be detailed in this discussion.

Figure 2-1:  General schematic of a lobule, the functional unit of the liver.  (Source: Cunningham, C.C. and Van Horn, C.G. Alcohol Research & Health (2003).76)
Blood flow is directed through the lobule from portal triads, carrying blood and bile, through sinusoids into the central vein. It is in the sinusoids, where the hepatocytes come in contact with the blood.

### 2.1.1 Hepatocyte

While there are a variety of cell types (i.e.: oval cells and Kupffer cells) in the liver, hepatocytes account for the majority of liver mass and are the metabolizing/detoxifying units of the liver. As stated previously, each layer of hepatocytes is surrounded by sinusoids. This maximizes the surface area of the hepatocyte and allows for a large amount of blood to be processed, necessary considering the rapid rate of blood flow though the liver. It is estimated that the total time for blood to pass from the portal triad to the central vein is only 8 seconds.

The filtering function of the hepatocyte is by the uptake of chemicals (via passive diffusion or active transport) from the blood plasma and followed by their processing/metabolism in the endoplasmic reticulum (ER) of the hepatocytes. All chemicals undergo the process of normal cellular metabolism, regardless of their status as, beneficial or harmful.

### 2.1.2 Metabolism

Cellular metabolism, occurring in the ER, is generally a two-phase process. Phase I involves a group of enzymes known collectively as cytochrome p450s. These enzymes, specifically located in the endoplasmic reticulum, occur in all
cells, however their concentration is highest in liver hepatocytes. The main role of these enzymes is to convert substances into more polar, soluble products that are able to interact with phase II enzymes, largely responsible for eliminating the substance via conjugation reactions with water-soluble macromolecules, such as glutathione, facilitating their excretion.

### 2.1.2.1 Phase I Metabolism-Cytochrome P450s

Phase I metabolism converts xenobiotics into polar intermediates. This conversion primarily results in oxidation of a chemical by enzymes known as cytochrome P450s (CYP450s).\(^{124}\)

CYP450s are thiolated heme enzymes, in that they are characterized by the presence of a central iron atom, and thus able to transfer oxygen to a particular previously bound substrate. The oxygenation of a xenobiotic is by electron transfer through the oxidation of NADPH to NADP\(^+\) and reduction of a CYP450 by cytochrome reductase, in the presence of molecular oxygen (O\(_2\)).

CYP450s are fairly substrate specific and are grouped as families and subfamilies based upon sequence similarity of 40% and 55%, respectively. The nomenclature for this group of enzymes now follows a strict naming scheme. This scheme begins with CYP, followed by a number indicating family, a letter representing subfamily, followed by another number for the individual isoform of the gene, (i.e. CYP1A2).

The families of CYP1, 2 and 3 are the most common of all the CYP450s for the conversion of xenobiotics.\(^{135}\) Despite the nomenclature for CYP450s being
standardized within individual species, it has not been standardized between species. Nelson et al.\textsuperscript{220} presented a comparison between human and mouse CYP450 genes, identifying 57 human genes and 102 mouse genes. The authors attribute the difference in quantity to gene duplication that occurred after the evolutionary split between humans and mice. Genes denoted by the same numbers and letters, due to sequence identity, are orthologs (i.e. have the same function) between the two species; however, it is still uncertain as to which of the remaining CYP450s act in orthologous manners between these species. This discrepancy presents a challenge for the study of toxicology in the extrapolation of mouse data and correlation to human toxicological significance.

Xenobiotics also have the potential to alter levels of the CYP450s. These alterations can have major effects on the metabolism of other chemicals as well, as can be seen by the induction of CYP2E1 by alcohol\textsuperscript{306} and its effects on acetaminophen toxicity. By increasing the amount of enzyme present, more metabolic reactions occur. Furthermore, certain xenobiotics, most notably phenobarbital, not only increase the levels of specific CYP enzymes, but can also induce ER proliferation allowing for a general increase in the expression of all CYP450s.\textsuperscript{323} These increases can be harmful since some of the byproducts from phase I can be bioactivated, thus producing toxic species, when compared with the original xenobiotic.\textsuperscript{175} This is exemplified in the case of methanol bioactivation. When ingested, it is converted to formaldehyde, which is extremely harmful.
2.1.2.2 Phase II Metabolism

The second step in the metabolic process involves the conjugation of the products created by phase I metabolism with larger molecules rendering them harmless and more water-soluble. These conjugation reactions result in the clearance and elimination of potentially biologically harmful products of phase I metabolism of xenobiotics via biliary or urinary excretion. Typical conjugations involve glutathione via glutathione-S-transferase (GST), glucuronidation by glucuronic acid, as well as sulfate conjugation via sulfotransferases.

If these conjugations do not occur, or are rate-limited, the reactive byproducts, from phase I, are able to interact with other macromolecules within the body. Several key events could increase this possibility. These are:

1. Phase I enzyme levels, and corresponding product levels, exceed the normal rate of phase II enzyme conjugation.
2. There is a rate limiting effect on the phase II enzymes by:
   a. a decrease in phase II enzyme levels
   b. a decrease in phase II enzyme function
   c. the supply of conjugatable moieties (i.e.: glutathione)
      is decreased or has reached a maximum threshold

Another important consideration for toxicology is the “dose” of a xenobiotic as well as the length of exposure to an individual, in that they both, either separately on in combination, can affect these events. The accumulation of phase I byproducts, by any of these means, is linked to liver carcinogenesis by the
formation of hepatic foci, as well as to toxicity considering many of these byproducts have been shown to cause mutation as well as induce a state known as oxidative stress.\textsuperscript{342}

\subsection{Hepatic Foci}

Liver tumors can have many causes. Excluding Europe and the United States where alcohol consumption is the primary cause\textsuperscript{268}, tumors of the liver are most commonly a result of underlying viral infection, as is the case with hepatitis viruses B and C, or there is a genetic basis, as observed in some familial inheritance patterns, both that precede hepatocellular carcinoma (HCC) development. In the absence of any virus or genetic basis, chemical compounds can, also induce liver tumors.

Many of the drugs that cause liver cancers, produce intermediate changes in liver morphology. These occur after treatment, but prior to any tumor development and are known as hepatic foci. Foci have been described\textsuperscript{241} and implicated in the development of liver neoplasms, not only of HCCs, but also of benign hepatomas. Hepatic foci appear in response to low dose administration of carcinogens. At these doses, no signs of toxicity are detectable either histologically (manifested as inflammatory infiltrates) or by regenerative response (as indicated by the lack of nuclear replication).\textsuperscript{66} These foci have characteristic features associated with them. They are histologically evident as localized regions of hepatocytes with enlarged nuclei. The enlarged nuclear size
(roughly double in volume) has been attributed to increases in aneuploidy, differing from the normal diploid or tetraploid status of hepatocytes\(^6,\)\(^{32}\)

In early studies in this laboratory, rats were treated with low dose carcinogens. Within the hepatic foci visible in the livers following treatment, specific RNAs, normally contained within the nucleus, were relocated to the cytoplasm\(^6,\)\(^{67}\) This altered compartmentalization was proposed to possibly be a result of leakage of the nuclear membrane or transport of these sequences from the nuclei, a characteristic feature in response to carcinogen treatment. Upon further investigation, these RNAs contained repetitive sequences, confirming the results of Patel et al.\(^{238}\) who found these nuclear RNAs tightly bound to chromatin in untreated rats. These sequences were reverse-transcribed from the hepatocyte cytoplasm of rats treated with thioacetamide and had significant similarity to rodent B2 transcripts, the rodent form of Alu-like repeats. More detail regarding this observation will be discussed later. In the same study, the authors identified regions of histologically-normal cells, surrounding foci, that exhibited positive nuclear staining for 4-hydroxy-2-nonenal (4-HNE), a product of oxidative damage.

### 2.2 Microsatellites

Regions of DNA that consist of short tandem units of repeated DNA bases are known as microsatellites. Microsatellites can appear as units of 1-6 nucleotides with repeat units of 1 and 2 nucleotides (known as mono- and di-nucleotide repeats, respectively) being the most predominant of all
microsatellites in the mammalian genome. Microsatellites make up 3% of the human genome and are randomly distributed. Upon stepwise analysis, both mono and di-nucleotide repeats show consistently higher densities in regions of DNA that do not code for protein. Of these, both types of repeats exhibit higher densities in introns than in intergenic DNA.\textsuperscript{288} In both mice and rats, similar observations are seen, although it has been noted that rodent stretches of microsatellites were longer.\textsuperscript{26}

Typical mutations of microsatellite repeats are classified as indels and are characterized by insertions or deletions of nucleotides composing entire repeat units. Alterations that occur in the reading frame of protein coding DNA can have dramatic consequences if the codons are altered. These types of mutations are more specifically called frameshifts, a subclass of indels. Particularly in the case of microsatellites that consist of repeat units other than 3 or 6 bases, which would maintain the integrity of the downstream coding sequence, the gain or loss of one repeat unit would have significant impact on the protein production. This impact could be observed ranging from aberrant folding, to premature truncation, or even to a complete loss of coded protein. It has also been noted that frameshift occurrences increase proportionally with the number of repeats in the microsatellite.\textsuperscript{47}

Increases in the rate of occurrence of these frameshifts, known as microsatellite instability (MIN), are frequently observed in many types of tumors. The incidence of MIN is extremely high in colorectal cancers associated with Lynch syndrome, a genetic variation causing deficient DNA repair that
predisposes individuals to developing the hereditary form of non-polyposis colorectal cancer (HNPCC).\textsuperscript{141, 262, 275}

The normal somatic mutation rate for any given nucleotide in a non-repetitive sequence is less than 1x10\textsuperscript{−9}.\textsuperscript{192, 194, 195, 296} In microsatellites, because of their susceptibility to indel mutation, thought to be due to “polymerase-slippage” during replication and repair processes,\textsuperscript{190, 196} the rate is 1x10\textsuperscript{−6}.\textsuperscript{264} MSI increases this rate further, as the average is 1x10\textsuperscript{−2}.\textsuperscript{264} In addition, as the length of repeated sequence increases, the rate of mutation increases.\textsuperscript{326}

### 2.2.1 Genomic Instability

Microsatellite instability is only one type of genomic instability. Genomic instability, as a whole, is a fundamental characteristic of cancer.\textsuperscript{59} Genomic instability can be classified as either mutagenic, affecting ribonucleotide sequence, or clastogenic. The latter type, also known as chromosomal instability (CIN), refers to the increased rate of change in the superstructures of the genome, i.e.: chromosomal alterations. These chromosomal alterations are evident as sister-chromatid exchanges, variable homologous recombination, etc. CIN results in aneuploidy, whereas MIN is a characteristic of cells that may maintain their diploid status. While these changes are evident in cancers, most non-heredity-based instabilities do not occur until tumors are more advanced according to pathology-based stage/grade.\textsuperscript{225} The process by which a specific tumor acquires such an abundance of alterations can be attributed to the presence of a mutator phenotype\textsuperscript{193}. 

Considering the low spontaneous mutation rate per cell division, the acquisition of multiple mutations in a non-dividing tissue, would not be possible in a short period of time. To explain this, two ideas have been proposed that describe how all these mutational events occur in a brief period of time. The first is described as a mutator phenotype. Briefly, there is a mutation that occurs in a specific gene, which conveys a growth advantage, or errors of DNA replication and/or repair, to a particular cell, thus providing either the loss of growth or repair control required for further mutations to occur and leading to an increased mutation rate. Rapid tissue proliferation, for instance in the case of developing tissues, can also convey something akin to a mutator phenotype, in as much as it also triggers an increased growth rate/DNA synthesis required for additional mutational events. Alternatively, there exists an explanation that does not involve increased mutation rates, but instead is based on selective advantage. This idea proposes that mutations, occurring as a result of a normal mutation rate, cause alterations in cells that confer specific advantages toward selection and clonal expansion. The most notable of these is evident as selection against apoptosis observed in cancers.

When considering chemical carcinogenesis, if a particular chemical affects one of the genes supporting growth regulation or repair, the model of a mutator phenotype is supported. There remains another possibility for tumor development, which can be explained more clearly by the initiation/promotion model. In this case, two steps (and usually two chemicals) are involved: initiation, by a chemical referred to as an “initiator”; and promotion, with another chemical known as a “promoter”. During initiation, a single
event/mutation occurs that remains latent until a second chemical, most often as a chronic exposure, allows for growth advantages for “initiated” cells. The underlying cause in both of these theories is that a permanent heritable change occurs. This change may be genetic, as in a direct mutational event in protein-coding DNA sequence, or epigenetic. Given this conclusion, most carcinogenicity assays often rely on the detection of DNA mutation.

2.3 Detection Assays-Carcinogenicity Estimation

Both in vitro and in vivo assays are currently in use by the National Cancer Institute (NCI) and National Toxicology Program (NTP) to study the carcinogenic potential of chemicals. The long-term bioassay, performed for 24 months with rats, and 18 months with mice, determines the propensity for tumor development. While these studies are informative, they have a number of drawbacks, they are not only costly, but time-consuming, low-throughput, and most notably, do not allow for determination of the underlying mechanisms, which is particularly important because differences may exist across species. In addition, these studies often use the maximum tolerated dose (MTD) of a chemical. Using the MTD, a dose that is just under a lethal dose, can cause tissue injury and inflammation, and may not reflect the usual effects of the compound at lower doses characteristically ingested. In doing so, the investigators may actually be inducing a mutator phenotype during proliferative repair following injury, thereby, a tumor formed may not be attributable to the chemical treatment alone.
Other assays for carcinogenicity over a shorter period of time include, the Ames *Salmonella typhimurium* assay\(^{10}\), the mouse lymphoma L5178Y \(tk^+/-\) assay, sister-chromatid exchange (SCE) detection in Chinese hamster ovary (CHO) cells, and the *in vivo* bone marrow micronucleus assay\(^{265}\). Most of these assays remove a key contribution to the carcinogenic potential of a chemical, namely individual tissue metabolism. To get around this difficulty, phage-based transgenic (Tg) mice were developed. These are marketed as the Muta™ Mouse\(^{117}\) and the Big Blue® Mouse\(^{172}\). Each of these mice contains either a lacZ or LacI gene reporter, respectively, inside a lambda (\(\lambda\)) phage within the tissues of the mice. This reporter can be altered by a chemical insult that may need bioactivation by the tissue. Following treatment, the phages can be isolated, incorporated into bacteria and mutants selected from blue plaques for further analysis.

In all of these *in vitro*, *in vivo*, and Tg mouse based assays, the endpoint is observed as a mutation, or DNA-based alteration, in the case of micronucleus formation. For this reason, chemicals producing positive results in these assays are referred to as “genotoxic”. There is however, another group of chemicals that yield negative results in these tests, but still cause tumor development in the long-term bioassays. Most often, these “non-genotoxic” chemicals are classified as promoters, which, by definition, would require a previous initiating mutation(s) to have taken place. This assumption is not always correct, considering certain non-genotoxins are carcinogenic by themselves. That is, they cause tumor development with repeated treatment in non-initiated animals.\(^{15,279}\)

Along a similar line, some chemical agents can yield a positive/mutagenic response in these assays, but are not carcinogenic. Explanations of this disparity
are supported by, efficient DNA repair, cell death, and sub-threshold levels. At these doses, the repair mechanisms and/or metabolism still remain functional and are able to eliminate the insulting chemical or its damaging effects.

2.3.1 Mouse Model Incorporating Plasmid-Based Microsatellite Sequence

In the late 1990’s another transgenic mouse line was created that contained G-repeat microsatellite, integrated directly into the genome without a λ-phage. By doing this, the mice could be studied for the spontaneous mutation frequency in the microsatellite sequence. This mouse line integrated approximately 40 copies of a transgene containing the placental form of alkaline phosphatase (PLAP) driven by the human β-actin enhancer/promoter. The PLAP gene was altered by the insertion of a tandem repeat of 7 guanine nucleotides (Gs) into a stretch of 4 Gs that were already within the coding sequence, thus rendering the gene inactive, since the codons would be shifted to an out-of-frame configuration in the +1 direction. In order to detect functional PLAP, the repeat sequence would need to lose one repeat nucleotide, in this case one single guanine base pair effectively shifting the PLAP back to an in-frame configuration. Additionally, this insertion of a microsatellite introduced a premature translational stop at codon 29. These mice were shown to have a high incidence of spontaneous mutation that was attributed to the natural instability of the G-repeat and the reversion rate for the microsatellite reflected a similar sequence (G)₇ in cells. In a fibroblast cell line, NIH3T3, this rate was estimated to be 2x10⁻³. In an additional study, the lab showed these mice exhibited very low
transcript and protein expression in the livers. In the context of studying chemically-induced hepatocarcinogenesis, this mouse model is simply not practical.

2.4 Oxidative Stress

The formation of radical oxygen species (ROS) is a byproduct of metabolism. At several points during the metabolism of a compound, several oxygen species can be generated and released (Figure 2.2 shows an example of such metabolism). Of particular note, is the decay of the reduced complex (d) formed after binding with molecular oxygen (O\(_2\)), thereby releasing a superoxide anion (O\(_2^•\)) and the formation of hydrogen peroxide (H\(_2\)O\(_2\)) caused from the protonation of complex (f). The superoxide radical is extremely unstable and undergoes a rapid conversion also to form H\(_2\)O\(_2\). Hydrogen peroxide is a small molecule that can rapidly move around intracellularly, due to its neutral charge. Both the superoxide anion and the hydroxyl radical (\(^•\)OH), a product of peroxide degradation, are incapable of this intracellular migration due to their electrophilic states. These charged compounds readily lend themselves to reaction, thus making them unstable.
Typical hydrogen peroxide levels within a cell are $10^{-7}$-$10^{-9}$ M. These levels are not altogether surprising considering that ROS are also formed by mitochondria during normal cellular respiration. In addition, H$_2$O$_2$ has been also implicated as an intracellular signaling molecule via its potential interactions with cysteine residues of critical enzymes such as protein phosphatases.

Cells utilize several key enzymes to control internal ROS levels; namely these are superoxide dismutase (SOD), catalase, and glutathione peroxidase. Glutathione itself can also directly bind ROS as a “scavenger” thus preventing an opportunity for damage. All of these serve as types of antioxidants. Additional antioxidants like vitamin E and selenium, can also contribute to the control of ROS levels. From this, it seems evident that the damage potential for a cell is dependent on its antioxidant capacity.

Figure 2-2: ROS release from CYP450 catalysis. (Source: Zangar, R.C. et al. Toxicology and Applied Pharmacology (2004).)
If the antioxidant levels drop or have hit a maximum threshold, radical accumulation induces a state of “oxidative stress” inside the cell where it is possible to incur ROS-induced damage. This damage can include lipid peroxidation, protein disruption, cellular membrane disruption, or direct DNA adduct formation. It should also be mentioned that oxidative stress is induced by the influx and response of phagocytic cells, i.e. macrophages, during inflammation or infection; however, in the case of chemically-induced carcinogenesis, this point is not necessarily applicable as it usually occurs as a later event in the carcinogenic process.

Cells maintain a certain level of repair mechanisms. For instance, protein conformational changes resulting from ROS-binding may trigger proteolytic degradation; however, it is through the binding of DNA that ROS have been primarily linked to cancer. Electrophilic compounds may intercalate directly into the DNA strand with backbone structures. However, most often, the result is direct oxidation of nucleotides. These lesions have been estimated to occur at a rate of 20,000/cell/day.

DNA surveillance machinery ensures the fidelity of its sequence. During replication, DNA polymerases, delta (δ) and epsilon (ε) are responsible for the correction of mismatched nucleotides with their 3’-5’ exonuclease activity and DNA polymerase beta (β) is responsible for repair of apurinic/apyrimidinic (AP) sites. Even though the error rates of these polymerases, is low (δ and ε, 1:10^7 and β, 1:5000), problems of replication fidelity may arise during high rates of DNA synthesis and proliferation.
Recent evidence also links these polymerases to the processes of base excision repair (BER). These processes are involved in non-replicating cells and act to efficiently remove 8-hydroxy-deoxyguanine (8OdG), an abundant adduct formed by hydroxylation of the nucleotide guanine. At other times, ROS damage occurs as strand breaks and in the case of a repeated sequence within a transcribed coding sequence, a frameshift may result. Additionally, cells that have defects in one or more components of their mismatch repair (MMR) systems (for example, colonic epithelial cells in patients with Lynch syndrome) exhibit a greater propensity for these frameshift mutations as a result of 8OdG. This may occur by 8OdG, or the misincorporation of an adenine (A) opposite the 8OdG, not getting recognized by the mismatch repair, MutS or MutY proteins. MutY, although considered a BER protein, has been shown to interact with MutS in MMR. Considering deficiencies in the processes of MMR, the connection between ROS lesions and frameshift mutations is further supported by the increase in microsatellite mutations, often observed as frameshifts, in MMR-deficient mice.

Additional links have been drawn between microsatellites and oxidative stress. In a process termed “molecular misreading”, identified in advanced neurodegenerative diseases such as Parkinson’s and Alzheimer’s, defective RNA surveillance has been implicated in aberrant protein expression. Molecular misreading would occur as a result of microsatellite slipping during transcription since the DNA sequence was found to be unchanged. In both the ubiquitin B (UbB) and beta amyloid precursor protein (βAPP) a dinucleotide frameshift was detected in the RNA transcript from the sequence GAGAG by the loss of a GA.
This deletion of two nucleotides caused a premature stop codon, which resulted in truncated proteins.\textsuperscript{310} In the case of UbB, the terminal end that is essential for poly-ubiquitination is not present. The protein, not able to be degraded, then accumulates in the brain tissue.\textsuperscript{311} This accumulation is one of the key features detected in neurodegenerative disease, to which increasing oxidative damage has also been linked. It was speculated that the molecular misreading during translation was a result of declining or defective RNA surveillance systems\textsuperscript{214, 286} that would normally act to detect premature stop codons.\textsuperscript{311} It could be assumed that other age-related diseases are results of a similar phenomenon. Contrary to this opinion, molecular misreading was originally observed in the vasopressin gene transcript from diabetic rats. In this case, the defective vasopressin gene function was actually restored.\textsuperscript{96} When considering this model in the context of cancer, the translational slipping events would result in protein differences; however, if the translational machinery was responsible for this slippage, most if not all of the mRNAs would be affected and within the defective gene there would need to be a repeated sequence that would be prone to slipping. These events in general have not been shown to be the case. The model does however bring to light the idea of RNA regulatory mechanisms, which may be affected in various disorders, such as cancer. This idea will be discussed in more detail later.

\textbf{2.5 EPigenetic Regulation}

Carcinogenesis due to oxidative damage is not solely based on DNA adduct formation.\textsuperscript{170, 171, 279, 303} In the case of non-genotoxic carcinogens, there must be
another explanation for the development of tumors. Observations from studies comparing tumor and normal tissues show distinct characteristics. Differential gene expression is perhaps one of the clearest examples. In 1993, Ray Tennant proposed that non-genotoxic carcinogenesis may be a result of transcriptional regulation. This almost certainly does not involve general defects in transcriptional machinery with uniform affects on all genes. Cancer development for instance, is not always accounted for by gene transcription alone. In fact, mRNA expression in tumor tissue of several cancers remains the same as that of normal tissue, but other modifications occur (i.e.: regulation of protein transcription caused by the anti-sense binding of microRNAs (miRNAs) to mRNA transcripts, a feature that will be discussed in more detail in the upcoming sections.)

Again, a general dysregulation of transcription is not observed in tumors, as certain genes show increased transcription, i.e.: oncogenes, while others do not, i.e. tumor suppressors. Support for this idea can be found in the examination of the transcription factor nuclear factor kappa B (NFκB). While NFκB is upregulated in response to oxidative stress, not all of the genes to whose promoters it binds have increased protein levels.

Epigenetic alterations are defined as changes in gene or protein expression, not involving DNA changes. One of the most studied forms of epigenetic modifications is that of modified gene expression based upon differential methylation patterns.
2.5.1 Methylation

One of the most studied forms of epigenetic modification is methylation. Methylation events occur by the incorporation of a methyl group on a cysteine residue that precedes (linked 5’-to-3’, by its phosphate backbone), a guanine (referred to as a CpG). Within the genome, CpGs exist as CpG islands where they are grouped, occurring as >1 CpG in every 8 nucleotides\textsuperscript{73}, or are also commonly found ungrouped, scattered throughout inter- and intragenic regions of DNA\textsuperscript{79}. \textsuperscript{197} CpG islands are usually located either in or near coding gene promoters and are typically unmethylated allowing for gene expression as methylation has been shown to silence gene transcription\textsuperscript{284}. One possibility for this silencing is the change in DNA structure from an open euchromatic state to a more condensed heterochromatic state by the recruitment of protein complexes, including histone deacetylases, by methyl binding domain proteins.\textsuperscript{197} The methylation of these cystines effectively condenses the chromatin structure, blocking the access of transcription factors to the DNA recognition sites, thereby inhibiting transcription.

In the case of oxidative damage, an adducted G (8OdG) is able to block methyltransferases from methylating, and silencing, a promoter.\textsuperscript{305,324} This decreased methylation could drive transcription, but this lesion has also been shown to interfere with transcription factor binding which, in turn would inhibit transcription.\textsuperscript{97} As DNA repair mechanisms in most cases, efficiently remove these bases, this could not be the sole basis of cancer development as it may transiently affect other elements involved in the process of tumor development.
Methyltransferases might also be implicated in cancer development, however alterations in their function would be detectable as a hyper- or hypomethylation of all CpGs. While cancers generally exhibit a pattern of global hypomethylation, there are regions of DNA that become hypermethylated, evident by the suppression of the tumor suppressor genes, and more so methylation profiling.\textsuperscript{72,95} The mechanisms underlying the establishment of these tumor-specific methylation patterns are still unclear, however, the differential methylation patterns observed within cells indicate that the disruption of methyltransferases, or their activity, is not the sole cause of cancer development.

2.5.2 Genomic Imprinting

Methylation is also involved in other forms of epigenetic modification. Genomic imprinting refers to methylation differences between alleles. Genomic imprinting defines the way genes are expressed. Each parent confers on their offspring not only the Mendelian idea of one allele, but also specific methylation patterns that dictate how the genes are to be expressed. Gains or losses of imprinting cause both alleles to function similarly. Thus, when methylation patterns change, resulting in gains or losses of imprinting, cancers may develop. This link between imprinting and cancer can been seen in the case of Wilms kidney tumors.\textsuperscript{102} These tumors have a dual allele expression of insulin-like growth factor 2 (IGF2). In normal tissue, only one allele expresses IGF2, whereas the other allele expresses H19, a transcribed gene, which does not encode a protein. When both alleles express the same protein, in this case IGF2, an
overexpression of protein occurs and resulting in tumor formation. Regions in between these two genes that exhibit differential methylation are thought to be responsible for this expression.

Once again, methylation patterns seem to have a role in cancers; however, in the case of liver carcinogenesis, these imprinting differences are not universal. If the patterns of methylation, or the processes affecting them, were always responsible, there would be a universal trend observed which is not the case.

2.5.3 Effects of Methylation on RNA

Methylation also affects RNA. Within the 3’ untranslated region (3’ UTR) of messenger RNA (mRNA) there are regions of sequence that have high concentrations of CpG. Again, these are known as CpG islands and are typically associated with Alu repeats. In fact, about 1/3 of all CpGs are located within these sequences. The potential for the sequences to be highly methylated in combination with their presence in 3’ UTRs may indicate a role in gene regulation.

These regions also are located near, or contain, microsatellites although these are typically found as repeats of adenines as would be found in poly-A tails of RNA. Other genes that contain microsatellites exhibit de novo methylation on expanses of CGG trinucleotide repeats (tri-nt) indicate a link between the methylation of RNA and human disease. This hypermethylation, most evident in the expanded tri-nt repeats in the 3’ UTR of the folate-sensitive fragile sites,
results in mental retardation syndromes. The most notable example is Fragile-X, resulting from the methylation-induced silencing of the FMR1 gene. It has been suggested that the increased length of the CGG tri-nt repeats induces hairpin folding of the DNA which attracts the dinucleotide methyltransferase protein, DNMT1, and accounts for the de novo methylation of these sequences.

Alu repeats have also been linked to human disease as in tumor development. Both adenomas and carcinomas from colon cancer show deletions in the poly-A tail, although, this most likely can be attributed to microsatellite instability. Once only thought of as transposable elements during DNA homologous recombination, these repeats are now recognized as transcribed non-coding RNA (ncRNA).

There is another modification that should be addressed at this point in our discussion involving epigenetic mechanisms affecting mRNA stability. These modifiers are proteins that act as post-transcriptional regulators through their interactions with the 3’ UTR sequence of mRNA. Proteins such as HuR and AUF1, known as ARE binding proteins since they interact with the AU rich element (ARE) located in the 3’ UTR, are involved with the nuclear-cytoplasmic shuttling as well as in the stability of mRNA. HuR protein is of particular note in that the binding of HuR to an ARE causes the dissociation of a translational repressor, miRNA-122, another non-coding RNA, from its target mRNA during a stress response.
2.5.4 Histone Modification

Another epigenetic mechanism involving methylation patterns is in heterochromatin rearrangement via histone modifications. Briefly, methyltransferases act on histone 3 on lysine 9 (H3K9), the methylated histone then acts to recruit other proteins like HP1, which in turn, condense chromatin. The condensed chromatin structure is then able to affect DNA methylation patterns according to which DNA segments are exposed. In studies with fission yeast, Schizosaccharomyces pombe (S. pombe), Volpe et. al. discovered yeast that had defective or missing RNA interference (RNAi) processing machinery exhibited altered methylation of H3K9 and were unable to recruit Swi6 protein, the yeast homologue to HP1. The processing machinery described here is responsible for the generation of short RNA molecules (~22 nt in length), often from longer pieces of transcribed non-coding RNA.

2.6 Non-Coding Genes

The recurring theme of non-coding genes keeps arising as commonality in altered gene regulation, with an association to cancer. The term “gene” presented here is in agreement with a new definition. No longer is a gene considered just a sequence of DNA that follows the central dogma of DNA to protein through mRNA, but it has now been expanded to include heritable sequences that are involved in regulatory functions of the genome. It should also be pointed out these regions may not necessarily be transcribed.
Non-coding RNAs, also known as riboregulators, are defined as genes that get transcribed (often including a poly-A tail), but do not contain an open reading frame (ORF); therefore, no protein is produced, but the transcripts can still affect some aspect of genome regulation.\textsuperscript{94} For more detail regarding non-coding RNAs and their function, see recent reviews.\textsuperscript{9, 60, 61, 165}

I have already mentioned several ways in which non-coding RNA can interact with the genome, for example binding 3’ UTR sequences of mRNA or regulating histone methylation patterns, both possibly including protein recruitment. Supporting this, a recent discovery of the non-coding RNA, NRON, showed that it interacted with several proteins including some transcription factors.\textsuperscript{327} It should not then come as a surprise that riboregulation might affect other aspects of genome control, the details of which have only begun to be examined.

\subsection*{2.6.1 Microsatellites in Non-Coding Genes}

As previously mentioned, microsatellites, and other repeated sequences, are found particularly in regions of DNA that do not code for a protein product. In these regions, it is estimated that they are 5-10 times more prevalent than random sequences.\textsuperscript{158} Due to this high frequency, repetitive sequences are believed to serve a regulatory function within the genome. In 1967, this idea was recognized by Shearer and McCarthy, in their observation that most repetitive RNA is restricted to the nucleus.\textsuperscript{272} While no microsatellite sequences have yet been identified directly in the sequences of known riboregulators, their
abundance as well as location, present an intriguing question as to a possible functional role, perhaps as part of riboregulator processing.

### 2.6.2 RNA Interference (RNAi)

Recently, studies examining transcriptional products have observed a pattern of transcription from both strands of DNA not just one as previously thought\(^{136,337}\) again, suggesting an important role for riboregulators. If transcription is driven from both strands of DNA, the complementary anti-sense strand could serve to silence the transcripts from the sense strand. This mechanism of riboregulation is termed RNA interference (RNAi) due to the effects on mRNA translation caused upon binding to the riboregulator. The binding acts to suppress translation of the particular mRNA by causing either mRNA degradation or translational “silencing” through the association of mRNA with a complex of proteins known as a RNA-interfering silencing complex (RISC).

The most commonly studied riboregulators in RNAi are known as the microRNAs (miRNAs) and the short-interfering RNAs (siRNAs), designated by their function and process of maturation. Typically, the miRNAs interfere with translation by sequestering mRNA in association with recruited proteins into RISCs, whereas the siRNAs are involved in the degradation of the mRNAs.

The maturation of miRNA occurs as follows. While transcription from the sense strand of DNA (providing mRNA) is a product of RNA polymerase II (pol II), transcribed non-coding genes result from both the sense and anti-sense
strands. The transcription of non-coding genes mostly results in a clusters of miRNAs generated by pol II in a single transcript\textsuperscript{166} that can be several kilobases in length.\textsuperscript{70} These long transcripts often contain a 7 methylguanosine cap on the 5' end and a poly(A) tail on the 3' end, much like mRNAs (generated by pol II), on transcripts of protein coding genes.\textsuperscript{185} In addition, transcripts of non-coding genes can also be generated by anti-sense strand transcription by RNA polymerase III (pol III) that recognizes promoter elements repeat sequences (i.e.: Alu repeats).\textsuperscript{38} The nuclear ribonuclease (RNase) III, Drosha in association with DGCR8 (a.k.a. Pasha\textsuperscript{338}), cleaves these long transcripts with secondary hairpin structures (pri-miRNA) into smaller double stranded fragments of approximately 70 nt in length, known as pre-miRNA.\textsuperscript{184} Following this, the pre-miRNAs are exported out of the nucleus by the protein Exportin 5\textsuperscript{35} where they are further processed into single stranded miRNA, of approximately 22 nt, in the cytoplasm by the complex of RNase Dicer\textsuperscript{28} and TRBP.\textsuperscript{161} The mature, processed miRNAs are then free to interact with mRNAs. During siRNA processing, the main difference in this process is that the original transcript is actually two from dual strands that anneal and thus become the template for Dicer/TRBP, thereby bypassing the need for Drosha/DGCR8 processing. Another difference between these RNAi molecules is related to their function. siRNA degradation of mRNA results from perfect complement matches, whereas for miRNA translational silencing, this perfect match is not essential.

Another key idea to discuss here is that of RNA editing. During RNA editing, nucleotides (C and A) are modified. The most common modification in mammalian cells is that of adenosine deaminated to an inosine residue carried
out by the family of adenosine deaminase acting on RNA (ADAR) proteins.\textsuperscript{20} It is estimated that at 6\% of all pri-miRNA are edited,\textsuperscript{33} although this number may be in fact be greater as more pri-miRNAs are studied. It has also been shown, that up to 50\% of adenosines can be modified in some miRNAs.\textsuperscript{243} ADARs act on double stranded RNA\textsuperscript{164} as is commonly found in the non-coding regions of RNA (3\’ and 5\’ UTRs as well as introns) at regions of inverted repeats. In this way, they compete with the miRNA processing pathway Drosha/DGCR8(Pasha) complex\textsuperscript{336} and the Dicer/TRBP complex\textsuperscript{161} for dsRNA templates.

The importance of this editing is multifold. Most importantly, inosine residues are “recognized” by the cellular machinery as guanines and base-pair with uracil. Edited RNA can adopt altered conformation of RNA structure thereby preventing recognition by cellular machinery as described above, can block RNA transfer via Exportin 5,\textsuperscript{245} or even alter the production of another type of riboregulator, the repeat-associated small interfering RNA (rasiRNA). The altered production occurs by the conformational change of the hairpin loops resulting from RNA fold-backs in repeat sequences.\textsuperscript{224} Another idea that has been proposed as a result of RNA sequence editing that in turn alters the binding patterns and recognition of miRNAs. This can result in a change in target mRNA(s) and would allow for a greater number of targets to be modulated by a single miRNA\textsuperscript{162} or even affect the selection of the “effective” strand of miRNA from the duplex cleaved by Dicer.\textsuperscript{33}
2.6.3 RNAi Expression in Hepatocarcinogenesis

While it would be understandable to imagine that defects in RNAi processing or transporter components are responsible for carcinogenesis, there are several key pieces of evidence that refute this idea. With regard to processing note, that following low dose carcinogen treatment, there is an altered compartmentalization of RNA. This would seem to indicate that key elements such as the RNA polymerases, (i.e.: Pol II and Pol III), responsible for transcribing RNA, as well as transport proteins (i.e.: Exportin 5), are functional. Additionally, studies of RNA, have shown that within tumor tissue RNAs are hypoedited on repetitive elements while a simultaneous gene specific pattern of editing occurs. This indicates that the proteins responsible for the editing are functional, however their regulation or targets somehow have been altered.

Editing in relation to RNAi is a fairly new field of study, so that further experimentation is required to elucidate more of these mechanisms.

The second key piece of evidence comes from studies on the expression profiles of miRNA. Using microarray technology several groups have indicated that there is a dysregulation of miRNA expression in hepatocellular carcinomas as compared to normal tissue, as well as in liver tissue following either acetaminophen or tamoxifen treatment. miRNAs are not specific to one unique target and instead target multiple mRNAs through perfect complementarity or imperfect matches, a nod to the importance of editing in miRNA recognizing a particular target. These miRNAs that can act in either an oncogenic or tumor suppressing fashion, depending on the specific target being affected, and they show patterns of dysregulation, with certain subsets of
miRNAs implicated in tumorigenesis. Often, these miRNAs are involved in processes of cell differentiation or growth. While some are commonly upregulated such as the mir-17 cluster (including miR-17.5, miR-20 and miR92), which targets tumor suppressor genes such as RB1 and E2F1, others show reduced levels of expression in HCC. For example, miR-16 and let-7, act as tumor suppressors, by modulating Bcl-2 and Ras protein levels, respectively.\textsuperscript{140}

Differential patterns of miRNA expression are not specific for liver cancer, as well-defined profiles have been reported for miRNAs in cancers of nearly all tissues, for example, from chronic lymphocytic leukemias\textsuperscript{50}, also, miRNA expression profiles are not universal for all cases of HCC, suggesting different origins. Several key miRNAs that show prognostic correlation between expression levels and HCC diagnosis\textsuperscript{186}, are often also observed in cases of prostate\textsuperscript{232}, breast\textsuperscript{142}, and thyroid cancers\textsuperscript{315}. Further evidence of this variability between cases of HCC has been noted\textsuperscript{14,186} and is speculated to be the result of background genetics, disease etiology, or simply a result of differences in experimental design, or a combination thereof.

It is also important to note that the human mir-17 cluster located on chromosome 13q31, which shows increased expression in cancers or in liver tissue following tamoxifen treatment\textsuperscript{242} and is thought to aid, if not drive cellular proliferation, is the same region commonly deleted (via loss of heterozygosity) in HCC\textsuperscript{189}.

Taken together, this evidence further indicates that this cluster along with other unique miRNA expression profiles participate in hepatocarcinogenesis. If the assumption is made that the processing components are damaged or
deficient, the observed downstream effects would lend themselves to a conclusion of global increases or decreases in miRNA levels. Considering the different patterns of expression observed, the disruption of RNAi processing cannot be inferred. Furthermore, due to the variation in patterns between cases of HCC, global miRNA maintenance/control also cannot be implicated in hepatocarcinogenesis. It is entirely possible that selective transcriptional regulation may account for the variation in expression, however, there is thus far no evidence backing this idea or indicating how certain miRNA genes are targeted for up or down regulation.

2.6.4 Alternate Riboregulation

While RNAi is perhaps the most well studied form of riboregulation, recent studies have begun to examine a new form of riboregulation having to do with chromatin structure and centromere function. As previously mentioned, initial studies identified a link between non-coding RNA and DNA methylation in S. pombe, which precedes protein recruitment required for chromatin condensation and silencing.\(^{150, 251}\) In fact, studies in several other model systems have also shown a similar link between small RNA transcripts and protein binding/chromatin structure through heterochromatin formation.\(^{19, 205}\) Just recently a role for epigenetics and structure directing the location of centromeres has been proposed, as the DNA sequence is not evolutionarily conserved.\(^{300}\) Despite the identification of non-coding transcripts participating in chromatin structure being a fairly recent observation, this idea of RNAs involvement has
been around for some time. In 1990, Vogt et al. proposed that his “chromatin folding code” may be based on not only DNA sequence, but also on the transcripts of satellite DNA, observed in both human and mouse systems.\textsuperscript{316}

Further examination into the fission yeast \textit{(S. pombe)} model system (which otherwise seems to lack miRNA and its regulation) detected small RNAs in the size range of \textasciitilde{}22 nt, that were complementary to pericentric DNA repeats.\textsuperscript{317} These transcripts have been demonstrated to be responsible for the recruitment of proteins into RNA-interfering transcriptional silencing (RITS) complexes\textsuperscript{312} or to the RNA-directed RNA polymerase complexes (RDRC).\textsuperscript{267} These complexes remain nuclear and appear to target centromeric DNA containing repeats.\textsuperscript{150} For clarity, these short (\textasciitilde{}22 nt) RNA transcripts that fulfill this specific role, will be denoted as centromere-associated silencing RNA (casRNA).

Given the fact that the size of these transcripts is within the range of miRNAs and siRNAs, it was proposed that Dicer was also involved in the processing of casRNAs. Upon elimination of Dicer, in mouse ES cells as well as \textit{S. pombe}, centromeric heterochromatin is disrupted causing an increase in transcription of centromeric repeats, indicating that Dicer is indeed required for the processing of casRNA.\textsuperscript{157, 216} Three other observations of note for this discussion resulted from these studies. The first is that the methylation of centromeric DNA as well as overall heterochromatin state, was unchanged by the loss of Dicer in a Cre/Lox system, thus suggesting a role for casRNAs in the establishment, but not maintenance, of centromeric heterochromatin. The second observation is that transcripts corresponding to both centromeric and pericentromeric satellite repeats increased following stress conditions and the
third is the observation that double-stranded RNA in Dicer-proficient cells exists as 2 distinct sizes and each of these binds on Northern blot with a probe generated from centromeric repeat DNA. It has thus become apparent through these studies that there seems to be another transcript generated from the centromeric region, in addition to the casRNA, and it also plays a role in centromere structure and/or function.

Centromeric regions are divided into two classes of repeated DNA sequences known as the minor satellite repeats, located in the primary constriction of condensed chromosomes and major satellite repeats that exist in the surrounding DNA classified as “pericentric”. Since it has already been established that casRNAs originate from pericentric repeats, we turn our attention in this discussion to those RNA transcripts from the “centromeric” minor satellite repeats.

Minor satellite repeats consist of long arrays of tandem repeats that are approximately 120 nt in length. Interestingly, transcription across these regions produces long transcripts (2 and 4 kb in length) which following time in culture or under stress conditions, seem to be cleaved into small RNAs of roughly the same length as the DNA repeat unit (120 nt), as 120 nt long RNAs appear with no observed increase in transcription from the minor repeats. Transcription across these repeats has been shown to occur in a bi-directional manner giving credibility to the idea that these regions form dsRNAs that can be templates for Dicer.

These transcripts that localize to the minor satellite repeats have also been observed to be differentially expressed during the cell cycle. During S phase,
when both centromeric DNA methylation and histone protein binding are low, RNA Pol II is localized to the centromeric regions and transcription levels of these regions are elevated. Following this transcription, both methylation and association with histone proteins (or histone variants specific to centromere architecture, such as CENP-A) are gradually restored\textsuperscript{57}, lending further credence to the idea that these transcripts are involved in establishing chromatin structure.

The idea that these transcripts are involved in centromere function is supported by studies, which involved overexpression of these 120 nt RNAs. When cells were transfected with a vector that drives the transcription a single minor satellite repeat unit, defects were detectable in both centromeric architecture (binding proteins) and its function during mitosis.\textsuperscript{44} Furthermore, this overexpression was observed to result in a wider range of histone methylation along chromosomes, which may cause ectopic foci of centromere formation, leading to impaired mitosis. In maize, CenpH3 (CENP-A homolog) co-precipitates with transcripts from the centromeric repeats\textsuperscript{299} and along with CENP-A’s established importance in kinetochore formation through the its recruitment of other kinetochore proteins\textsuperscript{56}, these facts combine to further support the idea that minor satellite transcripts are critical for proper centromere function.

Now that the interaction between non-coding transcripts of satellite RNA and centromere architecture and function is established, the question becomes one of mechanism: What is the exact contribution the 120 nt RNAs from minor satellites provide to the process of centromeric formation and function? It has been proposed that these transcripts function as a binding site/scaffolding for
the localization and recruitment of centromere-specific proteins. For this reason, these 120 nt RNAs will be referred to as centromere-associated binding RNA (cabRNA).

Given the observation that no RNAs in the size ranges of 22 nt are detected from the transcripts of minor satellite repeats, similar to those detectable from the pericentric repeats, yet northern blot results indicate RNA sizes of both 150 and 20-30 nt that bind a probe from minor satellite repeat DNA, I propose that cabRNA and casRNA function collaboratively to direct centromere localization. Although this interaction has yet to be elucidated, cabRNA through RNA-DNA binding would localize to centromeric satellite repeat DNA and would then be able through RNA-RNA binding to interact with casRNA. This second interaction would then allow for the proper localization of centromeric silencing via RITS recruitment, abrogating the transcription through the satellite regions (both major and minor). The cabRNA-casRNA interaction would also allow for the proper localization of proteins (i.e. CENP-A) responsible for centromere formation and function, an idea that supports the proposition that epigenetic mechanisms are responsible for centromere location. For this reason, it is proposed that the cabRNAs are ultimately responsible for the integrity of the genome through their role in centromere localization.

2.7 RNA Oxidative Damage

One key area has not been discussed thus far, that is, damage to RNA by ROS and the mechanisms underlying its repair. Radical oxygen species are
involved with all other intracellular macromolecules (i.e.: DNA, protein, and lipids). It would only be reasonable to assume that RNA molecules are modified as well. With the exception of tRNAs, nearly all RNAs are present in RNA-protein complexes. At certain times, RNA molecules may have a relative lack of associated proteins and during these times, the fact that it is single-stranded would imply that RNA may actually be more susceptible to damage.\textsuperscript{46,187} In a study on human atherosclerosis, RNase digestion abolished most of the cytoplasmic staining for 8-oxo-guanine (8OG) and 8-oxo-deoxyguanine (8OdG), indicating the abundance of damage produced by ROS actually involves RNA.\textsuperscript{203}

To date, 8OG is the only oxidatively-modified base that has been detected in RNA. This hydroxy-adduct can target the nucleotide pool prior to their incorporation during transcription, as well as directly attack RNA transcripts. Although this damage may be extensive, few details are known regarding any repair mechanisms. This relative lack of RNA damage control may simply be a result of the turnover rate for most RNA given that the average half-life of human mRNA is 10 hours\textsuperscript{335}. RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA) are much more stable and effects of damaged nucleotides may be more important for them.

8OG is measured to be 10-25 times more abundant than the corresponding lesion in DNA (8OdG) under oxidative stress conditions.\textsuperscript{137} Upon removal of a particular oxidative insult, these levels quickly revert (faster than the explanation of RNA turnover would indicate) to baseline implicating RNA damage surveillance mechanisms and selective removal of either these lesions, or the entire RNA transcript containing the lesion.\textsuperscript{274}
The damage mechanisms that have been elucidated are closely linked with degradation and contribute differently to the degradation of 8OG-modified RNAs. To combat this oxidative damage, *E. coli* has the proteins MutM, MutY, and MutT. As previously mentioned, the BER DNA glycosylases MutM and MutY remove integrated nucleotides sequence of either of 8-oxo-deoxyguanine or an adenine mis-integrated opposite of an 8OdG, respectively.\(^{218}\) The functional mammalian homologues to these proteins are 8-oxoG DNA glycosylase (OGG1)\(^{36,226}\) for MutM, and human MutY homolog (MUTYH)\(^{231,280,281}\) for MutY. While both repair DNA, to date, neither of these proteins nor any functional homologs, have been observed to act on RNA. The third, MutT, is the only one shown to “repair” RNA, through its action on the nucleotide pool and not by acting directly on the modified RNA transcripts themselves. MutT is responsible for the hydrolysis of the triphosphate, 8-oxo-dGTP (8OdGTP) to the monophosphate, 8-oxo-dGMP. Likewise, the human homolog MTH1 performs in much the same manner, and also hydrolyzes the ribonucleotide adduct,\(^{211}\) thus eliminating the possibility that 8OdG or 8OG from the nucleotide pool would get incorporated into a replicating DNA strand or into a transcribed RNA.

Another mechanism through which RNA damage is controlled, is indicated by the preferential binding (and incorporation) of nucleotides by RNA polymerase II. Considering the rate of incorporation for oxidatively-modified guanines is only 2% that of GTP\(^{127}\), it would be a relatively uncommon event for a damaged nucleotide to be incorporated into the nascent RNA. One final source of cellular maintenance as it relates to damaged RNA is through an interaction with a protein identified as polynucleotide phosphorylase (PnP).\(^{128}\) This protein
preferentially binds to 8OG, preventing translation and targets the RNA for
degradation. A similar protein, Y-box-binding protein 1 (YB-1), has been
reported to also bind 8OG.\textsuperscript{129}

The combined function of all these elements may be sufficient for the needs
of the cell; however, it would not be surprising to have a situation where the
levels of stress exceed the threshold of defense they provide. The concept of an
exceeded threshold is not unique, particularly in the case of non-genotoxic
carcinogens and their relationship to oxidative stress.\textsuperscript{246, 270} In fact, Rakitsky \textit{et. al.}
also added a hypothesis essentially stating that doses of compounds that
maximize this defense threshold could be deemed genotoxic, but in fact, are only
weakly so. At lower levels, these compounds would be monitored and potential
damage controlled by cellular defenses, thereby, they are observed to be non-
genotoxic as they would be removed before they were able to form adducts on
ribonucleotidisequences.

The consequent damage to RNA molecules provides numerous possibilities
for potential effects. It is speculated that through unknown RNA surveillance
mechanisms and controlled removal of damaged molecules a cell is able to
modulate the effects of RNA damage.

\subsection*{2.8 New Model of Non-Genotoxic Hepatocarcinogenesis-Altered Regulation}

In the absence of viral infection, the liver is a stable organ. Hepatocytes are
terminally differentiated, cell proliferation is low\textsuperscript{100}, and hepatocytes have a low
incidence of spontaneous mutational events\textsuperscript{117}. Hepatocytes also have defense
mechanisms in place to counteract most oxidative damage. With high levels of antioxidant enzymes, efficient DNA mismatch repair systems (evident by the absence of microsatellite instability seen in tumors), in combination with the high urinary clearance of modified DNA bases, the liver seems quite capable of dealing with the potentially harmful effects of oxidative stress.

Despite all of this, chemically-induced hepatic tumors and “pre-neoplastic” hepatic foci show indications of chromosomal alterations, differential gene expression, aberrant methylation patterns, and RNA compartmentalization, as well as high levels of ROS. Particularly in the case of non-genotoxic hepatocarcinogenesis like that of phenobarbital, it becomes increasingly apparent that the mechanisms preceding all of these changes, may involve oxidative stress and epigenetic regulatory mechanisms. Understanding the disruption in epigenetic regulatory patterns, an idea that has been supported by others, might be the key to understanding the mechanisms of chemically-induced hepatocarcinogenesis.

I propose that oxidative damage, possibly as unrepaired 8OG lesions to non-coding RNA (specifically cabRNA), targets these molecules for degradation and selectively relocates them to the cytoplasm. During their absence, the lack of DNA-cabRNA-casRNA interaction would have negative effects on chromatin structure through heterochromatin formation/centromere establishment. The lack of this interaction would eliminate RITS localization, providing an explanation for the increases in transcription observed from centromeric repeats during stress conditions.
The removal of damaged cabRNA from the nucleus, presents the potential opportunity for further damage to satellite regions responsible for cabRNA and casRNA generation causing impaired RNA-DNA or RNA-RNA interactions. It is also conceivable that during times of chronic stress, the transcription rate from minor satellite repeats (and generation of cabRNA) becomes constitutively activated, and similar to the studies with overexpression of cabRNAs, “ectopic centromeres” could be established. Chronic stress and constitutive generation of the cabRNAs thus provides the opportunity for genomic instability through improper localization of centromeres leading to mitotic errors, in turn causing aberrant gene expression of protein coding as well as non-coding genes (for example, those responsible for miRNA).

The altered compartmentalization detected in previous studies would be the result of protein binding, possibly by PnP, YB-1, or both in concert, shuttling the RNA out of the nucleus for degradation. Upon a repeated exposure to oxidative stress, the effective removal of damaged cabRNA from the nucleus (at a rate higher than that of its transcription and processing) presents an increased opportunity for further damage thus leading to a path of genomic instability. Identifying hepatocytes that have early signs of oxidative damage, an idea supported by Trush and Kensler,\textsuperscript{303} to satellite regions (using microsatellites as a representative tool) could provide the means by which to study these events.
2.9 Potential Hepatocarcinogens

2.9.1 Thioacetamide

Thioacetamide (TAA) is a chemical that has been used in the production of many materials due to its ability to donate a sulfur atom. These materials include leather, rubber, petroleum products as well as pesticides and fungicides. TAA-induced injury is predominantly detectable as centrilobular necrosis, a result of apoptosis by the activation of caspase-3.\(^{290}\)

2.9.1.1 Metabolism

Thioacetamide metabolism is primarily through its conversion by CYP2E1 to thioacetamide S-oxide (TASO). As originally proposed by Porter and Neal,\(^{244}\) a second downstream metabolite in the pathway of thioacetamide conversion was identified as thioacetamide S, S, dioxide (TASO\(_2\)). TASO\(_2\) forms adducts with the amino acid lysine. This adduct was identified as N-ε-acetyllysine\(^{90}\) \textit{in vitro} with preparations of rat microsomes. The same metabolite was identified \textit{in vivo} from livers of rats treated with thioacetamide where it was determined that 70\% of the dose was already bound to liver protein 12 hrs after dosing.\(^{89}\)

In a recent study, TAA administration to CYP2E1-null mice, exhibits no signs of liver injury/necrosis and verifies the importance of CYP2E1 to the metabolism, and thus bioactivation, of TAA.
2.9.1.2 Carcinogenicity and Mutagenicity

Thioacetamide has been linked to tumorigenicity via the development of primary hepatocellular tumors in rats fed thioacetamide. Following administration to chemically-initiated mice, “nodules” of focal hyperplasia (foci) are evident within 7 days. Under the pressure of regeneration following necrosis, these nodules may have distinct growth advantages and further develop into tumors. Within these foci, thioacetamide not only induces nuclear enlargement as previously discussed, but it also increases the transport of nuclear RNA into the cytoplasm of cells as evidenced by studies using radiolabeled RNA. It was also shown that this response was not due to diffusion across a disrupted nuclear membrane.

Direct DNA adduct formation, by TAA or its metabolites, has not been detected as a result of TAA administration, although chromosomal abnormalities have been shown. TAA is non-mutagenic as indicated by a negative result in the Ames assay.

2.9.2 Acetaminophen

Introduced in the 1950’s as an antipyretic and analgesic drug, Acetaminophen (AP) (para-acetamido-phenol), is recognized to be a hepatotoxin, causing centrilobular necrosis when taken at high doses. New warnings and recommendations were recently released by the FDA in light of hepatotoxic responses to large doses of AP. These responses are observed as an
increased elevation in liver enzymes over time. This increase is indicative of an inflammatory response that can lead to liver failure and death.

In response, the FDA panel provided recommendations that all products containing AP be appropriately labeled with the most severe warning, a black box warning, about the potential for liver damage from misuse leading to overdose. In addition, the panel also advised that, for adults, both the maximum daily dose as well as the milligram amount per tablet be reduced in accordance with these warnings.

2.9.2.1 Metabolism

Acetaminophen is primarily detoxified by phase II enzymes directly complexing it with glucuronic acid and sulfate. This primary path of detoxification has been shown to account for the presence of AP-conjugates, excreted in urine, corresponding to 80-90% of the original dose within 24 hours. AP is metabolized, predominantly by CYP2E1, although CYP1A2 and CYP3A4 have also shown metabolic conversion.

The metabolite formed is the electrophile N-acetyl-p-benzoquinone imine (NAPQI). This metabolite is detoxified by GSH conjugation; however, at high doses, can cause necrosis and apoptosis.

2.9.2.2 Carcinogenicity and Mutagenicity

The carcinogenicity and mutagenicity potential of AP has been nicely reviewed by Bergman et al. Acetaminophen and its metabolite, NAPQI, each
produce a negative result in the Ames assay as well as in CHO cells. While both have been shown to cause some genotoxic changes, namely single strand breaks and SCE, these have been attributed to the high doses used in these tests and most likely do not reflect a direct mutagenic response. NAPQI is able to cause DNA-adduct formation, but again only at high doses.

The toxicity of AP can be explained by either the depletion of GSH allowing NAPQI to preferentially bind thiol groups on critical cellular proteins, or by the metabolites’ induction of oxidative stress. Carcinogenicity studies in mice have been predominantly negative, and epidemiologic studies have shown a similar result indicating there is no link between AP and carcinogenicity.

2.9.3 Aflatoxin B₁

Aflatoxin B₁ (AB₁) is the most potent of all mycotoxins produced by several mold species, including Aspergillus flavus. This genus of molds is a common contaminant of food products, particularly grain, and has severe effects, both hepatotoxic and hepatocarcinogenic, in livers following ingestion.

2.9.3.1 Metabolism

Bioactivation of AB₁ by CYP3A4 creates the reactive epoxide, AB₁ exo-8, 9-epoxide. Glutathione-S-transferases (GSTs) are able to conjugate this metabolite and clear it from the body in urine. The majority (90%) of the ingested dose is metabolized in the liver to the non-toxic hydroxylated derivatives, aflatoxins M₁ and Q₁, and excreted in feces and urine within 24 hours.³²⁸
2.9.3.2 Carcinogenicity and Mutagenicity

AB$_1$ is considered to be one of the most potent hepatocarcinogens. The epoxide metabolite of AB$_1$ is mutagenic and forms a stable adduct at the N7 position of guanine. This adduct however, is efficiently removed from DNA and excreted in urine. Epidemiologic studies indicate only a causal relationship between AB$_1$ and hepatocellular carcinomas.$^{1,2}$ Often these relationships are confounded by the presence of a viral infection with HBV.$^{39}$ In this aspect it could be deemed non-genotoxic since at low dose levels the epoxide would only have a rare opportunity to bind ribonucleotides.

A similar phenomenon was observed in a study of another mycotoxin, Ochratoxin A (OTA), a toxin linked with renal cancer.$^{13}$ Not only did the authors observe OTA genotoxicity only at levels that are cytotoxic and would promote an inflammatory response, but they also showed a rapid response based on mRNA expression of key oxidative stress genes, such as SOD and catalase, following treatment, thereby further supporting the existing link between AB$_1$ induction of oxidative stress$^{273}$ and carcinogenesis.

2.9.4 Nodularin

Nodularin (Nod) is a hepatotoxin synthesized by the cyanobacteria Nodularia spumigena, originally identified in 1878 from a cyanobacterial bloom on Lake Alexandria, Australia, that caused massive animal death due to liver hemorrhage. N. spumigena blooms occur as a “scum” in brackish water and are
particularly problematic in the Baltic sea where the blooms have been associated with fish death.

Nod belongs to a family of other cyanobacterial hepatotoxins, namely the microcystins, though unlike them, Nod consists of a cyclic pentapeptide ring, whereas the structure of microcystins is that they are cyclic heptapeptides. While their compound structures are very similar (Figure 2-3), the mechanisms behind their toxicity, is unique.

2.9.4.1 Metabolism

Little is currently known regarding the bioactivation of Nod via the CYP450 family, however levels of glutathione are altered upon treatment.

2.9.4.2 Carcinogenicity and Mutagenicity

Both microcystins and Nod have been shown to inhibit protein phosphatases 1 (PP1) and 2A (PP2A), although the mechanisms by which they
act are independent. Microcystins bind covalently with PP1 and PP2A. Nod, on the other hand, does not\textsuperscript{18} and instead induces lipid peroxidation,\textsuperscript{41} a key indicator of ROS generation.

In the Ames assay, Nod yields a negative result. Several studies\textsuperscript{180, 198} have examined the genotoxic potential of Nodularin and found an increase in 8-OdG adducts. They also report a high incidence of micronucleus formation as well as strand breaks. Upon further analysis, these studies include models such as HepG2 cells, a hepatoblastoma cell line, or primary rat hepatocytes treated with high dose levels, each of which could obscure the true effects of Nod. Recently, Bouaicha, \textit{et. al.} concluded that the genotoxicity observed is most likely the result of secondary effects.\textsuperscript{42} This conclusion was based upon the dose required to yield guanine adducts, the rapid response of anti-oxidant enzymes, like GST, once again indicative of ROS generation, as well as the removal of 8-OdG within 24 hours following treatment.

The reports of genotoxicity would seem to support the carcinogenicity studies of Ohta \textit{et. al.}\textsuperscript{230} In this study, Nod was determined to induce GST hepatic foci following intraperitoneal (i.p.) injection, however, since this is a method that would not reflect a true route of exposure the results would most likely not be validated by a study using oral dosing.

Epidemiologic studies\textsuperscript{139, 341} that have associated primary liver cancer with the drinking of ground water, polluted with cyanobacteria, have only been reported in areas that also have high levels of Aflatoxin B\textsubscript{1} exposure or high incidence of HBV and HCV.\textsuperscript{84, 104} Again, the ingestion of Nod may not be causally related to the development of hepatocellular carcinomas. No other
carcinogenicity studies have been performed to study the effects of Nodularin. In 2006, the International Agency for Research on Cancer determined there was not enough evidence to confirm the carcinogenicity of Nodularin, deeming it “not classifiable”.\textsuperscript{122} Still, the induction of ROS following treatment might just be the key to link Nodularin exposure and hepatocarcinoma.
Chapter 3
Microsatellite-Containing Bicistronic Reporter Constructs: Generation and Validation

3.1 Introduction

The use of reporter genes has been used for several decades in order to examine proteins such as their localization within cells or tissues or to identify protein interactions with binding partners. In addition, reporter genes have been used to identify mechanisms such as those involved in mutation events. Some of the most commonly used reporter genes are genes that code for green fluorescent protein (GFP), beta-galactosidase (β-gal), and luciferase.

For these studies I selected GFP and β-gal. By utilizing the two genes that encode them, I would avoid any possibility of confounding data, as GFP is a naturally bioluminescent protein, while β-gal protein monitoring relies on additional manipulation such as an enzymatic reaction with a substrate or recognition with antibodies. Due to this distinction, the activity of each protein could be monitored uniquely.

It is my intention to observe microsatellite frameshift events as indicators of early genomic instability. The β-gal-encoding lacZ gene was modified to include microsatellite sequences. I devised a dual reporter system that incorporates the genes for each of the proteins on the same plasmid. By combining them onto a single plasmid that has only a single promoter, I am able to control for the transcription of the reporters, even if the sequence of lacZ was in an out-of-frame
configuration. Additionally, it is my intention to utilize these constructs to study early events in chemically-induced hepatocarcinogenesis. Therefore, I incorporated a liver specific promoter into the plasmids so that I could observe solely liver responses.

This chapter will show the process, by which, I developed these plasmids as well as early indication of their function.

3.2 Cloning of pβE Plasmid

Bacterial lacZ, the gene encoding the enzyme β-galactosidase, was subcloned into a bicistronic vector (pIRES2-EGFP). As indicated by the name, this plasmid contained the gene for EGFP (enhanced green fluorescent protein). GFP was originally isolated from the Pacific-northwestern jellyfish Aequorea victoria (A. victoria); however, it was then modified via a point mutation (Ser 65 → Thr) to create a protein that exhibited increased fluorescence, a single as well as longer wavelength of excitation and increased photostability.

3.2.1 Method

The vector pCMVβ (Figure 3-1:Panel A) (Clontech, Mountain View, CA) was digested with SacII. Another restriction enzyme, PstI, was used to distinguish the digested fragments of approximately equal size, created by SacII digestion, by further digesting one of the pieces that did not include the lacZ gene. The largest fragment then contained the lacZ gene. This fragment was
ligated into pIRES2-EGFP (Figure 3-1:Panel B) (Clontech), previously digested with SacII, located within the multiple cloning site (MCS).

**3.2.2 Function**

The new pIβE vector was transfected into cells to test the functionality of both reporter genes. GFP expression was microscopically detectable by excitation with a blue light. After fixation and staining of the cells using X-gal as the substrate, transfected cells were detectably blue in color (data not shown).

**3.2.3 Mutation of pIβE**

One additional modification needed to be performed in order to proceed with the creation of the microsatellite repeat vectors. This modification was that a

---

Figure 3-1: Vectors used in creation of bicistronic vector pIβE. Panel A: pCMVβ, Panel B: pIRES2-EFGP. (Source: Clontech product inserts)
novel restriction site was needed in the lacZ sequence to allow for further cloning manipulations.

A single nucleotide substitution (G→T), created by the QuikChange™ XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA,) in the upstream region of lacZ, produced the restriction site for Kpn21. This alteration, shown in Figure 3-2, is a silent mutation, in that the amino acid, Valine (Val), identified by the codon sequence that contains the mutation remains the same. By maintaining the original amino acid I was able to prevent any alterations in the shape of the β-gal protein ensuring that the functional activity of the protein would remain the same. I named this plasmid pIβE-mutKpn.

![Key: Blue ATG is start codon for β-gal gene Red T is mutated base (originally a G) Highlighted and underlined sequence is new restriction site, Kpn21](image)

**Figure 3-2:** Sequence of bicistronic vector pIβE-mutKpn. Single point mutation (shown as a red T) creates novel restriction site Kpn21.
3.3 Inclusion of Microsatellite Repeats –pIβE-ntR Vectors

Microsatellites are short tandem repetitive sequences of DNA that occur throughout the genome. They occur as lengths of 1-6 nucleotides, which are repeated as units. Alterations in microsatellite sequences show them to only be affected in terms of these units; that is, alterations in lengths of microsatellites occur as gains, or more commonly, losses of entire repeated units. The plasmids designed all contain microsatellite sequences of different lengths, unless specifically denoted, the vectors as a whole group, will be referred to as pIβE-nucleotide repeat (pIβE-ntR) vectors.

3.3.1 Di-nucleotide Repeats

Di-nucleotide (di-nt) repeats are units consisting of 2 base pairs that result in a repeating segment of DNA. The goal was to incorporate a repeating sequence \((CA)_n\) into the lacZ sequence, where “n” represents the number of units. The \((CA)_n\) dinucleotide was selected based upon its abundance within mammalian genomes.\(^{125}\)

The incorporation of microsatellite sequences would yield a positive β-gal protein expression by maintaining an “in-frame” codon sequence, or would alter the sequence so that additional repeat units caused a frameshift of \((n+1)\) resulting in an “out-of-frame” codon reading sequence from which no protein would be produced. In order for the protein to be produced, the repeated region would have to lose one (CA) repeat so the sequence was shifted back to an in-frame
codon configuration. Alternatively, the restoration of codons could also be obtained by the insertion of 2 repeat units.

3.3.1.1 Method

Di-nucleotide repeats ((CA)$_{17}$-in and (CA)$_{18}$-out) were incorporated (including an 8 nucleotide linker sequence) into the lacZ gene of pIβE-mutKpn via polymerase chain reaction (PCR) using Expand High Fidelity (Roche), a Taq-based kit that contains a proofreading component to eliminate the potential for mutations during PCR. Each of these repeats was placed immediately downstream of the methionine (ATG) start codon of the LacZ gene, so as to increase the chance that a truncated β-gal protein would not be created. The length of sequence selected for the out-of-frame repeat creates a premature stop codon at the newly created codon 20. Because of this, I am confident that the protein is not expressed even as a truncated form.

Primers were designed that included repeats as shown in Figure 3-3. Several of these primers also included a BamHI site so the amplification products could be combined through ligation. pIβE-mutKpn was used as the template for PCR amplification using primer pairs (1068/1072) or (1069/1072) to result in out-of-frame and in-frame amplicons respectively. Another PCR was performed with the primer pair 1070/1449 with pIβE-mutKpn again as template.
Once the amplicons were obtained, all three were digested with BamHI and ligated in pairs combining, 1072/1068 with 1070/1449, or 1072/1069 with 1070/1449. When the incubations were complete, each of the 2 newly ligated products were amplified via PCR with the primers 1072 and 1449, which now corresponded to the ends of the ligation products containing in- and out-of-frame microsatellite sequences.

These amplified products containing the di-nt microsatellite repeats were digested with XhoI and Kpn21 and ligated into pIβE-mutKpn vector (also digested with same enzymes), yielding the full plasmid containing di-nt microsatellite repeats.

Figure 3-3: Primers designed to incorporate di-nucleotide repeat sequences into plasmid pIβE-mutKpn.

Key:
Highlighted and underlined sequence is restriction site, BamHI
### 3.3.1.2 Sequencing

The di-nt repeat vectors pIβE-D\textsubscript{IN} (Figure 3-4) and pIβE-D\textsubscript{OUT} (Figure 3-5), corresponding to di-nt in-frame and di-nt out-of-frame respectively were transformed into *Escherichia coli* (*E. coli*) cultured, isolated, and sequenced.

5\'- CAACGCTAGC GCTACGGGAC TCAGATCTCG AGCTCAAGCT TCGAATTCTG  
  CAGTCCGCGG TACCAGGGCC GCAATTCCCC GGGATCGAAA GAGCCTGCTA  
  AAGCAAAAA GAAGTCACCA TGCGCACACA CACACACACA CACACACACA  
  CACACACAGG ATCTTCGTTT ACTTCTGACCA ACAAGAAGGT GATTTTCTGT  
  GGGATCCAGG GAGCAATTGG TCTGGACACC AGCAGGAGCC TGCTCAAGCC  
  CGATTTCAAA TGTGCAGTGT GCAGTTCCCT GCTGATCTGC GTTGACCCAC  
  GAACATCCAA CTATCCCTTT TCTGACAGCT GGGATCTAGC GCTGATCTGC  
  GGCTGAGTGC ATGCAGGCCC CACCTCTTCC CAGGCCCTGG AGCTGCTGCT  
  ATCCTCTTCC CAGGCCCTGG GCGTCGTCAC ATGGTTGATT GATTTTCTGT  
  GATTTTCTGT GGTGCAACGG GCGCTGGGTC NGTTACGGCC AGGACAGTCG TTTGCCGTCT  
  TGAGCGCATT TTACGCGCCG GAGAAAACCC CTGCGGCTGA TGGNGCTTNC  
  CTGGAGTGAC GGAAGTNTCT GGAAATCAG -3'  

**Figure 3-4:** Sequence of pIβE-D\textsubscript{IN}.
Mono-nucleotide Repeats

Similar to the di-nt repeats described above, mono-nucleotide (mono-nt) repeats are regions of DNA sequence consisting of a repeated single nucleotide. The goal was to incorporate a repeating sequence \((G)\_n\) into the lacZ sequence, where “n” again represents the number of units, in this case a single G. A repeating sequence of guanines was selected for its propensity to obtain damaging adducts, particularly that of oxidative damage. Further detail regarding this selection will be discussed in a later chapter.

Again, this incorporation would yield a β-gal protein that was maintained by an “in-frame” sequence of codons, or so that the number of units incorporated
caused a frameshift of (n+1) resulting in an “out-of-frame” sequence from which no protein would be produced. Similarly again to the di-nt sequences, in order for the protein to be produced, the repeated sequence would either have to lose one repeat, in this case one single guanine (G), or gain 2 guanines, so the sequence was shifted back to an in-frame codon configuration.

3.3.2.1 Method

Mono-nucleotide repeats (G$_{9}$-in and G$_{10}$-out) were incorporated into pIβE-mutKpn using a fusion PCR technique and Expand High Fidelity Taq polymerase (Roche). Each of these repeats was placed slightly downstream of the ATG of the lacZ gene, which, as in with the di-nt repeats, eliminates or significantly reduces the chance that a truncated protein would be created.

The repeats were also placed in the part of the lacZ sequence that already contained a stretch of 3 guanines. By locating the repeats within this triplet, I increased the length the final microsatellite to runs of 12 and 13 guanines respectively, which in turn, could increase the likelihood for a mutation within this sequence and thus its detection.$^{47,176}$ The length of sequence selected for the out-of-frame repeat creates a premature stop codon at the newly created codon 42. As with the di-nt repeats, I am confident that the β-gal protein is not expressed, even as a truncated form.

PCR was performed in pairs, with primers that were designed including desired repeats as shown (Figure 3-6) based on pIβE-mutKpn as a template. In the first pair of amplifications, for the in-frame repeats, one of the reactions
contained primers 1072 and 1452, while the other, used 1451 and 1449. Once the reactions were completed, the amplification products were run on agarose gels and extracted to eliminate any template vector.

Each of these products was combined into a single tube as a template for a subsequent round of PCR with primers 1072 and 1449. Based upon the sequences for each of the primers, the previous amplicons were able to anneal and the PCR reaction continued.

\[
\begin{align*}
\text{CLAW 1072(F):} & \quad 5'\text{- GACTCAGATCTCGAGCTCAAGCTTC } -3' \\
\text{CLAW 1449(R):} & \quad 5'\text{- CAGCCAGCTTTCCGGAACCGC } -3' \\
\text{CLAW 1451(F):} & \quad 5'\text{- GTTGCCGCTCTGGGGGGGGGGGAGGCATTGGTCTG } -3' \\
\text{CLAW 1452(R):} & \quad 5'\text{- CAGACCAATGCCTCCCCCCCCCAGACCGGAAC } -3' \\
\text{CLAW 1453(F):} & \quad 5'\text{- GTTGCCGCTCTGGGGGGGGGGGAGGCATTGGTCTG } -3' \\
\text{CLAW 1454(R):} & \quad 5'\text{- CAGACCAATGCCTCCCCCCCCCAGACCGGAAC } -3'
\end{align*}
\]

Figure 3-6: Primers design for incorporation of mono-nucleotide repeat sequences into plasmid pIβE-mutKpn.

Likewise, the primers that contained the out-of-frame sequence were used in a similar manner as described above, with primers pairs of 1072/1454 and 1453/1449.

Both amplicons (in- and out-of-frame) resulting from the round of fusion PCR were digested with the restriction enzymes, XhoI and Kpn21 and ligated into pIβE-mutKpn vector, once again also digested with same enzymes.
3.3.2.2 Sequencing

The mono-nt repeat vectors pIβE-M\textsubscript{IN} (Figure 3-7) and pIβE-M\textsubscript{OUT} (Figure 3-8), corresponding to mono-nt in-frame and mono-nt out-of-frame respectively, were transformed into \textit{Escherichia coli} (\textit{E. coli}) cultured, isolated, and sequenced.

![Figure 3-7: Sequence of pIβE-M\textsubscript{IN}.

![Figure 3-8: Sequence of pIβE-M\textsubscript{OUT}.](image)
3.3.3 Function of pIβE-ntR Vectors

pIβE-ntR vectors were transfected with Lipofectamine Plus transfection reagent (Invitrogen) into human embryonic kidney cells (HEK-293T) in order to examine their activity by the presence or absence of β-galactosidase protein activity. Following transfection, cells were incubated, fixed, and stained using X-gal as a substrate. Blue cells indicative of the presence of β-gal from the cells transfected with the in-frame construct were evident. The lack of blue cells in those cells transfected with out-of-frame constructs indicated that indeed the frameshift created was sufficient to eliminate β-gal protein expression (data not shown).

3.4 Incorporation of Liver-Specific Enhancer/Promoter -pαβE-ntR Vectors

Based upon the initial observations in this laboratory, which showed focal changes in rat livers following low dose carcinogen treatment, I decided to direct the expression of the microsatellite-containing reporter constructs to liver cells. As attempts to use the mouse albumin enhancer-promoter were unsuccessful, I opted to use a newly-constructed, liver-specific, chimeric enhancer-promoter (EALB/Pα1AT).

3.4.1 Design and Function of EALB/Pα1AT

The liver-specific chimeric promoter, EALB/Pα1AT, developed by Dra. M. Gabriela Kramer, was constructed using the mouse albumin (ALB) enhancer region (GenBank Acc. #U04199) linked to the human α1-antitrypsin (α1AT) promoter.
(GenBank Acc. #D38257) with a BglII restriction site (Figure 3-9). Incorporating non-viral sequences into this new chimeric enhancer/promoter sequence, she was able to increase the potential for this sequence to be utilized in future plasmids by avoiding the possibility for viral promoter silencing that has been shown to occur in cells as well as transgenic mice.\textsuperscript{329}

3.4.1.1 Activity of Chimeric Enhancer/Promoter in Cell Culture

In liver cell lines, the activity of $E_{\text{ALB}}/P_{\alpha_{1}\text{AT}}$ has been previously studied directing the expression of luciferase and exhibited an activity that was approximately 77\% of the activity of the cytomegalovirus (CMV) promoter (Figure 3-10). In non-liver derived cell lines, no activity from this promoter was evident.

---

Figure 3-9: Chimeric liver-specific enhancer/promoter $E_{\text{ALB}}/P_{\alpha_{1}\text{AT}}$. Underlined sequences indicate restriction sites for XhoI, BglII and HindIII, respectively.
3.4.1.2 \textit{In Vivo-Mouse Tail Vein Injection Kinetic Activity Study}

The functional activity of $E_{ALB}/P_{a1AT}$ was also assessed \textit{in vivo}. Hydrodynamic tail vein injections were performed in mice with various enhancer-promoter constructs driving transcription of the human $\alpha 1$AT gene. As an indication of promoter activity, blood serum levels of $\alpha 1$AT were monitored over time. Figure 3-11 shows the continually high levels of serum $\alpha 1$AT driven by $E_{ALB}/P_{a1AT}$ indicated by the red line, while levels decreased over time from constructs that included viral sequences, as evidenced by the green line, representing a construct including a sequence for the viral enhancer, EII, from the Hepatitis B virus (HBV).
In order to confirm this function, the promoter small plasmid was designed to test not only the function but also the specificity of this promoter in cell culture. $E_{ALB}/P_{\alpha1AT}$ was inserted upstream of EGFP as shown in Figure 3-12. This new vector, named EPE-pVAX1, was transfected into cells and GFP fluorescence activity was monitored after 48 hours.

**Figure 3-11:** Kinetic activity of $E_{ALB}/P_{\alpha1AT}$ in vivo. Red line indicates the activity of $E_{ALB}/P_{\alpha1AT}$. Figure taken from Kramer, M.G. *et al*. Molecular Therapy (2003).

### 3.5 $E_{ALB}/P_{\alpha1AT}$-EGFP/pVAX1 (EPE-pVAX1)

In order to confirm this function, the promoter small plasmid was designed to test not only the function but also the specificity of this promoter in cell culture. $E_{ALB}/P_{\alpha1AT}$ was inserted upstream of EGFP as shown in Figure 3-12. This new vector, named EPE-pVAX1, was transfected into cells and GFP fluorescence activity was monitored after 48 hours.
Examination of cells after transfection showed expression of EGFP in liver-derived cell lines only. In HepG2 cells, which are a human derived tumor cell line, the activity, based upon percent transfection confirmed the activity of the chimeric E_ALB/P_α1AT as approximately 77% of the activity from a similar-sized vector (CMV-EGFP/pVAX1), designed using the CMV promoter to drive expression of EGFP.

In normal mouse liver cells that were derived from weanling NAMRU mice (NMuLi), promoter activity was reduced, but still present (data not shown). The weaker expression observed is attributable to the fact that P_α1AT is a human
sequence-based promoter that could have weaker activity in the context of the mouse cells.

Figure 3-13: Promoter activity test in Liver (HepG2) and non-liver (293T) cells. Panels A and B: HepG2 and 293T cells (respectively) transfected with CMV-EGFP/pVAX1, Panels C and D: HepG2 and 293T cells (respectively) transfected with EPE-pVAX1.

3.6 CMV Replacement with $E_{ALB}/P_{\alpha 1AT}$ in “-ntR” Vectors

$E_{ALB}/P_{\alpha 1AT}$ was amplified from graciously donated $E_{ALB}/P_{\alpha 1AT}$–luc vector using PCR and primers that contained AseI and XhoI recognition sites. When the amplification product was ligated into pI$\beta$E-D$_{OUT}$, previously digested with the same enzymes, a single point mutation was located in the promoter sequence. This was corrected by replacing the BgIII-EcoRI fragment with the same fragment subcloned from EPE-pVAX1. Sequencing of the new plasmid (p$\alpha$$\beta$E-D$_{OUT}$) yielded the intended correct sequence.
The replacement of the CMV promoter in the remaining plasmids was performed by subcloning the AseI-EcoRI fragment from pαβE- D\textsubscript{OUT} into the three remaining plβE-ntR vectors previously digested with the same restriction enzymes. The sequences of the newly created pαβE-ntR vectors are shown in Figure 3-14.

<table>
<thead>
<tr>
<th>727</th>
<th>3.6.1 Function</th>
</tr>
</thead>
</table>
| Cells were transfected using Lipofectamine Plus reagent (Invitrogen). 48 hours after transfection, cells were fixed and stained overnight using X-gal. β-gal activity and thus presence, was assessed by the detection of blue cells. Numerous

Figure 3-14: Microsatellite repeat sequences of pαβE-ntR plasmid sequences.
transfections with the out-of-frame vectors indicated no presence of blue cells, confirming that the vectors were not mutated during their replication in *e. coli*.

### 3.7 Discussion

Based upon the function of these reporter constructs confirming the expected results, I have effectively created a plasmid-based model system for the assessment of frameshift mutations in both di-nucleotide and mono-nucleotide microsatellites. Future studies using cell lines and transgenes are now possible; each of which will be of primary focus in the upcoming chapters.
Chapter 4

Utility of Reporter Constructs for Assessing Effects of Hepatocarcinogen Treatment in Murine Liver Cells

4.1 Introduction

The underlying mechanisms of chemically-induced hepatocarcinogenesis are not always clear. During metabolism, most xenobiotics are converted by the CYP450 family of enzymes into polar intermediates that are able to react with larger biomolecules like glutathione and, in turn, excreted. Sometimes however, this conversion yields a product known as an “ultimate carcinogen” that directly interacts with DNA sequence causing adducts. These compounds are referred to as being “genotoxic”. These adducts, if not properly repaired, induce mutations in the targeted genes causing gene expression differences. If the genes affected play a role in cellular homeostasis, a pattern of increasing gene alterations may develop. However, there are other compounds that cause hepatocarcinogenesis in rodent 2-year bioassays that show no indication of DNA adduct formation, mutation via the Ames assay, or chromosomal alterations, such as sister-chromatid exchanges or strand breakages. These compounds are referred to a “non-genotoxic”. Many of these compounds are shown to be tumor promoters in DEN-initiated rodent studies, and are thought to act mainly in this capacity, yet as they alone can form tumors, they are carcinogenic. Given that no adducts or DNA alterations occur, the mechanism(s) of their action remains a mystery.
Early studies in this laboratory, reported that rats treated with low, non-toxic doses of carcinogens developed hepatic foci, morphologically-altered regions of cells, that had characteristic enlarged nuclei. These foci also showed altered cytoplasmic compartmentalization of RNAs normally located only in the nucleus. Upon further investigation, these RNAs were shown to consist of repetitive sequences from the B2 sequence family. Additionally, the histologically-normal cells surrounding these foci stained positive for 4-hydroxy-2-nonenal (4-HNE), a product of oxidative damage.

Genomic instability has been linked to the development of cancer. In cells that show indications of genomic instability, microsatellites, are particularly prone to frameshift mutations. While robust DNA replication and errors during DNA repair have been implicated in numerous cancers involving microsatellites, in the case of chemically-induced hepatocarcinogenesis, these mechanisms have not clearly factored into the process. The induction of oxidative stress has also been associated with non-genotoxic carcinogenesis, although again, the exact mechanisms are unknown.

In this chapter I aim to examine the role of oxidative stress and frameshift events in the context of non-genotoxic carcinogen treatment by using novel dual-reporter, microsatellite-containing plasmids.
4.2 Methods

4.2.1 Mouse Metabolism Analysis

600 ng of RNA, extracted from NMuLi cells with RNeasy Mini Kit (Qiagen), was treated with DNase to eliminate all genomic DNA, and reverse-transcribed using the RT² First Strand Kit (C-03) (SABiosciences, Fredrick MD). The RT reaction product was combined with SuperArray RT² qPCR Master Mix and loaded onto an RT²Profiler™ PCR array specific for mouse metabolism (Cat# PAMM-002). Following 40 cycles of QPCR amplification using SYBR green detection, data was analyzed in combination with data obtained from a separate RNA preparation from the same cell line using the PCR Array Data Analysis Web Portal (www.sabiosciences.com). Quality control analysis was also run via the same application utilizing the housekeeping genes integrated into the array plate.

4.2.2 Microsatellite-Based Frameshift Reporter Cell Line Design

Liver-specific, bicistronic, dual reporter plasmids (pαβE-ntR vectors) containing genes for EGFP and β-galactosidase, were constructed as previously described (see chapter 3). NMuLi cells (Cat# CRL-1638, ATCC, Manassas, VA), originally isolated from NAMRU mice, were transfected with plasmids using Lipofectamine 2000 transfection reagent (Invitrogen) and incubated at 37°C overnight. Each vector was used for two independent transfections into NMuLi cells. Stable populations of pαβE-ntR were derived by selecting transfected cells
with 500 µg/mL Geneticin (Cat# 10131, Gibco-Invitrogen) and further enriched, based upon GFP expression by sorting the selected populations on a Cytomation MoFlo High Performance Cell Sorter. NMuLi enriched populations were maintained in selection media (DMEM containing 10% FBS and 500 µg/mL Geneticin).

### 4.2.3 Analysis and Characterization of Morphology Changes in NMuLi Cell Line

NMuLi cells displaying morphological changes were split into 2 groups for analysis. The first was plated on Lab-Tek II chamber slides (Nunc, Gibco, Invitrogen, Carlsbad, CA), fixed with 70% ice-cold methanol, and Hemotoxylin/Eosin (H+E) stained. Alternatively, the second group was plated into wells of a 24-well plate (Falcon, Gibco, Invitrogen) and fixed with 2% formaldehyde and 0.05% glutaraldehyde in 1x PBS for 5 minutes. After fixation, cells were stained for β-gal overnight at 37°C using X-gal (Invitrogen) as substrate in a 1x PBS staining solution containing 50mM each potassium ferricyanide (K₃Fe(CN)₆) and potassium ferrocyanide (K₄Fe(CN)₆).

The mixed population was sorted using a Cytomation MoFlo High Performance Cell Sorter one cell each into individual wells of a 96-well plate. These individual cells were grown under selection media (as described previously) until clonal populations were obtained. These cloned cells were subsequently stained for β-gal activity using X-gal as before.
4.2.4 Immunohistochemical Analysis

NMuLi cells as well as the mixed population, arising from NMuLi cells in culture, were plated onto chamber slides, fixed in 70% ice-cold methanol for 5 minutes, rehydrated in 1x PBS for 1 hr. at RT, and stained for albumin expression with the Vector ImmPress Kit (Cat #MP-7401, Vector Laboratories, Inc., Burlingame, CA) with rabbit anti-mouse albumin antibody (ICN cappel #55462, now MP Biomedicals, Solon, OH).

The staining procedure included an endogenous peroxidase block using 0.3% H$_2$O$_2$ solution for 30 minutes as well as the 2.5% normal horse serum block included in the ImmPress kit. The primary antibody was diluted in a 1:1 solution of normal horse serum and wash buffer, consisting of 1x PBS+0.3% Triton X-100. Vector VIP Substrate Kit (Cat #SK-4600, Vector Labs) was used to develop the staining resulting in a purple reaction product. Once staining was complete, chambers were removed and slides were coverslipped using VectaMount (Cat #H-5000, Vector Labs).

4.2.5 Establishment of LD$_{50}$s in NMuLi Cells

NMuLi cells were plated in 96-well plates and incubated at 37°C overnight. Cells were treated with compounds at several doses. After 24 hours, the treatment media was removed via plate inversion and blotting. Cells were then incubated at 37°C in 100 µL media containing 10 µL CCK-8 reagent (Dojindo Molecular Technologies, Gaithersburg, MD) for 1 hour before being measured at
450 nm in a Synergy HT plate reader (BioTek Instruments, Inc, Winooski, Vermont).

**4.2.6 Measurement of ROS Generation**

NMuLi cells, in a clear bottom-black sided 96 well plate, were incubated with 100 µM 2’, 7’-Dichlorofluorescin Diacetate (DCFH-DA) (Calbiochem, EMD Chemicals, San Diego, CA) for 30 minutes in DMEM normal growth medium. Following incubation at 37°C, after media was removed, by plate inversion and blotting, Krebs-Ringer-HEPES buffer was added to each well containing several dilutions of compounds. Plates were immediately placed in the Synergy HT plate reader previously pre-heated to 37°C, and measured at 30-minute intervals for 2 hours.

**4.2.7 Chemicals**

The chemicals used for these experiments are listed: Aflatoxin B₁ (AB₁) (Cat #A6636, Sigma, St. Louis, MO); Acetaminophen (AP) (Cat #A7085, Sigma); Hydrogen Peroxide (H₂O₂) (Cat #216763, Sigma); Nodularin (Nod) (ICN 15838090, Fisher, Waltham, MA); and Thioacetamide (TAA) (Cat #163678, Sigma). All chemicals were selected based on their unique properties: inducer of frameshift mutation (H₂O₂), potential liver carcinogens (TAA, AB₁, and Nod), and liver toxin (AP).
4.2.8 Compound Treatments

Stable pαβE-ntR\textsubscript{OUT} cell lines, expressing microsatellite repeats in an out-of-frame-configuration, were plated into 48-well plates and incubated at 37°C overnight. The following day, the cells were incubated for 30 minutes in growth media containing 100 µM DCFH-DA. After incubation, media was removed from wells and KRH buffer containing compounds was added to wells. Plates were immediately placed in the plate reader, pre-warmed to 37°C, monitored for 2 hours by measurements at 30-minute intervals in order to confirm ROS generation.

Once the generation of ROS was confirmed, plates were removed from the plate reader. KRH buffer containing compounds was removed from wells and replaced with fresh growth media and cells were then incubated for 72 hours at 37°C. After 72 hours, cells were trypsinized and divided into 2 groups. The first group was replated and incubated for another 48 hours. The second group was fixed and stained for flow cytometric analysis.

4.2.9 Flow Cytometry

Cells treated with compounds or vehicle controls were trypsinized and fixed for 5 minutes at room temperature with Reagent A of the CalTag Fix & Perm\textsuperscript{®} Kit (Invitrogen) diluted 1:3 in intracellular staining buffer (ISB), PBS containing 1% heat-inactivated FBS and 0.09% (w/v) sodium azide. Cells were washed by the addition of 1 mL ISB, vortexed briefly, centrifuged, and supernatant removed.
Cells were resuspended with anti-β-gal primary antibody (ab616-Abcam) at a 1:750 dilution in 50 µL of the cell permeabilizing Reagent B, of the CalTag kit, diluted 1:4 with ISB (Reagent B*). Resuspensions were incubated on ice for 45 minutes and washed as described above. Secondary antibody, R-phycoerythrin (RPE) anti-rabbit (Cat# P2771MP-Invitrogen) was diluted 1:200 in Reagent B*. Cell pellets were resuspended and incubated on ice again for 20 minutes. Finally, cells were washed, resuspended in ICS and run on a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA) to assess PE (red) and GFP (green) fluorescence.

4.3 Results and Discussion

4.3.1 Gene Expression Profile of NMuLi Cells

In order to assay for chemically-induced frameshift events, I selected the normal mouse liver (NMuLi) cell line based upon it being a spontaneously-immortalized, non-transformed cell line isolated from the NAMRU mouse line. Since frameshift events occur with greater frequency with genomic instability, this cell line removed the possibility for additional frameshifts created by genomic instability induction from the SV40 T antigen or inherent genomic instability, a characteristic of tumor-derived cell lines.

To determine the “normal” metabolic function of this cell line, a gene profile for mouse metabolism was established using RNA isolated from untreated NMuLi cells. Gene expression, showing the presence or absence of key phase I
and phase II enzymes, is summarized in Table 4-1. (For the entire metabolism profile: see Appendix)

### Table 4-1: Phase I and II gene expression in NMuLi cells

Presence/Absence scores are displayed corresponding to Ct values, as follows: +++, Ct < 30; +, 30 < Ct < 35; -, Ct > 35. Values based on recommendations by SABiosciences (personal communication).

<table>
<thead>
<tr>
<th>Gene description</th>
<th>Symbol</th>
<th>Refseq</th>
<th>Ct</th>
<th>Presence/Absence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol dehydrogenase 1 (class I)</td>
<td>Adh1</td>
<td>NM_007409</td>
<td>28.725</td>
<td>+++</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 4 (class II), pi polypeptide</td>
<td>Adh4</td>
<td>NM_011996</td>
<td>30.09</td>
<td>+</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 5 (class III), chi polypeptide</td>
<td>Adh5</td>
<td>NM_007410</td>
<td>19.225</td>
<td>+++</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase family 1, subfamily A1</td>
<td>Aldh1a1</td>
<td>NM_013467</td>
<td>24.67</td>
<td>+++</td>
</tr>
<tr>
<td>Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1</td>
<td>Chst1</td>
<td>NM_023850</td>
<td>27.27</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 1, subfamily a, polypeptide 1</td>
<td>Cyp1a1</td>
<td>NM_009992</td>
<td>28.38</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 1, subfamily a, polypeptide 2</td>
<td>Cyp1a2</td>
<td>NM_009993</td>
<td>34.46</td>
<td>+</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily b, polypeptide 10</td>
<td>Cyp2b10</td>
<td>NM_009998</td>
<td>38.77</td>
<td>-</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily c, polypeptide 29</td>
<td>Cyp2c29</td>
<td>NM_007815</td>
<td>33.69</td>
<td>+</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily e, polypeptide 1</td>
<td>Cyp2e1</td>
<td>NM_021282</td>
<td>29.08</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 4, subfamily b, polypeptide 1</td>
<td>Cyp4b1</td>
<td>NM_007823</td>
<td>24.38</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 11, subfamily b, polypeptide 2</td>
<td>Cyp11b2</td>
<td>NM_009991</td>
<td>21.22</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 17, subfamily a, polypeptide 1</td>
<td>Cyp17a1</td>
<td>NM_007809</td>
<td>No Ct</td>
<td>-</td>
</tr>
<tr>
<td>Cytochrome P450, family 19, subfamily a, polypeptide 1</td>
<td>Cyp19a1</td>
<td>NM_007810</td>
<td>27.08</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 27, subfamily b, polypeptide 1</td>
<td>Cyp27b1</td>
<td>NM_010009</td>
<td>31.57</td>
<td>+</td>
</tr>
<tr>
<td>Epoxide hydrolase 1, microsomal</td>
<td>Ephi</td>
<td>NM_010145</td>
<td>20.45</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione reductase</td>
<td>Gsr</td>
<td>NM_010344</td>
<td>19.065</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 1 (Ya)</td>
<td>Gsta1</td>
<td>NM_008181</td>
<td>19.125</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 3</td>
<td>Gsta3</td>
<td>NM_010356</td>
<td>25.455</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 4</td>
<td>Gsta4</td>
<td>NM_010357</td>
<td>17.02</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 1</td>
<td>Gstm1</td>
<td>NM_010358</td>
<td>22.18</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 2</td>
<td>Gstm2</td>
<td>NM_008183</td>
<td>23.795</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 3</td>
<td>Gstm3</td>
<td>NM_010359</td>
<td>28.455</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 4</td>
<td>Gstm4</td>
<td>NM_026764</td>
<td>32.675</td>
<td>+</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 5</td>
<td>Gstm5</td>
<td>NM_010360</td>
<td>22.695</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, pi 1</td>
<td>Gstp1</td>
<td>NM_013541</td>
<td>32.91</td>
<td>+</td>
</tr>
<tr>
<td>Glutathione S-transferase, theta 1</td>
<td>Gstt1</td>
<td>NM_008185</td>
<td>23.425</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase zeta 1 (maleylacetoacetate isomerase)</td>
<td>Gstz1</td>
<td>NM_010363</td>
<td>22.915</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 1</td>
<td>Mgst1</td>
<td>NM_019946</td>
<td>23.62</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 2</td>
<td>Mgst2</td>
<td>NM_174995</td>
<td>18.73</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 3</td>
<td>Mgst3</td>
<td>NM_025569</td>
<td>21.96</td>
<td>+++</td>
</tr>
<tr>
<td>N-acetyltransferase 1 (arylamine N-acetyltransferase)</td>
<td>Nat1</td>
<td>NM_008673</td>
<td>30.81</td>
<td>+</td>
</tr>
<tr>
<td>N-acetyltransferase 2 (arylamine N-acetyltransferase)</td>
<td>Nat2</td>
<td>NM_010874</td>
<td>23.355</td>
<td>+++</td>
</tr>
</tbody>
</table>
These data confirm the presence of many key enzymes, especially the alcohol dehydrogenases and the CYP450s, particularly Cyp2e1 which metabolizes TAA and AP, indicating the high probability for biotransformation of the selected chemicals by NMuLi cells and supports the possibility for this cell line to be implicated as model system in future toxicology studies. Supporting this conclusion, Landolph et al. demonstrated that this same cell line (NMuLi) was able to bioactivate chemicals, such as benzo(a)pyrene.\textsuperscript{178}

### 4.3.2 Establishment of LD\textsubscript{50}

For this study, I selected several compounds that are either hepatocarcinogenic, i.e.: thioacetamide (TAA) or hepatotoxic: acetaminophen (AP) and Nodularin (Nod). Since oxidative stress is often linked to both non-genotoxic carcinogenesis and frameshift mutational events, hydrogen peroxide (H\textsubscript{2}O\textsubscript{2}) was included as a control.

Dose response curves were obtained for each compound using CCK-8 reagent, which measures cell viability via mitochondrial dehydrogenase function through the conversion of a tetrazolium salt, WST-8, into a colorimetric formazan dye that can then be quantitated on a plate reader. A representative curve is shown in Figure 4-1. From these, the LD\textsubscript{50}, the dose required to kill 50 percent of the cells was established and is shown in Table 4-2.
Figure 4-1: **Dose response curve for hydrogen peroxide.** The LD\(_{50}\) was measured as the concentration at which 50% of the cells survived. This value was estimated to be 4.2mM.

Table 4-2: **LD\(_{50}\) values for compounds in NMuLi cells.**

<table>
<thead>
<tr>
<th>Compound</th>
<th>LD(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen Peroxide</td>
<td>4.2 mM</td>
</tr>
<tr>
<td>Thioacetamide</td>
<td>150 mM</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>30 mM</td>
</tr>
<tr>
<td>Nodularin</td>
<td>170 μM</td>
</tr>
</tbody>
</table>
4.3.3 Oxidative Stress Induction

Measurement of ROS generation by the compounds of interest was determined by the preincubation of the cells with DCFH-DA. This substrate diffuses into the cells and is cleaved by intracellular esterases to yield the non-fluorescent DCFH. Upon exposure to ROS, DCFH is oxidized to the fluorescent compound DCF, which can then be measured. I chose to examine a range of doses surrounding the LD$_{50}$ values for each of the compounds in order to see if a dose-dependent ROS curves were observable (Figure 4-2, panel A). I observed ROS-generating potential for two of the compounds, H$_2$O$_2$ and Nodularin. H$_2$O$_2$ yielded the anticipated result as it was used as a positive control. Nodularin induction of ROS showed a nice dose-dependent response (Figure 4-2, panel B). From this data, I opted to continue onto the frameshift mutational studies using values approximating the LD$_{50}$s.
Figure 4-2: **Induction of ROS in NMuLi cells.** Panel A: Kinetic curves of ROS generation for each compound dose monitored over time. Treatments performed as follows: Hydrogen Peroxide (H$_2$O$_2$) at 1, 5, 10 and 20 mM; Thioacetamide (TAA) at 50, 100, 200, and 400 mM; Aflatoxin B$_1$ (AB$_1$) at 1, 5, 10, and 20 µM; Acetaminophen (AP) at 5, 10, 20, and 40 mM; and Nodularin (Nod) at 25, 50, 100, and 200 µM. P-values shown as indicated for Nodularin-treated cells relative to cells in media alone. Panel B: Dose response curve for Nodularin measured at three time intervals: 15 min, 45 min, and 2 hours.
4.3.4 Morphology Changes in NMuLi Cell Line

Following stable selection of the NMuLi cells transfected with a mono-nucleotide reporter construct that maintained an in-frame configuration of the codons of the lacZ gene, two unique morphologies became apparent. These morphologies are shown with H+E staining (Figure 4-3, Panel A). The first morphology appears to be more hepatocyte-like and mimics the morphology of the parental cell line. The growth patterns of these cells are also similar to the parental cell line in that they grow in tight clusters. On the contrary, the second morphology detected appears to be more fibroblast-like. These cells have larger nuclei and elongated cytoplasm and their growth patterns are scattered in between the clusters of growth from cells exhibiting the alternate morphology.
Figure 4-3: **Distinct morphologies observed from stably-selected NMuLi cells transfected with pαβE-MIN**. Cells displaying distinct morphologies are depicted stained for morphology (H+E), β-gal activity (positive = blue cells), and albumin expression. Panel A: H+E (20x), Panel B: β-gal positive cells (blue) in mixed population, Panels C and D: individual clones showing β-gal activity in clustered hepatocyte-like cells (C) and no activity in diffusely growing cells (D). Panels E and F: Albumin IHC results in NMuLi cells (E) and mixed population (F). In Panels A, B, and F (mixed populations), arrowheads indicate cells displaying hepatocyte-like clustered growth and arrows show cells of alternate morphology displaying diffuse growth. (Panels B-F are shown at 10x magnification.)
Following staining for β-gal enzyme activity using X-gal as a substrate, the only cells that displayed β-gal activity, evident by blue cells, were those cells that appeared as more hepatocyte-like (Figure 4-3, Panel B). No blue cells were detectable from any of the cells that had a more diffuse growth pattern. These cells were then sorted individually and clones displayed both morphologies. Staining the clones for β-gal activity further supported this finding that blue cells were only evident in the hepatocyte-like cells exhibiting clustered growth patterns (Figure 4-3, Panel C) and not in the cells displaying more diffuse pattern of growth (Figure 4-3, Panel D). Further investigations into the morphologic change observed utilized immunohistochemical staining for albumin given that albumin expression is a key marker for differentiated hepatocytes. The results show clear expression differences between these cell types. Similarly to the parental (NMuLi) cell line (Figure 4-3, Panel E), the cells exhibiting more clustered growth (Figure 4-3, Panel F-arrowhead) showed positive staining for albumin whereas the more diffuse-growing cells (Figure 4-3, Panel F-arrow) do not. It is clear from this data that this confirms the specificity of these constructs. More so, that the constructs utilizing a “liver-specific” chimeric enhancer/promoter is functional in hepatocytes or hepatocyte-like cells and thus would only be actively transcribed in the liver.

The observation of two different morphologies emanating from the parental NMuLi cell line is not altogether surprising, as it has been shown that these cells are capable of differentiating into cells with alternate morphology. For instance, treatment with hepatocyte growth factor (HGF) invokes a morphological change and causes the cells to alter their morphology and form tubule structures. This
differentiation is reminiscent of that arising from a population of oval cells first described in models of liver regeneration following chemical carcinogen treatment. Oval cells, often thought to be bipotent or even multipotent cells originating from hepatoblasts, are capable of differentiating into both hepatocytes as well as biliary epithelial cells. For reviews further detailing the nature and characteristics of these unique cells, please refer to Alison, M et al., 1997 and Fausto, N and Campbell, J, 2003.

Another possibility remains to explain the morphological change observed. This is the process known as EMT (epithelial-mesenchymal transition) in which epithelial cells undergo both a morphologic and functional change toward characteristics of cell types resulting from mesenchymal origin. It should be noted, this process is distinct from the process of de-differentiation observed in epithelial cells, which is characterized solely by the loss of epithelial-specific markers (such as cytokeratins, like CK18).

EMT was first observed as a spontaneous event in primary cultures of fetal/neonatal rat hepatocytes and has only recently been observed in both primary cultures of adult mouse hepatocytes as well as in the mouse hepatocyte cell line, AML12, following treatment with TGFβ1. EMT is a normal process present in eukaryotes and plays a role in fetal development, but during adult years, this process is not observed. EMT is thought, however, to play a role in the process of tumorigenesis, particularly involved in invasion and migration of cells considering the tendency for cells of mesenchymal origin to migrate. Adding to this idea is the fact that EMT has been linked to tumor progression and of particular interest to these studies, to HCC.
Given that this cell line was derived from weanling (not fetal) hepatocytes, was isolated from normal tissue (not of tumor-origin), is non-tumorigenic (tumors result only upon injection into isogenic NAMRU mice), and exhibits EMT without the addition of any growth factors, this cell line may be of great benefit in furthering the study of the EMT process; more specifically, how it relates to tumorigenesis.

### 4.3.5 Frameshift Detection

Preliminary studies used the microsatellite-reporters (pIβE-ntR\textsubscript{OUT}), driven by a CMV promoter, which were transiently-transfected into HEK-293T cells then subsequently treated with hydrogen peroxide, confirmed the anticipated result. Briefly, cells were treated with 20 mM \( \text{H}_2\text{O}_2 \) and allowed to grow and divide for 0, 24, 72, and 168 hours. At each of these time points, the cells were fixed and stained using X-gal as a substrate. Plates were then monitored via microscopy and blue cells, indicating that frameshift mutations had occurred, were observed. From transfections with the di-nucleotide repeat vector (pIβE-D\textsubscript{OUT}), the results did not show a clear induction of mutations (data not shown). However, using the mono-nucleotide repeat constructs (pIβE-M\textsubscript{IN}) I observed blue cells from all time points (Figure 4-4). Interestingly, these blue cells exhibited what appeared to be clonal growth over time. This indicates that the frameshift mutations induced were heritable during cell divisions, although the exact mechanism(s) are still to be determined. The frameshift observed with the mono-nt vector was expected. Since the repeat is a \((G)_n\) and shows frameshift
mutations, as supported by other studies\textsuperscript{30, 177} that show frameshift mutations in runs of guanines, caused by carcinogen treatment, shift in the -1 direction, it is likely that there is a single G lost. It is however possible that the restoration of β-gal protein observed is caused by the addition of two G nucleotides into the repeat sequence. Further experimentation and DNA sequencing will be necessary to distinguish the mutational specifics of this restoration.

Once frameshift mutation detection was established it became clear that a better way to detect rare mutational events in large numbers would be by using fluorescence-based flow cytometric methods. Upon mild fixation and permeabilization, GFP fluorescence is retained inside the cell, however in order
to detect the presence of β-gal protein, a detection method was needed. Attempts were made using directly-converted β-gal substrates (Table 4-3) that had fluorescent properties and thus could be detected by flow cytometry.\textsuperscript{71, 227, 304}

<table>
<thead>
<tr>
<th>Substrate (Common Name)</th>
<th>Chemical Name</th>
<th>Vendor</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ImaGene Red\textsuperscript{TM}</td>
<td>C\textsubscript{12}RG (dodecylresorufin β-D-galactopyranoside)</td>
<td>Invitrogen / Molecular Probes (Cat # I-2906)</td>
<td>Tung, C.H. \textit{et al.}\textsuperscript{304}</td>
</tr>
<tr>
<td>DDAO galactoside</td>
<td>9H-(1,3-dichloro-9,9-dimethylacridin-2-one-7-yl) β-D-galactopyranoside</td>
<td>Invitrogen (Cat # D-6488)</td>
<td>Corey, P. \textit{et al.}\textsuperscript{71}</td>
</tr>
<tr>
<td>ImaGene Green\textsuperscript{TM}</td>
<td>C\textsubscript{12}FDG (fluorescein di-β-D-galactopyranoside)</td>
<td>Invitrogen / Molecular Probes (Cat # I-2904)</td>
<td>Nolan, G.P. \textit{et al.}\textsuperscript{227}</td>
</tr>
</tbody>
</table>

In my hands, these substrates were not effectively able to distinguish between β–gal positive and negative transiently-transfected cells and were eliminated from further use in favor of immunofluorescent staining.

Intracellular staining of β-gal protein with fluorescently-tagged antibodies was confirmed. The anti-β-gal antibody was able to specifically recognize β-gal positive cells (Figure 4-5), as shown here with a biotinylated secondary antibody/streptavidin-TRITC-labeled detection system.
After confirmation of primary antibody specificity, the secondary antibody was subsequently changed to a direct-labeled R-phycoerythrin conjugate due to the lack of detection of TRITC emission wavelengths on the FACSCalibur.

4.3.6 Frameshift Detection in pαβE-ntR Cell Lines

Cell lines bearing microsatellite sequences shifting the lacZ gene into an out-of-frame configuration were monitored for ROS generation during treatment with compounds, or equivalent-volume of vehicle, for 2 hours. Following treatment, cells were incubated in fresh growth media for 72 hours at which time they were divided. One half of the treated cells were re-plated to continue cell growth for another 48 hours. At the end of the two time points, 72 (Day 3) and 120 (Day 5) hours the cells were stained with an anti-β-gal antibody then detected by an RPE-conjugated secondary antibody. Frameshift mutations occurring within the microsatellites shifted β-gal back to an in-frame

Figure 4-5: Anti-β-gal antibody validation.
configuration, thereby causing protein expression, and were detectable as red-fluorescently-labeled cells via flow cytometric analysis.

The percent of cells containing frameshift mutations was obtained by the ratio of β-gal positive, red-fluorescent cells to GFP positive cells. Since the cell lines used for these studies were enriched by selection based upon GFP expression, each was estimated to contain only GFP-positive cells. Therefore, this ratio was obtained by dividing the number of β-gal positive cells by the total number of cells measured. The data from each independently-transfected cell line was pooled and the percent mutation was obtained for each microsatellite repeat sequence (Figure 4-6).

![Figure 4-6: Frameshift mutations in NMuLi cell lines expressing pαβE-ntR_OUT.](image)

The combined data from two independent cell lines containing the same microsatellite repeat sequence show increases in frameshift mutations. These increases evident in both types of microsatellites are not directly correlated to the generation of ROS measured during treatment, as hydrogen peroxide treatment
generated the highest level of ROS and would then be expected to show the highest ratio of mutations. The results also confirm that the resultant frameshifts are specific to the compound treatment, as the vehicle controls did not show a similar increase. The higher incidence of mutational events in the mono-nucleotide repeat is not unexpected as a single nucleotide repeat has been found to mutate at a higher frequency than in a di-nucleotide repeat. A decrease in the percent of mutation was observed for most of the samples on Day 5 as compared with Day 3. This may be explained by the return of cell homeostasis of the cell population following removal of a chemical insult. This return is attributed to the restored functioning of intracellular repair mechanisms as well as cell death of mutant cells. Taken in accordance with NMuLi cells being non-transformed and containing no viral entities, this finding is not surprising.

Figure 4-7 further examines this mutational specificity by showing the treatment response shown as fold change as the ratios of treatments to their respective vehicles. From this data there seems to be no strong correlation between ROS generation and number of mutations observed. While the detection of mutants in the nodularin-treated cells containing the mono-nt repeat may in fact, be a result of direct oxidative DNA damage occurring from its metabolism and ROS generation, a discrepancy exists. This discrepancy can be observed by comparing the hydrogen peroxide (H₂O₂) treatment with nodularin (Nod) treatment. In these experiments, hydrogen peroxide was used as a mutagenic positive control for direct DNA damage by adduct formation from ROS; therefore, I expected that cells treated with H₂O₂ would show a higher
number of mutants than those from Nod treatments based upon their respective levels of ROS generation.

From the data, it would be also be expected that if a direct correlation between ROS and mutation number could be seen, based upon direct DNA damage, that the mono-nt reporter would show an increased number of mutations over the di-nt, especially evident between both types of \( \text{H}_2\text{O}_2 \)-treated cells. This idea is supported by the fact that guanines are the prime targets for DNA oxidative damage\(^68\) in addition, to the fact that mono-nt mutations occur at higher frequencies than di-nt.\(^45\) Yamada \textit{et al.}, however, showed an increased frequency of mutations following hydrogen peroxide treatment in a cell line containing a di-nucleotide repeat of (CA\(_{17}\)) greater than that observed for any other repeat type.\(^332\) Due to the lack of mutational variation observed within these sequences while comparing treated to untreated cells, the authors speculated that this increase might be due to the mechanisms underlying

Figure 4-7: Response of treatments in NMuLi cell lines expressing p\(\alpha\beta\text{R}\text{-nt}^{\text{OUT}}\). Bars depict fold changes as ratios of treatments to their respective vehicles.
spontaneous mutation. While spontaneous mutation may also be the mechanism underlying what is observed in this data it is unlikely given the lack of a similar mutation ratio in cell treated with vehicle. It is also possible that the mutations observed result from mechanisms unrelated to direct DNA damage by adduct formation, this will however, only be apparent upon DNA sequencing results.

Comparisons between of the mutant ratio obtained from Nodularin (Nod) treatment to that of H$_2$O$_2$ treatment, in the cells containing the mono-nucleotide repeat, it’s clearly evident that the mutations are not reflective of their ROS-generating ability. It is entirely possible that this treatment could be acting via another mechanism to cause the mutations observed. Nodularin’s primary intracellular change is that it inactivates protein phosphatases (PP) 1 and 2A.$^{230}$ Nod also causes an increased activation of nuclear factor kappa B (NFkB),$^{289}$ possibly through hyperphosphorylation (resulting from inactivation of the PPs), rapid ubiquitination and degradation of the family of inhibitor κB (IκB) proteins, which regulate NFκB.$^{154}$ NFκB upregulation also increases the expression of tumor necrosis factor alpha (TNFα). TNFα, which is also observed having increased expression following Nod treatment,$^{289}$ is a pro-inflammatory protein. Interestingly, it has been shown to cause a decrease in the levels of activity of superoxide dismutase (SOD) enzymes by reducing the transcription of these genes.$^{322}$ All of these actions, ROS generation by Nod, the potential increase in inflammation caused by increased TNFα, increases in NFκB, and the decrease of the anti-oxidant levels (particularly that of SOD), may all contribute to the increased mutation ratio observed from treatments with nodularin as compared to H$_2$O$_2$. 

Of particular interest in this data is the large number of mutations (seen on Day 5) caused by treatment with acetaminophen (AP) observed in the cells carrying the di-nucleotide reporter construct. This finding is especially surprising considering the lack of ROS generation from this treatment. This seems to indicate that the underlying mechanism(s) creating these mutations is(are) distinct from those behind mutations in H$_2$O$_2$ and Nod treatments. The equal number of mutations observable between the di-nt repeats and the mono-nt on Day 3, may indicate that there is similar mechanism of mutation on both types of repeat sequences and could be the result of AP-induced DNA strand breaks$^{200}$, elevated calcium (Ca$^{2+}$) levels$^{138, 249}$, which may or may not invoke mitochondrial damage$^{114, 173}$, or even impaired redox balance of glutathione (GSH)$^{43}$.

What is more intriguing is the increase in mutations observed on Day 5 was only seen in the di-nt containing cells. Since the treatment times for both stable cell lines were identical, and the compounds were removed after 2 hours, it is unlikely that this increase is due to additional direct mutations within these sequences caused by the treatment. Instead, two possible explanations may account for this result. The first may be that there is an increased growth rate of these mutated cells. As it is known that AP induces inflammation$^{146}$, the result observed could be generated by either, increased growth, upregulation of inflammatory cytokines, such as interleukin-6 (IL-6) which causes the dysregulation of GSH balance$^{43}$, or combination of both. If this is the case, the increased mutations would most likely be observed in both repeat types. A second possibility to explain this increase is that the cells containing mutations in di-nt repeats conferred some other sort of selective advantage, possibly the result
of a different mechanism resulting from the treatment, allowing for these cells to continue dividing. This selective advantage may occur by the blocking of apoptotic pathways\textsuperscript{5,302} or alternatively may be a growth advantage, resulting from the activation of oncogenic or signaling pathways, like those observed in some patients diagnosed with hepatocellular carcinoma (HCC)\textsuperscript{51,204}.

Several genes that contain dinucleotide repeat sequences in their promoter/intronic regions are important in growth regulation.\textsuperscript{109,163,276,318} In particular, genes that contain (CA)\textsubscript{n} repeats, are the epidermal growth factor receptor (EGFR)\textsuperscript{112}, matrix metalloproteinase 9\textsuperscript{276}, and the gene encoding cysteine-rich heparin binding protein, Cyr61\textsuperscript{318}. In support of the idea that growth regulation is in some way disrupted in these cells showing mutations thus allowing for additional mutations to be observed on day 5, each of these 3 genes have been demonstrated to have altered expression levels in HCC.\textsuperscript{48,103,151} Specifically, that alterations in the lengths of the (CA)\textsubscript{n} repeats contained within each of their sequences, affects their expression levels.\textsuperscript{269} Given this correlation, it would be intriguing to study the expression of each of these genes in the context of the mutated cells.

4.3.7 Closer Examination of Flow Cytometry Results

An interesting “peak” was observed in the scatter-plots from certain treatment groups corresponding to those treatments showing ROS generation during the DCFH-DA assay. Each peak indicated an increase in intensity for both red and green fluorescence. Figure 4-8 depicts a representation of this described
peak shown for Nodularin-treated cells (panel A) along with its vehicle, DMSO (panel B). The peak of dual fluorescence is not observed in DMSO-treated cells.

The increase in GFP fluorescence observed is intriguing. The GFP fluorophore is photoactivated to fluoresce through oxidation. In this sense GFP fluorescence can actually release oxidative radicals and stimulate oxidative stress. Due to the limited amount of time the cells were exposed to light of the correct wavelength to activate GFP, I do not believe this increased the frameshift events. In a similar manner, the GFP fluorophore has been shown to act in an antioxidant capacity accepting free oxygen radicals. This action however, is apparent as a slight decrease in fluorescence not as the increase observed.

Figure 4-8: Representative double-fluorescence peak. Peak denoted by R2 gate (indicated by arrows). Panel A: Nodularin-treated cells; Panel B: DMSO-treated cells.
An increase in GFP fluorescence not associated with induction of GFP protein levels has recently been described by Dabrowska et. al. from colorectal adenoma cells treated with methotrexate.\textsuperscript{77} This finding is interesting considering methotrexate has been shown to induce oxidative stress in several cell lines.\textsuperscript{17,53}

In light of these studies, I was interested to see how the mutational percentages correlated with the increase in GFP fluorescence. The percentages were calculated from the ratio of cells gated, as indicated in Figure 4-8 by arrows, relative to the total number of cells analyzed. These results show a similar pattern of microsatellite frameshift events among all of the treatment groups; however, an altered trend within each treatment from that seen in the previous analysis was observed. These data show a higher percentage of frameshift mutations on Day 5 than were detected on Day 3 (Figure 4-9).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4-9.png}
\caption{Frameshift mutations from EGFP-fluorescence-enhanced cells.}
\end{figure}
Once again the mutants from treatments were compared to their respective vehicle controls to obtain the treatment response ratio (Figure 4-9).

![Graph](image)

**Figure 4-10:** Response of treatments in EGFP-fluorescence-enhanced NMuLi cell lines expressing pαβE-ntR\textsubscript{OUT}. Bars depict fold changes as ratios of treatments to their respective vehicles.

In adherence with the original frameshift percentages, these data do not show a correlation between ROS generation and frameshift mutation especially evident by the relative comparison of responses between hydrogen peroxide and nodularin treatments. I would again expect that the number of mutations detected in the hydrogen peroxide-treated cells to be higher than the nodularin-treated, due to their respective ROS generation curves (see Figure 4-2, panel A). Additionally, this correlation in terms of direct DNA damage by adduct formation would more evident in cells containing the mono-nt repeats than the di-nt.
The data also demonstrates that this isolated population of cells exhibiting higher GFP fluorescence, corresponds well with the treatment results from the entire population, especially in the cells treated with nodularin. More so, that this smaller population represents the cells that are the direct result of these mutational mechanisms. It would be interesting to examine these cells and further characterize them as they specifically relate to ROS generation and the mutational response observed within the entire population.

4.4 Conclusion

In this chapter I have described a novel bicistronic-reporter gene system that is able to detect frameshift mutations in microsatellite sequences. Future studies will be necessary to determine the exact nature of these mutations given the lack of association between oxidative stress and mutational events. These studies would involve sorting frameshifted-cells based upon their red fluorescence, isolating DNA, and proceeding with direct sequencing or PCR-based microsatellite frameshift analysis in order to detect microsatellite sequence differences. If DNA alterations were not observed (considered unlikely), the frameshifts observed may alternatively be a result of errors during transcription or translation. mRNA analysis would be necessary to examine if the frameshift was included directly in the sequence. Again, this could be ascertained by cDNA synthesis followed by direct sequencing or PCR analysis.

If no sequence variations are apparent, differences in CpG methylation status of the mRNA may be responsible and could be assessed by methylation-
specific PCR. Regardless of the mutational events, I have described a model system that could provide the basis for future studies into epigenetic regulation by examining cells that have altered expression patterns. In addition, these future studies could be supported by the incorporation of a fluorescent β-gal substrate that would maintain cell viability. In doing so, cells showing early frameshift mutations, that may or may not be corrected by mechanisms such as mRNA turnover and restoration of normal translation, could be followed over time.

Based upon the mutations detected, I have also presented possible avenues for expanding these studies from ROS generation into other areas of inflammatory and cell signaling pathway response to these treatments.

Alternatively, I have shown here that cells with different morphologies and different expressions, as preliminarily characterized by albumin expression, may arise from the parental NMuLi cell line in the absence of external growth factors like TGFβ1. These changes provide evidence that this cell line would a strong candidate for the study of EMT and its potential impact on tumorigenesis.
Chapter 5

Development of Transgenic Mouse Model for Determining
In Vivo Effects of Hepatocarcinogen Treatment

5.1 Introduction

Determining the carcinogenic potential for a particular compound in humans is one of the key considerations for successful drug development. Rodent long-term (2-year) bioassays are currently used for this determination and also influence regulatory policy; however, these have several disadvantages, namely, high-cost, lengthy time until analysis, as well as the non-practicality assumed considering the acquisition of data from large numbers of synthetic, as well as natural compounds.

Assays intended to study carcinogenic potential in shorter-term, present additional complications for analysis. For instance, short-term carcinogenic studies require treating mice with chemicals at the maximum tolerated dose (MTD) and provide little translational information regarding low dose exposure in humans.

While providing a link between compound exposure and human carcinogenesis, these studies provide no information regarding the mechanisms underlying tumor formation. Elucidating these mechanisms would not only assist future drug development, but also provide essential understanding as to the etiology of cancer.
Genotoxic agents, chemicals that directly bind DNA thereby causing mutations, show a strong link to carcinogenesis and the process has been studied mechanistically. This however, leaves an informational gap considering the large number of chemicals that are not genotoxic as determined by the Ames, in vitro micronucleus, or mouse lymphoma assays.\textsuperscript{15, 283} Gene expression profiles from the livers of rats treated with certain non-genotoxic chemicals show responses that correlate with the induction of oxidative stress as well as genes inducing DNA replication and cellular proliferation.\textsuperscript{93}

Genomic instability has been linked to the development of cancer.\textsuperscript{59, 88} In cells that have genomic instability, regions of DNA known as microsatellites, are particularly prone to frameshift mutations. DNA replication and errors during DNA repair and have been implicated in numerous cancers involving microsatellites.\textsuperscript{167, 174}

Compounds that cause frameshift mutations may act in a variety of ways. First, they may be direct mutagens whereby they, or their metabolized product, bind covalently to DNA and create a strand-slipage effect during DNA repair. Alternatively, as may be the case in carcinogenesis caused by non-genotoxins, the compound may induce changes in other key processes that have the ability to alter gene expression. This may occur by causing errors in the processing or regulation of RNA transcripts.

Early studies in this laboratory indicated that treating rats with low doses of non-genotoxic carcinogens showed the development of enzyme-altered foci in livers. The cells found within these foci showed altered intracellular localization of RNA, which contained repetitive sequence. I hypothesize that it is the altered
metabolism of these repetitive RNA molecules that alter cellular processes, which may be involved in hepatic carcinogenesis.

Detection of early genomic instability presents a challenge since by the time chromosomal alterations are detectable the cells have already undergone such significant chromosomal change that elucidating any mechanism(s) that are responsible becomes impractical. Here I present the development of a transgenic mouse line that was designed to detect this early genomic instability using a frameshift-prone microsatellite sequence within a reporter gene to locate these cells altered by non-genotoxic treatment. Other mouse lines aimed at detecting frameshifts in microsatellite regions will also be discussed.

5.2 Materials and Methods

5.2.1 Transgene Purification

10 µg pαβE-M\textsubscript{OUT} plasmid DNA was digested with AseI and AflII restriction enzymes (New England Biolabs, Ipswich, MA) to obtain a linearized fragment of DNA. The digested DNA was run on a 1.5% agarose gel and band of correct size (5774 bp), corresponding to the size of the desired transgene, was gel extracted using the QIAEX II Agarose Gel Extraction Kit (Qiagen Inc., Valencia, CA). In order to remove any other contaminants that might interfere with the success of the injections, the transgene was phenol:chloroform extracted, followed by ethanol precipitation and resuspended in 50 µL TE buffer.
5.2.2 Creation of EαβM_{OUT} Mouse Line

Purified, linearized transgene DNA, EαβM_{OUT} was injected at a concentration of 19.6 µg/µL, via pronuclear injection into B6D2 F2 hybrid mouse embryos. Weaned pups were screened using DNA isolated from tail biopsies and PCR analysis using transgene-specific primers. Transgene-specific primers used:
Forward: 5'-ATCGAAAGAGCTGCTAAAG-3', Reverse: 5'-CAGCCAGCTTTCCGGAACCGC-3'

B6D2 F2 female founder was mated with a wildtype C57Bl/6 male mouse. The EαβEM_{OUT} mouse line is maintained as hemizygous, created by breeding a transgenic mouse with a wild type C57Bl/6.

5.2.3 Mouse Characterization

5 week old mice used for characterization were sacrificed and select organs were dissected. Tissue was snap frozen in liquid nitrogen before being further processed for DNA, RNA, or protein analysis.

5.2.3.1 DNA Analysis

Mouse tissues 3 mm³ in size were added to AquaGenomic solution (MultiTarget Pharmaceuticals, Salt Lake City, UT) and homogenized in a microfuge tube using a teflon microtube pestle. Isolated DNA was resuspended in 20 mM Tris-HCl and 200 ng used in PCR utilizing the same primers used for the detection of the founder mouse.
5.2.3.2 RNA Analysis

RNA was isolated from 3-5 mm³ pieces of mouse brain, kidney, lung, spleen, heart and liver tissues. Briefly, each tissue was processed independently by snap freezing the tissue in liquid nitrogen, wrapped in foil sachets and disrupted by impact with a hammer. The resulting powder was placed in a 1.5 mL Eppendorf tube containing 600 µL buffer RLT and 6 µL β-mercaptoethanol (β-ME). Contents were then placed in a QIAshredder, which was subsequently centrifuged in order to homogenize the tissues. Following homogenization, the RNA was extracted using the RNeasy Mini Kit (Qiagen, Valencia, CA) including the optional on-column DNase digestion step.

Primers and probes for analysis via quantitative RT-PCR (RT-QPCR) were designed using Primer Express software (Applied Biosystems Inc, Foster City, CA). Each RNA analyzed was done so utilizing 3 primer/probe sets. The first set was designed to a region in the early, coding sequence of the lacZ gene. These sequences are: F: 5’-TCCTGAGGCGGACTGTCG-3’; R: 5’-GCGGATTGACCGTAATGGG-3’; Corresponding probe: 5’d FAM-TCCCCTCAAACGTGCAATGG-3’. The second set was designed to a region of the IRES sequence. The sequences of this set are: F: 5’-GGCTCTCCTCAAGCGTATTTCA-3’; R: 5’-ACATGTAAGCATGTGCACCG-3’; Corresponding probe 5’d Quasar 670-AAGGATGCCCAGAAGGTACCCCATTG-3’. Finally, the third primer/probe set was used as a RNA quality control and was designed to a region of the mouse TATA-box binding protein (TBP). These sequences are: F: 5’-ACGGACAACTGCCTGTTGATT-3’; R: 5’-
ACTTAGCTGGGAAGCCCAAC-3'; Corresponding probe 15 5’ T JOE d-TGTGCACAGGAGCCAAGAGTGAAGA-BHQ-1-3’.

RT-QPCR was performed using the QuantiTect Probe RT-PCR kit (Qiagen). This kit also includes a HEX dye within the 2x Master Mix which controls for consistency of readings from wells across the entire plate. 100 ng of each RNA sample was run in triplicate and all 3 primer/probe sets were multiplexed within the tubes. Along with a no template control, each tissue was run in a well with master mix that contained no reverse transcriptase (RT). Positive RNA controls were obtained from RNeasy Mini Kit preparations of NMuLi cells transiently-transfected with either $P_{\text{CMV}}$-driven $P_{\beta_\text{E-M} \text{OUT}}$ or $E_{\text{ALB}}/P_{\alpha_1 \text{AT}}$-driven $P_{\alpha_\beta_\text{E-M} \text{OUT}}$. Relative expression levels were obtained by the standard curve method. This method results in ng values for each sample based upon the Ct values compared to the Ct values resulting from a standard curve consisting of serial dilutions of known template amounts.

5.2.3.3 Protein Analysis

Isolated tissues were lysed in Igepal Lysis Buffer containing Mammalian Protease Inhibitor Cocktail (Sigma, St Louis, MO) and homogenized using Tissue Tearor. 20μg per lane of total protein was run under reducing conditions on NuPage 4-12% Bis-Tris gel (Invitrogen, Carlsbad, CA) with 250ng of recombinant GFP protein (Clontech, Mountain View, CA) run as a positive control. Protein was then transferred to PVDF membrane and blocked with 5% milk/TBS-T. The primary antibody, Living Colors® A.v. Peptide Antibody (Clontech), was used at
a 1:200 dilution followed by a 1:2000 dilution of the secondary antibody (Anti-rabbit-HRP, Cell Signaling Technologies, Danvers, MA). Antibodies were detected using the West Femto Kit (Pierce, Rockford, IL).

Cryostat frozen sections were cut and examined by an Olympus microscope for the presence of GFP under the FITC fluorescent filter. Sections were also observed using a TRITC fluorescent filter to detect possible autofluorescence or sectioning artifact.

5.2.4 Mutation Detection

A single gavage dose of 1.8mg/kg Aflatoxin B₁ (Cat #A6636, Sigma-Aldrich, St. Louis, MO) representing 20% of the LD₅₀ or volume matched DMSO (control), in 0.2cc canola oil, was administered to three mice (30g each). 48 hours after carcinogen treatment, the acute phase response was induced in each of the mice by subcutaneous injections in both hind legs with 0.1cc turpentine. Mice were sacrificed 72 hours after induction of the acute phase response, livers isolated, and snap frozen prior to cryosectioning. Frozen liver sections were fixed with 0.5% glutaraldehyde in 1x PBS for 10 minutes, rinsed, and stained overnight at 37°C in a humidity chamber with a 1x PBS solution containing 50 mM K₃Fe(CN)₆ and 50 mM K₄Fe(CN)₆ and X-gal (Invitrogen). The following day, sections were rinsed and counterstained with Nuclear Fast Red and coverslipped. Sections were examined on an Olympus microscope for the presence of blue cells indicating a frameshift mutation had occurred.
5.3 Results

5.3.1 Design of Transgene

A bicistronic reporter transgene (Figure 3-10) was developed incorporating a frame-shifted version of the bacterial beta-galactosidase (β-gal) gene as well as the enhanced Green Fluorescent Protein (EGFP) gene from *Aequorea victoria* (*A. victoria*) following an Internal Ribosome Entry Site (IRES). The frameshift in the β-gal gene was created due to the incorporation of a single nucleotide microsatellite region \((G)_{10}\) within a stretch of 3 guanine residues, which yielded a mononucleotide repeat sequence of \((G)_{13}\). When incorporated, this sequence shifted the β-gal gene coding sequence, rendering it out-of-frame. By the incorporation of these repeats, no β-gal protein could be produced due to the presence of a premature stop codon. The lack of protein produced from this construct was confirmed *in vitro* by transiently-transfecting the construct into mouse liver cells (NMuLi). Using a similar DNA sequence including a repeat that maintained the codon integrity, β-gal protein could be detected. In order for the expression of β-gal protein, the microsatellite would need to lose a repeat unit, in this case a single G) or gain 2 repeat units 2 Gs, which would create a frameshift back to an in-frame configuration. These changes would restore the reading frame, and result in the translation of the β-gal protein.
To target the expression of transgene to the liver, I included a liver-specific chimeric enhancer/promoter, \( E_{ALB}/P_{\alpha 1AT} \) (courtesy of Dra. Gabriela Kramer, Universidad de Navarra, Pamplona, Spain). \( E_{ALB}/P_{\alpha 1AT} \) combines a mouse albumin enhancer \( (E_{ALB}) \) linked with the human alpha-1-antitrypsin promoter \( (P_{\alpha 1AT}) \) and has been shown to not only direct liver-specific expression of a reporter gene, but also maintains it high level of function over time, as compared with a viral-based enhancer-promoter, following hydrodynamic tail-vein injection.

### 5.3.2 Identification of Founder Mouse

Mouse pups resulting from pronuclear injection into B6D2 F2 hybrid mouse embryos, were tail clipped and screened using PCR. One female mouse was identified as a founder by a band ~340 bp in size (Figure 5-2). This female became the founder and was mated to a wild-type (WT) C57Bl/6 male.
I selected the C57Bl/6 mouse background due to its low propensity for spontaneous tumors. This strain is also one of the selected backgrounds chosen by the NTP for use in their long-term bioassays and is also the strain used in the creation of the BigBlue® Mouse. This similarity provides the opportunity for direct comparisons among these tests. The possibility for severe, sometimes lethal, genome alteration due to the removal or disruption of key developmental or regulatory genes was avoided by maintaining a hemizygous transgenic line, thereby retaining the expression of one normal allele.

Figure 5-2: PCR screening analysis for founder pup identification. Positive mouse DNA, indicated by arrow, shows band at ~340 bp. Lane 1: DNA ladder, Lane 2: WT DNA (negative control); Lanes 3-8: Tail clip DNA from weanlings; Lane 9: empty; Lane 10: Transgene spiked into WT gmDNA (positive control)
5.3.3 Characterization of EαβM\textsubscript{OUT} Mouse Line

5.3.3.1 Presence of Transgene DNA

High copy numbers of DNA transgene provide increased opportunity for the microsatellite sequence to be affected, thereby increasing the likelihood that a frameshifted protein will be produced. The liver-targeted transgene, which contains a non-functional, frameshifted version of the lacZ gene followed by a functional GFP sequence, is detectable during PCR analysis. Each tissue contains more than 100 copies of the transgene. (Figure 5-3) Due to the abundance of the transgene, future studies concerning rare mutational (frameshift-inducing) events are supported.

Figure 5-3: PCR-based copy number analysis of EαβM\textsubscript{OUT} DNA. Positive mouse DNA shows band at ~340 bp. Lanes 1 and 13: 1 kb+ DNA ladder (Gibco), Lane 2: WT DNA (negative control); Lanes 3-5: WT DNA spiked with 1, 10, and 100 copies of transgene DNA respectively; Lanes 6-11: Tissue DNA; Lane 12: No template (PCR control)
5.3.3.2 Presence of GFP Protein

The presence of GFP protein was assessed by detection via western blot (Figure 5-4). While a number of non-specific bands were observed in these mouse tissues, bands at 27 kDa were evident in liver samples from two different mice indicating the presence of GFP. A similar-sized band was also identified in the protein isolated from the heart, but was absent from other tissues such as brain, kidney lung, etc.

![Figure 5-4: GFP protein expression in EαβM\_OUT tissues. Arrows indicate liver bands of correct size. Lanes 1 and 9: MagicMark protein marker (Invitrogen), Lanes 2-8: tissue samples; Lane 10: recominant GFP protein (250 ng)](image)

Upon microscopic analysis of frozen sections, a few GFP positive cells were detectable. (Figure 5-5) Panel A shows an overlay of FITC and TRITC images. Red fluorescence is usually indicative of autofluorescence; however, the intensity of this fluorescence indicates that it is actually due to sectioning artifact.
Increasing the exposure time can detect a background level of red fluorescence although the signal is too low to be a contributing factor to the red shown. The GFP signal in this section was very bright and showed no corresponding location of red fluorescence under the TRITC filter, implying that this signal is indeed GFP fluorescence and not additional artifact. The GFP expression shown in panel A was acquired using a lower exposure time in order to approximate morphology. Panel B combines Panel A and a brightfield image to show location of these cells in the tissue section. Most of the cells detected here are located along the edge of the liver section, however a couple other sections showed an equivalent GFP expression from clusters of cells located within the tissue, although still near the edge (data not shown).

Figure 5-5: **Tissue Expression of GFP-containing cells.** Panel A: Fluorescent overlay of image captured using TRIT-C (red) and FIT-C (green) filters. Green cells indicate GFP protein, as autofluorescence (shown as red) within the tissue would be visible in this overlayed image as yellow. Panel B: The location of GFP-positive cells can be seen in this image the result of overlaying the image in Panel A and a bright field image (shown without contrasting stain).
The limited number of GFP cells observed here may reflect a limited sensitivity due to microscopic detection. It is also possible that the GFP expression, which is dependent on an IRES site, is not translated from all transcripts, substantially lowering the expression of GFP protein and additionally limiting its detection.

5.3.4 Mutation Detection

In initial studies treating EαβM<sub>OUT</sub> mice with a low dose of Aflatoxin B<sub>1</sub>, a single blue cell was detected in the liver of only one mouse, out of 3 individuals treated, by staining with X-gal. (Figure 5-6), whereas no blue cells were detected in the livers of mice treated with DMSO. Thus, if this is a true positive, the reporter may be functional although the low level and expression compromises the utility of the model. This provides good indication that the reporter is functional and able to identify cells that have had mutational events.
5.3.5 Further Characterization of EαβM_OUT Mouse Line

Due to the concerns from the lack of clear data from the previously described experiments, I continued the characterization of the EαβM_OUT mouse line by evaluating RNA from all tissue samples as well as reassessing the presence of GFP protein in the liver tissue.

5.3.5.1 RNA Analysis

Based upon analysis using the standard curve method comparing sample Ct values to those obtained from known amounts of template, RNA transcript levels were calculated. These transcription levels were then normalized to the

Figure 5-6: Frame-shifted β-gal positive cell in liver tissue. Arrow indicates single blue cell identified, the result of staining with X-gal as the substrate for βgal. Tissue was counterstained with Nuclear Fast Red.
transcript level of a housekeeping gene, in this case, TBP. The QPCR performed on cDNA samples reverse transcribed from RNA isolated from mouse tissues showed no evidence of detectable levels of transgene-derived RNA transcripts from either of the two designed transgene-specific primer-probe sets (Figure 5-7, panel B). In contrast, RNA from transfected cell lines showed cycle threshold (Ct) values from the transgene-specific sets that fell within the range of the standard curve (Figure 5-7, panel A), indicating that the primer-probe sets were accurately measuring their specific target. Furthermore, the levels obtained from these two cell lines (used as positive controls) reflected the relative activities of their promoter sequences with the CMV promoter, yielding an approximately 10-fold higher RNA level than that from the mammalian E\textsubscript{ALB}/P\textsubscript{a1AT} promoter.

Among all tissues analyzed, liver RNA had the highest Ct value, thus the expression was one of the lowest. This result could, theoretically, be due to a decreased RNA quality from tissue samples as compared to those that were cell line-derived. However, the corresponding results from the primer/probe set designed to an internal control, the housekeeping gene TBP, indicated that the quality of tissue-derived mRNA is acceptable. The good quality was further confirmed by the detection of clear 18s and 28s bands and little to no smearing, indicative of degradation, seen when 600 ng of the mRNA preps was run on an agarose gel (data not shown).
Figure 5-7: RT-QPCR analysis. Panel A: Sample data points shown against plot of three standard curves representing lacZ (blue), IRES (green), and TBP (red) primer-probe sets; Panel B: Bar graph depicting relative mRNA fold change of tissues and cell lines as compared to pαβE. Two unique primer probe sets to lacZ (blue bars) and IRES region (green bars) were analyzed. p-βE denotes NMuLi cell line transiently-transfected with the transgenes constructed with the CMV promoter (pIβE) or liver specific enhancer/promoter E_{ALB}/P_{αAT} (pαβE).
5.3.5.2 Confirmation of Protein Analysis

Several attempts at detecting GFP protein using multiple antibodies on paraffin-embedded tissues for immunohistochemistry, were made with no success. Returning to western blotting techniques, protein isolations were again run, however, instead of solely running liver tissues against other tissue types, a negative control was included. This negative control protein was isolated from a transgene-negative littermate of the $E\alpha\beta M_{\text{OUT}}$ mouse being characterized. This was done to ensure the consistency of results given that both mice would have the same genetic background and thus the potential for similar non-specific binding. The results indicated that indeed the ~27 kDa band observed in the liver of the transgenic mouse was similar in size and intensity as a band seen from the negative littermate (Figure 5-8).
Among 3 additional antibodies against GFP protein, one other showed a band in the livers at 27 kDa consistent with mature GFP; however, the remaining two did not detect this band. All of the antibodies were able to detect the recombinant EGFP protein run as a positive control.

Figure 5-8: Suspected GFP protein. Arrows indicate liver bands of correct size and confirm the band size observed on previous westerns. Lanes 2, 4 and 6: MagicMark protein marker (Invitrogen), Lane 1: EαβM_OUT liver #2 (as denoted in Figure 5-4); Lane 3: EαβM_OUT liver; Lane 5: negative liver; Lane 7: recombinant GFP protein (250 ng).
5.4 Discussion

The development of mouse models for the study of cancer, have steadily been increasing since its conception in the late 1970’s and early 80’s. Often these mice have altered protein levels for one particular protein and while indicating the importance for that protein in normal cellular development or disease processes, these mice do not aid the identification of any underlying mechanisms that might affect multiple proteins.

Transgenic mice able to detect frameshift mutations in DNA sequences, have been previously created, however the majority of these mouse lines utilize a reporter containing λ-phage incorporated into the genome. While these mice offer a means for analyzing the effects of metabolic conversion following compound treatment, they rely on the isolation of the λ-phage from the original tissue for the detection and sequencing of mutants. Therefore, no data can be obtained from these cells in terms of morphology, chromosome, or enzyme level status and attempts at deciphering mechanisms become unattainable.

In order to simultaneously assess the mutation and morphologic status another transgenic mouse line was created by another group with a (G)$_{11}$-repeat microsatellite. This mouse contains a microsatellite-disrupted human placental alkaline phosphatase (PLAP) gene driven by a human β-actin enhancer/promoter. In order for PLAP to be expressed the microsatellite would need to lose one repeat, a single guanine. The construct they selected has several limitations compared to my model and considering the intended study. First, in cells, they reported a high spontaneous mutation frequency from this
microsatellite.\textsuperscript{52} In mice, this frequency was similarly observed.\textsuperscript{52} For the intended study of chemically-induced hepatocarcinogenesis, I wanted to eliminate most spontaneous mutation and focus on only those cells that are direct targets. Second, the strain of mouse they selected is FVB/N, which has a higher potential for spontaneous mutation than does C57Bl/6. Due to the background strain alone, the mutation rate they observe may be a reflection of this selection. Other factors give my construct an added advantage, namely the presence of a second reporter gene and the inclusion of an IRES sequence. Combining these, serves to control for transcriptional and translational fidelity, respectively within the cell.

Finally, the main advantage of my construct is that it is relatively liver specific. In an additional study with the PLAP mouse line, the lab reported low expression of PLAP, both in levels of transcripts and protein, from mouse livers containing a repeat sequence that was in-frame.\textsuperscript{135}

While the presence of GFP protein, identified initially from protein isolates of the liver supported the fact that my transgene is predominantly liver-specific, subsequent experimentation showed no presence of mRNA derived from the transgene and a similarly-sized band to that of EGFP, was detected in liver protein of a mouse that bore no transgene. The suspected presence of GFP from protein isolated from heart tissue presented some interest. I initially thought this was due to the promoter used in the development of the transgene, the promoter for human $\alpha$-1-antitrypsin. This same promoter has also been identified in macrophages and peripheral blood monocytes.\textsuperscript{62} Supporting this finding, a recent study found blood macrophages were capable of generating hepatocyte–like
cells. Given these two reports, it was entirely likely that blood monocytes were the source for the expression of GFP observed in heart tissue.

It was anticipated that the RNA isolated from the liver tissue would have yielded one of the lowest Ct values equivalent to the high nanogram quantities of RNA transcripts more so than any other tissue. This value however, was not expected to be as high as mRNA levels resulting from that of a transient transfection due to relative DNA copy number. Since QPCR analysis demonstrated no mRNA expression, as well as confirmation of a band detected in a mouse bearing no transgenic DNA, I am now certain that the transgene is not being expressed and that the 27 kDa band identified via western blot, is non-specific and occurring by antibody cross-reactivity with other cellular proteins.

There exist several possibilities as to why the transgene that is integrated into the genome in over 100 copies within each mouse of the line fails to express RNA and protein. Another common explanation is one of viral genome silencing, typically observed in the case of promoters where the viral genome becomes inactive due to methylation on CpG dinucleotides, thus preventing transcriptional factor binding. In this transgene, this is not the case as I replaced the viral promoter sequence with one that was constructed from mammalian sequences.

This transgenic sequence does however carry viral sequence in the form of the IRES. This sequence was isolated from the encephalomyocarditis virus (EMCV). While this sequence has been shown to be effective in vivo by infection into chicken embryos, electrotransfer into mouse muscle, and in a transgenic mouse, the expression of the gene following this sequence is reduced
compared to that of a gene directly following the promoter sequence. In particular, the EMCV IRES sequence shows minimal activity in livers of adult transgenic mice as compared to other tissues. Since these experiments yielded no detectable mRNA levels, the lack of GFP protein cannot be explained by the lack of in vivo activity of the EMCV IRES sequence in livers.

Provided another attempt was made to create a functional mouse line expressing these reporter genes, it would be an important consideration to remove the viral IRES and replace it with one of mammalian origin. To date, several mammalian IRES sequences have been identified and it would be important to select a sequence that shows relatively high activity in adult mouse liver given that some, as previously mentioned, show minimal activity. One strong contender for a mammalian sequence that exhibits IRES activity is the sequence constructed of 10 repeated copies of the 9 nt sequence from the 5' UTR of the Gtx homeodomain RNA. This construct, deemed a “super IRES” (SIRES), shows much higher levels of activity than the EMCV IRES, although its activity in the context of the liver of an adult transgenic mouse has yet to be determined.

Other explanations as to the lack of transcription must be considered. One of these is perhaps that the point of integration for the transgene is in a silenced region of the chromatin structure. Silencing of transgenes occurs around centromeric as well as telomeric regions. The mechanisms of this silencing involve DNA protein binding as well as methylation status. In order to assess the circumstances surrounding the lack of transcription, it would be of interest to determine where the exact point of integration for the transgene is. Experiments
using fluorescence in-situ hybridization (FISH) could be designed to approximate the transgene’s location. Further experimentation would involve sequencing of the adjacent DNA from the junction fragment observed on a Southern blot, to identify the exact location of the transgene. Continuing in this vein, the methylation status or other protein binding resulting in a silenced gene could be determined by isolating primary hepatocytes, culturing them, and treating them with various inhibitors which might include 5-azacytidine (5-azaC), an inhibitor of DNA methyltransferases, or trichostatin A (TSA), a histone deacetylase (HDAC) inhibitor. It would be of particular interest to determine if any of these treatments induced expression of the transgene in primary hepatocyte cultures, determinable by RT-QPCR and GFP fluorescence, the latter indicative of protein expression.

With all of the aforementioned possibilities, one other possibility remains. It has been shown that repeated sequences, such as would be present as the “head to tail” concatemers of transgenic DNA in a mouse, are able to be silenced by inducing a heterochromatic state, involving protein associations, such as HP1, and CpG methylation. Interestingly, this induction and maintenance of this heterochromatin does not involve the Dicer-dependent RNAi pathway in differentiated mammalian cells. If heterochromatin formation is occurring, thereby silencing the transgene, primary hepatocyte cultures from the mice could also be used to study the mechanisms behind its development and maintenance.

In this chapter, I attempted to design a transgenic mouse line that contains a bicistronic reporter transgene. This mouse line would present a greater opportunity for the analysis of early liver carcinogenic events compared to the
PLAP-reporter mouse. Since the EαβM\textsubscript{OUT} mouse line developed contains many copies of the DNA, yet never expressed the reporter transgene, I am unable to utilize this mouse line for the intended studies; however, through the use of primary hepatocyte cultures isolated from these mice, several new options for additional avenues of study have come to light.
6.1 Summary

In this dissertation, I aimed to examine key early events in carcinogenesis. Based upon the observation of repetitive sequences contained in RNA that exhibits altered compartmentalization in hepatic foci, it is speculated that these sequences play a key role in the development of genomic instability. Therefore, it is hypothesized that hepatocarcinogenic compounds, which may or may not be associated with genotoxicity, alter these repetitive sequences and that these sequences are key components for genomic regulation. Furthermore, based upon its association to non-genotoxic carcinogenesis as well as the observation of 4-HNE staining, indicative of oxidative damage, in hepatocytes surrounding morphologically-altered foci, I examined the oxidative stress-inducing potential of the selected compounds.

To address this hypothesis, I developed novel liver-specific, bi-cistronic reporter plasmids containing microsatellite repeat sequences. These constructs are capable of detecting frameshift events from either a \((G)_{13}\) mono-nucleotide repeat or a \((CA)_{18}\) di-nucleotide repeat, through the restored expression of beta-galactosidase protein. A second gene encoding enhanced green fluorescent protein (EGFP) was also included on the same plasmid and transcribed into RNA from the same promoter. EGFP, which can be translated using an internal ribosome entry site, is then able to act as a control indicating the presence of the
plasmid despite the presence or absence of β-gal protein. After preliminary studies determining I could induce frameshift events, I aimed to design two model systems for the expression of these constructs.

The first model system was developed using a model liver cell line. The primary advantage of this approach is ease of experimental manipulation and data interpretation. Considering the lack of other cell types, the data collected from a single cell line can be directly attributed to the response of those cells and is not confounded by the interactions of the other cells that might be found to interact in vivo. On the other hand, while this approach is effective at describing the response of the selected cells, the result may not be indicative of the global effects of the compound given the contextual interaction that exists between cell types within an organ. In order to address the latter, attempts were also made to create an in vivo mouse model.

6.1.1 pαβE-ntR Cell Lines

In the first system, I stably-transfected normal mouse liver cells with each construct which were first sorted based upon the fluorescence and intensity of EFGP so that all cells should have high levels of expression. The cells were then used in a series of experiments to detect oxidative stress induction and frameshift events from a variety of possibly hepatocarcinogenic compounds.

After confirming that the metabolism gene profile of the cells reasonably resembled the profile expected from a mouse liver in vivo, the cells were treated with various compounds and monitored for induction of oxidative stress.
Hydrogen peroxide was used as a control for the ROS induction. A dose-dependent ROS induction was observed with Nodularin, a cyanobacterial toxin that causes liver toxicity. No free radical formation was observed from either the vehicle controls, or treatments of acetaminophen or thioacetamide.

Cells were treated at dose levels approximating the LD$_{50}$ values previously determined. Through immunofluorescent staining and flow cytometry, I was able to detect cells that expressed β-gal, which indicated a frameshift mutation had occurred. The frameshift mutations observed, particularly in the cells with the mononucleotide repeat, did not correlate well with the overall induction of ROS by each of the treatments. Similarly, the anticipated result that direct DNA damage by ROS would be more prevalent in cells with the mononucleotide construct, given its guanine sequence, and the propensity for ROS adducts to be located on this nucleotide was also not observed. The data suggest that ROS generation is not directly damaging the DNA sequence thereby inducing the mutations observed.

While DNA damage by ROS does not seem to be the cause of the mutations detected, an interesting observation was noted in the cell populations with treatments that showed ROS generation, most notable with Nod. That is, GFP fluorescence increased in certain cells along with detectable β-gal fluorescence. A similar increase in GFP fluorescence has also been observed in methotrexate-treated cells.$^{77}$ Upon further examination, the increase in GFP fluorescence correlated strongly with Nodularin-induced mutations indicating that this population of cells represents those cells directly affected by Nod treatment.
Acetaminophen induced a large treatment-response ratio compared with its vehicle control in the cells containing the di-nucleotide repeat. From the mononucleotide repeat containing cells, Nodularin treatment induced the largest ratio of mutations. Considering these treatment-specific mutations are detectable from unique types of reporter constructs, the data suggest that the mutational events occur via different mechanisms of action and that each may be due to an induction of a compensatory response in the cells.

In addition to the mutation experiments, the NMuLi cells exhibited changes in morphology and assumed a more “fibroblast-like” appearance. Immunohistochemical analysis for the expression of albumin (a marker of hepatocyte function) resulted in the differential expression between these cell types. That is, the cells with the morphology similar to the parental NMuLi cell line stained positive for albumin, whereas those cells with a fibroblast-like morphology did not.

6.1.2 \textbf{EαβM_{out} Transgenic Mouse Line}

Despite the utility of cell lines expressing the microsatellite-containing \textit{pαβE-ntR} plasmids, I also aimed to study the effects of potential hepatocarcinogens \textit{in vivo}. Again, this model system was designed so that the data obtained would account for any context-specific response on the action of each compound. The creation of an \textit{in vivo} model system also provides the opportunity for the compound needing bioactivation by the CYP450s to exert its true effects, due to the normal metabolic function of the liver. In cell lines, some
of the CYP450s may remain active, but many enzymes have reduced or even absent expression compared to that of an intact liver. Studying the compounds in an \textit{in vivo} model also allows the opportunity for studying if there is a cooperative or synergistic effect between enzymes that may alter these effects.

Through the use of pronuclear injections of the purified bicistronic transgene, I established a transgenic mouse line on a C57Bl/6 background. This background was selected to corroborate any data I obtained with that of other mice used in the study of toxicology to detect mutational events.

After several attempts to create both di-nucleotide and a mono-nucleotide microsatellite containing mice I was only successful in obtaining one transgenic mouse. This mouse I developed harbored the mono-nucleotide (G)$_{13}$ repeat sequence that placed the bacterial lacZ gene that encodes β-galactosidase protein, into an out-of-frame codon configuration.

PCR screening of this mouse confirmed that she carried over one hundred copies of the transgene integrated into her DNA. Pups obtained from the mating of this mouse indicated that the transgene was transmitted in the germline and PCR analysis again confirmed the copy number to be well over 100 copies. So as to prevent the disruption of any genes that may be critical in development or maintenance I opted to maintain the transgenic line as heterozygous. Although preliminary results indicated that the transgene was expressed (via protein analysis), upon further characterization, I find no consistent evidence as to its expression, particularly through RNA analysis, was identified. Given this unexpected result, this transgenic mouse line was not useful.
6.2 Future Directions

Now that the models are established, there are many possibilities for future experiments, I will consider each separately. I should begin by acknowledging that the plasmids I have designed, could be altered by subcloning methods to incorporate a different promoter sequence, thereby making these reporters specific to any tissue or cell type or broad enough, as would be the case using a β-actin promoter, to be expressed universally within an entire organism.

6.2.1 pαβE-ntR Cell Lines

Studies incorporating the pαβE-ntR cell lines would be greatly supported by the development of new or improved methodologies whereby, the cells would not have to be fixed and could instead remain viable. In doing so, it would then be possible that mutated cells could be monitored over time and mutations observed through cellular generations.

In terms of elucidating the mechanism causing the frameshift, it will be necessary to determine where the exact frameshift event is, whether it is DNA or RNA-based or whether neither apply and there is another cause for the mutation. Direct DNA sequencing or PCR-based microsatellite screening would be advantageous in this regard. Similarly, RNA could be examined through reverse transcription into cDNA and screened the same way. Methylation detection PCR would also be of interest if no observable mutation was seen in the DNA sequence or the transcribed RNA.
Other future studies will include treating cells continually over longer periods of time, or in a series of multiple treatments to observe if there is a threshold of cellular defense against chemical insult that might be surpassed. This would be of special interest in the case of ROS generating compounds as oxidative stress has been associated with non-genotoxic carcinogenesis. Studies aimed at examining the mechanisms behind these mutations might also be designed. These could include the incorporation of antioxidant treatment and its effects on the mutations observed. Particularly in the case of nodularin and its ability to induce ROS generation, antioxidants, i.e. melatonin (a scavenger of ROS) have shown to be effective in reducing the effects of treatment and if related to the mechanisms of mutation could reduce the mutations observed in the studies.

Expansion of the utility of the NMuLi cell line beyond solely that of this model system, comes from the observation that this cell line is capable of exhibiting morphological, as well as functional, changes in cell type. These changes provide good evidence that this cell line is a good candidate for studying the process of epithelial-mesenchymal transition (EMT). Additional studies will be designed to further characterize these cell types. These will include additional immunohistochemistry with antibodies such as those against vimentin (a mesenchymal cell protein) and select cytokeratins that are markers of epithelial cells.
6.2.2 EαβM\textsubscript{OUT} Transgenic Mouse Line

I have shown that the EαβM\textsubscript{OUT} transgenic mouse line, despite having a high copy number integrated into its genome, lacks expression of transgene in both RNA and protein. While future studies intended to look at the transgene expression in the context of foci development, these experiments are no longer possible. Through the isolation and culture of primary hepatocytes from the livers of these transgenic mice, future experiments have been suggested. These experiments would aim to study whether the lack of transgene was due to positional effects within the genome or if there was a novel induction of a heterochromatic state that could be surpassed by treatments with 5-azaC or TSA, methyltransferase and histone deacetylase inhibitors respectively. If a heterochromatic state is suggested, additional experiments could be designed to study the induction and maintenance of this state and incorporate the effects and contribution of microRNA and RNAi in this process.

6.3 Conclusion

Throughout this dissertation, a hepatocyte cell line model system was developed and validated for detecting frameshift mutations that would result from a chemical insult and attempts were made at creating a \textit{in vivo} mouse model system. The \textit{in vitro} model system proved valuable, in that induction of microsatellite instability (seen as alterations the gene containing the microsatellite sequence) could be observed after treatment with compounds such as nodularin with mononucleotide repeats and acetaminophen with dinucleotide
repeats. Unfortunately, the in vivo model was not validated as a useful model due to the lack of expression of the transgene. However through the design of other experiments, it may be useful for other areas of study, such as that of microRNA involvement in the induction and maintenance of a heterochromatic state.

I have focused on compounds that are potential hepatocarcinogens, but are not genotoxic, that is, they do not cause direct DNA adducts, but assumedly influence or disrupt regulatory pathways thus preceding hepatic foci and/or liver tumor development. While much work needs to be done, the microsatellite models I describe could provide the basic framework for understanding the dysregulation of the genome in hepatocarcinogenesis.
## Appendix

### Mouse Metabolism Gene Profile of NMuLi Cell Line

<table>
<thead>
<tr>
<th>Gene description</th>
<th>Gene Symbol</th>
<th>Refseq</th>
<th>Unigene</th>
<th>Average Ct</th>
<th>Expression Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP-binding cassette, sub-family B (MDR/TAP), member 1A</td>
<td>Abcb1a</td>
<td>NM_011076</td>
<td>Mm.207354</td>
<td>26.295</td>
<td>+++</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family B (MDR/TAP), member 1B</td>
<td>Abcb1b</td>
<td>NM_011075</td>
<td>Mm.146649</td>
<td>30.055</td>
<td>+</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family B (MDR/TAP), member 4</td>
<td>Abcb4</td>
<td>NM_008830</td>
<td>Mm.297825</td>
<td>30.04</td>
<td>+</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family C (CFTR/MRP), member 1</td>
<td>Abcc1</td>
<td>NM_008576</td>
<td>Mm.196634</td>
<td>22.55</td>
<td>+++</td>
</tr>
<tr>
<td>Amiloride binding protein 1 (amine oxidase, copper-containing)</td>
<td>Abp1</td>
<td>NM_029638</td>
<td>Mm.213898</td>
<td>31.125</td>
<td>+</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 1 (class I)</td>
<td>Adh1</td>
<td>NM_007409</td>
<td>Mm.2409</td>
<td>28.725</td>
<td>+++</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 4 (class II), pi polypeptide</td>
<td>Adh4</td>
<td>NM_011996</td>
<td>Mm.158750</td>
<td>30.09</td>
<td>+</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 5 (class III), chi polypeptide</td>
<td>Adh5</td>
<td>NM_007440</td>
<td>Mm.3874</td>
<td>19.225</td>
<td>+++</td>
</tr>
<tr>
<td>Aryl-hydrocarbon receptor</td>
<td>Ahr</td>
<td>NM_013464</td>
<td>Mm.341377</td>
<td>24.64</td>
<td>+++</td>
</tr>
<tr>
<td>Aminolevulinate, delta-, dehydratase</td>
<td>Aldad</td>
<td>NM_008525</td>
<td>Mm.6988</td>
<td>22.085</td>
<td>+++</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase family 1, subfamily A1</td>
<td>Aldh1a1</td>
<td>NM_013467</td>
<td>Mm.457973</td>
<td>24.67</td>
<td>+++</td>
</tr>
<tr>
<td>Arachidonate 12-lipoxygenase</td>
<td>Alox12</td>
<td>NM_007440</td>
<td>Mm.12286</td>
<td>31.7</td>
<td>+</td>
</tr>
<tr>
<td>Arachidonate 15-lipoxygenase</td>
<td>Alox15</td>
<td>NM_009660</td>
<td>Mm.4584</td>
<td>32.565</td>
<td>+</td>
</tr>
<tr>
<td>Arachidonate 5-lipoxygenase</td>
<td>Alox5</td>
<td>NM_009662</td>
<td>Mm.41072</td>
<td>30.62</td>
<td>+</td>
</tr>
<tr>
<td>Apolipoprotein E</td>
<td>Apoe</td>
<td>NM_009696</td>
<td>Mm.305152</td>
<td>27.15</td>
<td>+++</td>
</tr>
<tr>
<td>Aryl hydrocarbon receptor nuclear translocator</td>
<td>Arnt</td>
<td>NM_009709</td>
<td>Mm.250265</td>
<td>21.95</td>
<td>+++</td>
</tr>
<tr>
<td>ArsA (bacterial) arsenite transporter, ATP-binding, homolog 1</td>
<td>Asna1</td>
<td>NM_019652</td>
<td>Mm.41475</td>
<td>21.215</td>
<td>+++</td>
</tr>
<tr>
<td>Biliverdin reductase A</td>
<td>Blvra</td>
<td>NM_026678</td>
<td>Mm.22028</td>
<td>23.995</td>
<td>+++</td>
</tr>
<tr>
<td>Biliverdin reductase B (flavin reductase (NADPH))</td>
<td>Blvrb</td>
<td>NM_144923</td>
<td>Mm.24021</td>
<td>24.24</td>
<td>+++</td>
</tr>
<tr>
<td>Carboxylesterase 1</td>
<td>Ces1</td>
<td>NM_021456</td>
<td>Mm.22720</td>
<td>37.025</td>
<td>N/A</td>
</tr>
<tr>
<td>Carboxylesterase 2</td>
<td>Ces2</td>
<td>NM_145603</td>
<td>Mm.28191</td>
<td>26.44</td>
<td>+++</td>
</tr>
<tr>
<td>Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1</td>
<td>Chst1</td>
<td>NM_023850</td>
<td>Mm.38021</td>
<td>27.27</td>
<td>+++</td>
</tr>
<tr>
<td>Catechol-O-methyltransferase</td>
<td>Comt</td>
<td>NM_007744</td>
<td>Mm.100940</td>
<td>20.825</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 1, subfamily a, polypeptide 1</td>
<td>Cyp1a1</td>
<td>NM_009992</td>
<td>Mm.14089</td>
<td>28.38</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 1, subfamily a, polypeptide 2</td>
<td>Cyp1a2</td>
<td>NM_009993</td>
<td>Mm.15537</td>
<td>34.46</td>
<td>+</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily b, polypeptide 10</td>
<td>Cyp2b10</td>
<td>NM_009998</td>
<td>Mm.218749</td>
<td>38.77</td>
<td>N/A</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily c, polypeptide 29</td>
<td>Cyp2c29</td>
<td>NM_007815</td>
<td>Mm.20764</td>
<td>33.69</td>
<td>+</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Accession</td>
<td>Symbol</td>
<td>NM_</td>
<td>Mm.</td>
<td>Ct</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-----------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Cytochrome P450, family 2, subfamily e, polypeptide 1</td>
<td>Cyp2e1</td>
<td>NM_021282</td>
<td>Mm.21758</td>
<td>29.08</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 4, subfamily b, polypeptide 1</td>
<td>Cyp4b1</td>
<td>NM_007823</td>
<td>Mm.1840</td>
<td>24.38</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 11, subfamily b, polypeptide 2</td>
<td>Cyp11b2</td>
<td>NM_009991</td>
<td>Mm.377079</td>
<td>21.22</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 17, subfamily a, polypeptide 1</td>
<td>Cyp17a1</td>
<td>NM_007809</td>
<td>Mm.1262</td>
<td>No Ct</td>
<td>N/A</td>
</tr>
<tr>
<td>Cytochrome P450, family 19, subfamily a, polypeptide 1</td>
<td>Cyp19a1</td>
<td>NM_007810</td>
<td>Mm.5199</td>
<td>27.08</td>
<td>+++</td>
</tr>
<tr>
<td>Cytochrome P450, family 27, subfamily b, polypeptide 1</td>
<td>Cyp27b1</td>
<td>NM_010009</td>
<td>Mm.6216</td>
<td>31.57</td>
<td>+</td>
</tr>
<tr>
<td>Cytochrome b5 reductase 3</td>
<td>Cyb5r3</td>
<td>NM_029787</td>
<td>Mm.22560</td>
<td>19.98</td>
<td>+++</td>
</tr>
<tr>
<td>Epoxide hydrolase 1, microsomal</td>
<td>Ephx1</td>
<td>NM_010145</td>
<td>Mm.9075</td>
<td>20.45</td>
<td>+++</td>
</tr>
<tr>
<td>Fatty acid amide hydrolase</td>
<td>Faah</td>
<td>NM_010173</td>
<td>Mm.256025</td>
<td>28.06</td>
<td>+++</td>
</tr>
<tr>
<td>Fructose bisphosphatase 1</td>
<td>Fbp1</td>
<td>NM_019395</td>
<td>Mm.423078</td>
<td>31.25</td>
<td>+</td>
</tr>
<tr>
<td>Glutamic acid decarboxylase 1</td>
<td>Gad1</td>
<td>NM_008077</td>
<td>Mm.272120</td>
<td>30.88</td>
<td>+</td>
</tr>
<tr>
<td>Glutamic acid decarboxylase 2</td>
<td>Gad2</td>
<td>NM_008078</td>
<td>Mm.4784</td>
<td>31.24</td>
<td>+</td>
</tr>
<tr>
<td>Glucokinase regulatory protein</td>
<td>Gckr</td>
<td>NM_144909</td>
<td>Mm.100043</td>
<td>33.82</td>
<td>+</td>
</tr>
<tr>
<td>Glucose phosphate isomerase 1</td>
<td>Gpi1</td>
<td>NM_008155</td>
<td>Mm.589</td>
<td>26.46</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione peroxidase 1</td>
<td>Gpx1</td>
<td>NM_008160</td>
<td>Mm.1090</td>
<td>20.94</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione peroxidase 2</td>
<td>Gpx2</td>
<td>NM_030677</td>
<td>Mm.441856</td>
<td>30.00</td>
<td>+</td>
</tr>
<tr>
<td>Glutathione peroxidase 3</td>
<td>Gpx3</td>
<td>NM_008161</td>
<td>Mm.20916</td>
<td>24.70</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione peroxidase 5</td>
<td>Gpx5</td>
<td>NM_010343</td>
<td>Mm.1332</td>
<td>No Ct</td>
<td>N/A</td>
</tr>
<tr>
<td>Glutathione reductase</td>
<td>Gsr</td>
<td>NM_010344</td>
<td>Mm.283573</td>
<td>19.06</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 1 (Ya)</td>
<td>Gsta1</td>
<td>NM_008181</td>
<td>Mm.197422</td>
<td>19.12</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 3</td>
<td>Gsta3</td>
<td>NM_010356</td>
<td>Mm.394593</td>
<td>25.45</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, alpha 4</td>
<td>Gsta4</td>
<td>NM_010357</td>
<td>Mm.2662</td>
<td>17.02</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 1</td>
<td>Gstm1</td>
<td>NM_010358</td>
<td>Mm.37199</td>
<td>22.18</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 2</td>
<td>Gstm2</td>
<td>NM_008183</td>
<td>Mm.44086</td>
<td>23.79</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 3</td>
<td>Gstm3</td>
<td>NM_010359</td>
<td>Mm.44085</td>
<td>28.45</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 4</td>
<td>Gstm4</td>
<td>NM_026764</td>
<td>Mm.31203</td>
<td>32.67</td>
<td>+</td>
</tr>
<tr>
<td>Glutathione S-transferase, mu 5</td>
<td>Gstm5</td>
<td>NM_010360</td>
<td>Mm.282351</td>
<td>22.69</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione S-transferase, pi 1</td>
<td>Gstp1</td>
<td>NM_013541</td>
<td>Mm.378930</td>
<td>32.91</td>
<td>+</td>
</tr>
<tr>
<td>Glutathione S-transferase, theta 1</td>
<td>Gstt1</td>
<td>NM_008185</td>
<td>Mm.2746</td>
<td>23.42</td>
<td>+++</td>
</tr>
<tr>
<td>Glutathione transferase zeta 1 (maleylacetocetate isomerase)</td>
<td>Gsto1</td>
<td>NM_010363</td>
<td>Mm.29652</td>
<td>22.15</td>
<td>+++</td>
</tr>
<tr>
<td>Hexokinase 2</td>
<td>Hk2</td>
<td>NM_013820</td>
<td>Mm.255848</td>
<td>22.24</td>
<td>+++</td>
</tr>
<tr>
<td>Hydroxysteroid (17-beta) dehydrogenase 1</td>
<td>Hsd17b1</td>
<td>NM_010475</td>
<td>Mm.188939</td>
<td>28.13</td>
<td>+++</td>
</tr>
<tr>
<td>Hydroxysteroid (17-beta) dehydrogenase 2</td>
<td>Hsd17b2</td>
<td>NM_008290</td>
<td>Mm.276466</td>
<td>23.73</td>
<td>+++</td>
</tr>
<tr>
<td>Hydroxysteroid (17-beta) dehydrogenase 3</td>
<td>Hsd17b3</td>
<td>NM_008291</td>
<td>Mm.5109</td>
<td>No Ct</td>
<td>N/A</td>
</tr>
<tr>
<td>Lactoperoxidase</td>
<td>Lpo</td>
<td>NM_080420</td>
<td>Mm.41236</td>
<td>39.53</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Gene Name</strong></td>
<td><strong>Gene Symbol</strong></td>
<td><strong>Gene ID</strong></td>
<td><strong>Accession</strong></td>
<td><strong>Value</strong></td>
<td><strong>Significance</strong></td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Myristoylated alanine rich protein kinase C substrate</td>
<td>Marcks</td>
<td>NM_008538</td>
<td>Mm.30059</td>
<td>20.45</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 1</td>
<td>Mgst1</td>
<td>NM_019946</td>
<td>Mm.14796</td>
<td>23.62</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 2</td>
<td>Mgst2</td>
<td>NM_174995</td>
<td>Mm.24679</td>
<td>18.73</td>
<td>+++</td>
</tr>
<tr>
<td>Microsomal glutathione S-transferase 3</td>
<td>Mgst3</td>
<td>NM_025569</td>
<td>Mm.21826</td>
<td>21.96</td>
<td>+++</td>
</tr>
<tr>
<td>Myeloperoxidase</td>
<td>Mpo</td>
<td>NM_010824</td>
<td>Mm.4668</td>
<td>No Ct</td>
<td>N/A</td>
</tr>
<tr>
<td>Metallothionein 2</td>
<td>Mt2</td>
<td>NM_008630</td>
<td>Mm.147226</td>
<td>21.63</td>
<td>+++</td>
</tr>
<tr>
<td>Metallothionein 3</td>
<td>Mt3</td>
<td>NM_013603</td>
<td>Mm.2064</td>
<td>31.955</td>
<td>+</td>
</tr>
<tr>
<td>5,10-methylenetetrahydrofolate reductase</td>
<td>Mthfr</td>
<td>NM_010840</td>
<td>Mm.89959</td>
<td>27.135</td>
<td>+++</td>
</tr>
<tr>
<td>N-acetyltransferase 1 (arylamine N-acetyltransferase)</td>
<td>Nat1</td>
<td>NM_008673</td>
<td>Mm.14125</td>
<td>30.81</td>
<td>+</td>
</tr>
<tr>
<td>N-acetyltransferase 2 (arylamine N-acetyltransferase)</td>
<td>Nat2</td>
<td>NM_010874</td>
<td>Mm.4695</td>
<td>23.355</td>
<td>+++</td>
</tr>
<tr>
<td>Nitric oxide synthase 3, endothelial cell</td>
<td>Nos3</td>
<td>NM_008713</td>
<td>Mm.258415</td>
<td>35.375</td>
<td>N/A</td>
</tr>
<tr>
<td>NAD(P)H dehydrogenase, quinone 1</td>
<td>Nqo1</td>
<td>NM_008706</td>
<td>Mm.252</td>
<td>27.5</td>
<td>+++</td>
</tr>
<tr>
<td>Pyruvate kinase liver and red blood cell</td>
<td>Pklr</td>
<td>NM_013631</td>
<td>Mm.383180</td>
<td>21.39</td>
<td>+++</td>
</tr>
<tr>
<td>Pyruvate kinase, muscle</td>
<td>Pkm2</td>
<td>NM_011099</td>
<td>Mm.326167</td>
<td>18.185</td>
<td>+++</td>
</tr>
<tr>
<td>Paraoxonase 1</td>
<td>Pon1</td>
<td>NM_011134</td>
<td>Mm.237657</td>
<td>33.75</td>
<td>+</td>
</tr>
<tr>
<td>Paraoxonase 2</td>
<td>Pon2</td>
<td>NM_183308</td>
<td>Mm.126984</td>
<td>24.705</td>
<td>+++</td>
</tr>
<tr>
<td>Paraoxonase 3</td>
<td>Pon3</td>
<td>NM_173006</td>
<td>Mm.9122</td>
<td>27.07</td>
<td>+++</td>
</tr>
<tr>
<td>Swi/SNF related matrix associated, actin dependent regulator of chromatin, subfamily a-like 1</td>
<td>Smarcal1</td>
<td>NM_018817</td>
<td>Mm.274232</td>
<td>22.09</td>
<td>+++</td>
</tr>
<tr>
<td>Stannin</td>
<td>Snn</td>
<td>NM_009223</td>
<td>Mm.325800</td>
<td>21.745</td>
<td>+++</td>
</tr>
<tr>
<td>Steroid 5 alpha-reductase 1</td>
<td>Srd5a1</td>
<td>NM_175283</td>
<td>Mm.422833</td>
<td>22.7</td>
<td>+++</td>
</tr>
<tr>
<td>Steroid 5 alpha-reductase 2</td>
<td>Srd5a2</td>
<td>NM_053188</td>
<td>Mm.38933</td>
<td>27.075</td>
<td>+++</td>
</tr>
<tr>
<td>Glucuronidase, beta</td>
<td>Gusb</td>
<td>NM_010368</td>
<td>Mm.3317</td>
<td>22.76</td>
<td>+++</td>
</tr>
<tr>
<td>Hypoxanthine guanine phosphoribosyl transferase 1</td>
<td>Hprt1</td>
<td>NM_013556</td>
<td>Mm.299381</td>
<td>19.035</td>
<td>+++</td>
</tr>
<tr>
<td>Heat shock protein 90kDa alpha (cytosolic), class B member 1</td>
<td>Hsp90ab1</td>
<td>NM_008302</td>
<td>Mm.2180</td>
<td>16.295</td>
<td>+++</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>Gapdh</td>
<td>NM_008084</td>
<td>Mm.343110</td>
<td>15.925</td>
<td>+++</td>
</tr>
<tr>
<td>Actin, beta, cytoplasmic</td>
<td>Actb</td>
<td>NM_007393</td>
<td>Mm.328431</td>
<td>14.73</td>
<td>+++</td>
</tr>
</tbody>
</table>
References


VITA

Sarah Jean Byers Gramling

Contact Information:
612 Chatham Way
Lititz, PA 17543-8234
Phone: (717) 560-7861
Email: sjb163@psu.edu

Education:
Doctor of Philosophy, Molecular Toxicology, December 2009
The Pennsylvania State University, College of Medicine, Hershey, Pennsylvania

Bachelor of Science, Biology-Vertebrate Physiology, December 1999
The Pennsylvania State University, University Park, Pennsylvania

Minor in Marine Science, December 1999
The Pennsylvania State University, University Park, Pennsylvania

Employment:
Associate Scientist, January 2000 to August 2000
Department of Safety Assessment (Toxicology), SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania

Associate Technician Co-op, January 1999 to August 1999
Department of Safety Assessment (Toxicology), SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania

Meetings Attended:
May-June 2005 Human Genome Organization (HUGO) Mutation Detection 2005, Santorini, Greece
March 2000 Society of Toxicology, Philadelphia, Pennsylvania
June 2000 Conference on Laser Capture Microdissection (LCM), National Institutes of Health, Bethesda, Maryland

Presentations/Abstracts:


Honors/Awards:
2000-2002 Life Sciences Consortium Fellowship

Professional Memberships:
2005-current The American Association for the Advancement of Science (AAAS)