AD_____

GRANT NUMBER DAMD17-98-1-8019

TITLE: Development of General Methods for Identification of Genes Regulated by Oncogenic Stimuli or DNA Damage

PRINCIPAL INVESTIGATOR: Stephen J. Elledge, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030-3498

REPORT DATE: May 1999

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1999	Annual (13 Apr 98 - 12 Apr 99)

4. TITLE AND SUBTITLE
Development of General Methods for Identification of Genes Regulated by Oncogenic Stimuli or DNA Damage

5. FUNDING NUMBERS
DAMD17-98-1-8019

6. AUTHOR(S)
Stephen J. Elledge, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030-3498

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The ability of cancer cells to proliferate inappropriately is due the loss of tumor suppressor genes and the gain of function of oncogenes. Two classes of tumor suppressor genes exist, mutators and growth regulators. Mutators are genes that when mutant cause an increase in the rate of genomic instability and hence the rapid accumulation of mutations of all classes. Genes of the mutator class include mis-match repair genes involved in HPCC, DNA repair genes such as X.P., cell cycle checkpoint genes such as ATM and p53. Recently two breast cancer genes BRCA1 and BRCA2 have been implicated in DNA repair and fall into this class. The second class of tumor suppressors, growth regulators, are those directly involved in regulating cell proliferation or the ability of tumor cells to survive and metastasize. Oncogenes such as myc or ras are dominant and act in opposition to tumor suppressors. Many of the genes in these groups directly or indirectly regulate transcription and the identification of target genes is important for elucidation of these regulatory pathways. In addition, the transcriptional induction of genes in response to DNA damage is a critical component of the cellular response to DNA damage and the prevention of mutagenesis. This grant proposes to identify genes induced by DNA damage and in response to oncogenic stimuli by conventional and novel methods.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	11
	16. PRICE CODE

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

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18 298-102

USAPPC V1.00

FOREWORD

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

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

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

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

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

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

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

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

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

5-11-99

PI - Signature Date

Table of Contents

(5) **Introduction**

The ability of cancer cells to proliferate inapproriately is due the loss of tumor suppressor genes and the gain of function of oncogenes. Two classes of tumor suppressor genes exist, mutators and growth regulators. Mutators are genes that when mutant cause an increase in the rate of genomic instability and hence the rapid accumulation of mutations of all classes. Genes of the mutator class include mis-match repair genes involved in HPCC, DNA repair genes such as X.P., cell cycle checkpoint genes such as ATM and p53. Recently two breast cancer genes BRCA1 and BRCA2 have been implicated in DNA repair and fall into this class. The second class of tumor suppressors, growth regulators, are those directly involved in regulating cell proliferation or the ability of tumor cells to survive and metastasize. Oncogenes such as myc or ras are dominant and act in opposition to tumor suppressors. Many of the genes in these groups directly or indirectly regulate transcription and the identification of target genes is important for elucidation of these regulatory pathways. Many oncogenes and tumor suppressors are involved in signaling pathways that activate the expression of a particular subset of relevant genes. Furthermore, the transcriptional induction of genes in response to DNA damage is a critical component of the cellular response to DNA damage and the prevention of mutagenesis. The identification of transcriptional target genes is critical to understanding how these regulatory pathways function. In this grant we propose the identification of target genes that are differentially transcribed in response to oncogenic stimuli or DNA damage to better our understanding of these processes and to provide tools for future analysis.

(6) **Body**

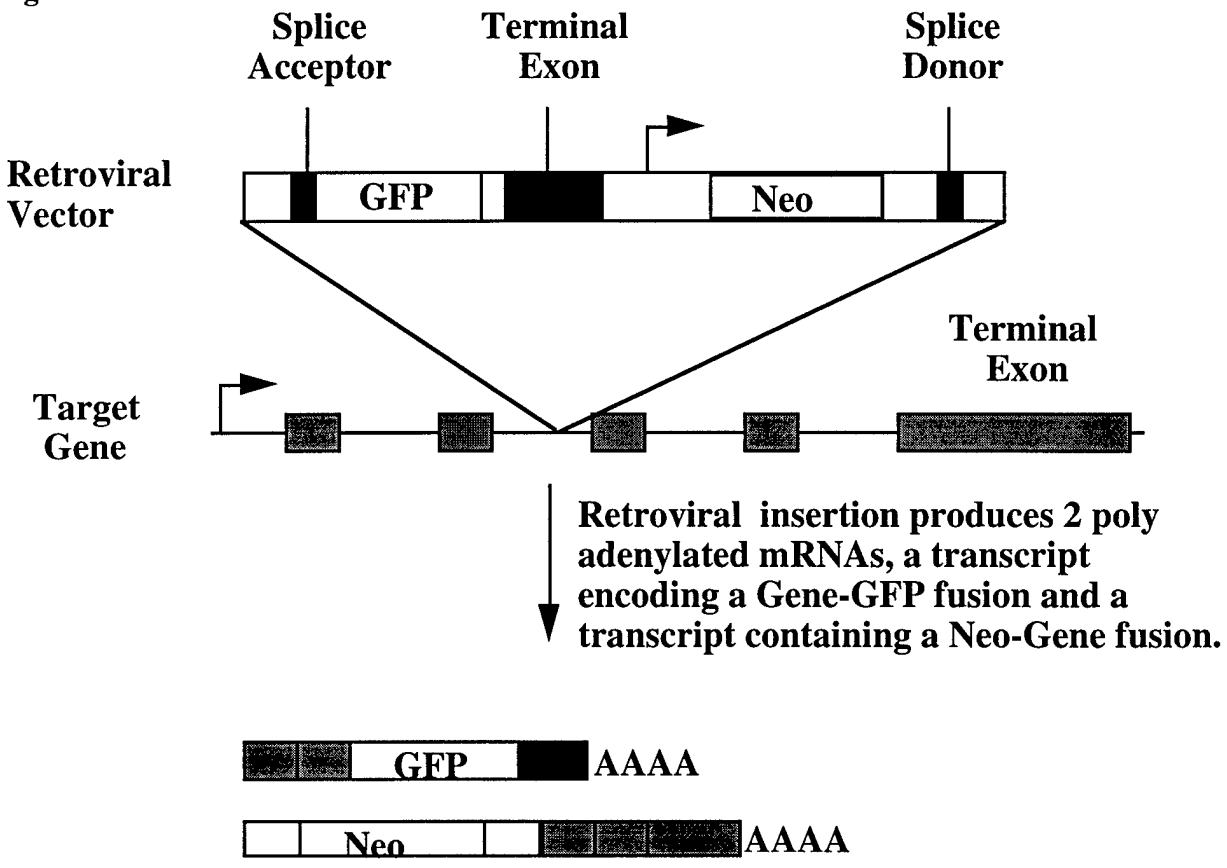
We have proposed to identify genes that are transcriptionally activated in response to DNA damage or potentially in response to oncogenic stimuli such as activation of ras or loss of Rb. We initially proposed to develop new technology to identify such genes by using retroviral insertions into genes that would place GFP under the control of endogenous promoters. We have run into problems with the design of these vectors as proposed in Task 1 and Task 2.

As an initial test of our concept, we placed GFP[1,2,6] under the control of the retroviral LTR. This construct provided a strong promoter and an optimal expression context for GFP. We were able to successfully package this control vector and showed that it could infect cells. We next infected cells containing an IRES sequence and GFP and could show that IRES GFP was able to express well when run from the LTR promoter as had been previously shown for other promoters carrying IRES and reporters such as lacZ and luciferase[3,4,5]. We next tried to simulate the expression from and insertion into an open reading frame. We did this by inserting a long cDNA between the LTR and IRES-GFP. The long cDNA we chose was the coding sequence for the BRCA1 gene. When the construct LTR-BRCA1-IRES-GFP was infected into cells, we were unable to see GFP expression. To test whether this was due to an inability to properly package because the construct was too big, we transfected the construct and one containing only the LTR-IRES-GFP as a control. Under these circumstances, the LTR-IRES-GFP construct expressed GFP well but the LTR-BRCA1-IRES-GFP failed to express significant levels of GFP. This has caused us to question the overall design of the vectors and for this reason we have had a delay in making the vectors we proposed in Tasks 1 and 2.

We are still making the original vector now using as a backbone some previously tested vectors. To do this we have modeled these vectors on the previously tested ROSA-β-gal vectors used for promoter trapping in ES cells. However, we feel that it might be better

to have protein fusions rather than transcriptional fusions because our experience with long transcripts in front of the IRES-GFP construct is likely to mean that we will miss a significant number of genes because of long transcripts. We have now redesigned our fusion vectors more along the lines originally outlined in Task 2 with a modification shown below in Figure 1. The problem with the translational fusion vectors as originally designed is that the frequency of productive fusions would be quite low. However, we have come across a modification that will significantly improve the frequency of isolating in frame retroviral insertions[7,8]. This modification allows one to select for the initial retrovirus to insert into a transcription unit. This is achieved by having the selectible marker, neo in this case, to have a splice donor sequence at the 3' end instead of a terminal exon with a poly-A addition site. In order for the retrovirus to express neo, it must integrate into a transcription unit in the proper orientation to splice into the 3' terminal exon. This means that each neo-resistant colony contains an insertion into a transcription unit in the correct orientation such that the 5' end of the retrovirus is in a position to act as a splice acceptor site for making fusions between GFP and the endogenous gene. This should increase our efficiency of insertion to productive GFP fusion by about 20-fold.

Figure 1



Since the goal of this project was to identify DNA damage inducible genes from eukaryotes, we thought it would be useful to also know the spectrum of DNA damage-inducible genes in a model eukaryote S. cerevisiae. We have established a collaboration with Pat Brown's lab to establish the complete list of DNA damage inducible genes from yeast using chip hybridization technology[9]. We have performed a number of analyses and have identified approximately 300 transcripts that are significantly increased in response to DNA damage. Many of these have obvious human homologs that we can test for DNA

damage-inducibility. Interestingly we found that they fall into several categories based on time of induction during a time course and into groups of genes based on function. These genes are shown grouped into functional units in Tables 1 and 2 below. In addition to genes involved in DNA replication, nucleotide synthesis and DNA repair we were surprised to see a number of genes involved in electron transport, carbohydrate metabolism, Amino acid biosynthesis, protein folding and degradation. This significantly changes the way we now look at the likely genes we will find in our screen for mammalian damage inducible genes.

Table 1

Types of yeast genes induced by DNA damage by MMS through a 2 hr time course.

Total number of genes induced over 2-fold	number of genes peaking at 15-30 min	number of genes peaking at 45-60 min	number of genes peaking at 90-120 min
Amino Acid Biosynthesis			
22	14	4	4
Protein Degradation			
39	6	13	20
Protein Folding			
7	6	1	0
Transport			
20	7	4	9
Nucleotide synthesis			
15	0	1	14
Stress-induced genes			
20	3	2	15
Electron transport			
8	0	2	6
Carbohydrate metabolism			
45	0	7	38
DNA replication and repair			
13	0	6	7

Wild type cells were treated with 0.03%MMS (80% viability)

Table 2. Names of 4 classes of yeast genes induced by DNA damage

DNA replication and/or repair	**Nucleotide Biosynthesis**
DUN1 Checkpoint protein kinase	THI1
RNR1* Ribonucleotide reductase	THI5
RNR2* Ribonucleotide reductase	THI12
RNR4* Ribonucleotide reductase	TRR1
TRX2*	TRR2
MAG1 Methyl transferase	UGP1
RFA1 DNA Replication	YNK1
RFA2 DNA Replication	ADE17
POL30 DNA Replication	STF2
PHR1 Photo-repair	YBR284W (similar to AMP deaminase)
RAD51 Recombination	RNR1*
TRR1	RNR2*
YHR029C *(similar to thymidylate synthase)	RNR4*

	TRX2* YHR029C *(similar to thymidylate synthase)
Stress-induced genes	**Genes involved in protein degradation**
SOD1 Superoxide dismutase	RPN6
HYR1	RPT4
YAP1 Transcription factor	RPT6
YHB1	RPN7
TIR1	RPT2
XBP1 Transcription factor	RPN12
TIR2	RPT1
HSP42 Heat shock protein	RPN3
HSP104 Heat shock protein	RPT3
HSP82 Heat shock protein	RPT5
HSP12 Heat shock protein	PRE5
HSP26 Heat shock protein	PRE9
SSE2 Heat shock protein	PRE10
STI1 Heat shock protein	PUP3
SIS1 Chaperone	PRE3
YGP1	PRE2
USV1	PRE8
DDR48	PUP1
GRE2	PRE1
GRE3	UMP1
	UBA1
	UBC5
	UBC8
	UBC13
	UBI4
	UFD1
	YPS6
	YPS6
	APE2
	PRB1
	PEP4
	PRD1
	LAP4
	CDC48
	UBP6
	YAP180
	AUT7
	APG1
	APG7

Another goal of this project is to look at genes inducible by oncogenic stimuli. We are still planning to do this using the retrovirus vectors we are constructing. However, we have also taken an alternative approach to identify such genes. We have set up a collaboration with Dr. Ronald Davis's lab[10] to use human cDNA chips made by Affymetrics to look at the expression pattern of genes in wild-type diploid fibroblasts compared to the same fibroblasts that express the Adenovirus E7 gene that inactivated the retinoblastoma protein. Inactivation of Rb is normally controlled by cyclin-dependent kinases and cyclin-dependent kinase activation is a downstream event in many oncogenic pathways. We have compared the gene expression profiles for 6000 human cDNAs from these two cell lines throughout the cell cycle. We have found clusters of coordinately regulated genes through the cell cycle and the genes maximally expressed in G2/M and in S phase are shown in Tables 3 and 4. We are currently analyzing these genes for differences between the E7 expressing cell and normal fibroblasts. We do not have this data yet.

Table 3 Human genes expressed maximally in G2/M.

Accession Number		Gene function
X66945_at	3	H.sapiens N-sam mRNA for fibroblast growth factor receptor
X54942_at	3	H.sapiens ckshs2 mRNA for Cks1 protein homologue
X79201_at	3	H.sapiens mRNA for SYT
Z36714_at	3	H.sapiens mRNA for cyclin F
X87613_at	3	H.sapiens mRNA for skeletal muscle abundant protein
Z49989_at	3	H.sapiens mRNA for smoothelin
L25876_at	3	Homo sapiens protein tyrosine phosphatase (CIP2)mRNA,
M62994_at	3	Homo sapiens thyroid autoantigen (truncated actin-binding protein)
U47635_at	3	Human D13S824E locus mRNA, complete cds
X14850_at	3	Human H2A.X mRNA encoding histone H2A.X
U14518_at	3	Human centromere protein-A (CENP-A) mRNA, complete cds
M25753_at	3	Human cyclin B mRNA, 3' end
U73379_at	3	Human cyclin-selective ubiquitin carrier protein mRNA,
M96803_at	3	Human general beta-spectrin (SPTBN1) mRNA, complete cds
M86699_at	3	Human kinase (TTK) mRNA, complete cds
U37426_at	3	Human kinesin-like spindle protein HKSP (HKSP) mRNA,
D26361_at	3	Human mRNA for KIAA0042 gene, complete cds
D31885_at	3	Human mRNA for KIAA0069 gene, partial cds
D43948_at	3	Human mRNA for KIAA0097 gene, complete cds
X51688_at	3	Human mRNA for cyclin A
D78514_at	3	Human mRNA for ubiquitin-conjugating enzyme, complete cds
U63743_at	3	Human mitotic centromere-associated kinesin mRNA, complete cds.
U28386_at	3	Human nuclear localization sequence receptor hSRP1alpha mRNA,
U05340_at	3	Human p55CDC mRNA, APC activator
U01038_at	3	Human pLK mRNA, complete cds
M90657_at	3	Human tumor antigen (L6) mRNA, complete cds
M91670_at	3	Human ubiquitin carrier protein (E2-EPF) mRNA, complete cds

Table 4. Human genes maximally expressed during S phase.

Accession Number		Gene function
X68277_at	7	H.sapiens CL 100 mRNA for protein tyrosine phosphatase
X66087_at	7	H.sapiens a-myb mRNA
X89750_at	7	H.sapiens mRNA for TGIF protein
Z46967_at	7	H.sapiens mRNA for calicin (partial)
X62535_at	7	H.sapiens mRNA for diacylglycerol kinase
Z24725_at	7	H.sapiens mitogen inducible gene mig-2, complete CDS
M77140_at	7	H.sapiens pro-galanin mRNA, 3' end
Z24727_at	7	H.sapiens tropomyosin isoform mRNA, complete CDS
U62015_at	7	Homo sapiens Cyr61 mRNA, complete cds
Z46629_at	7	Homo sapiens SOX9 mRNA
J02854_at	7	Human 20-kDa myosin light chain (MLC-2) mRNA, complete cds
U73960_at	7	Human ADP-ribosylation factor-like protein 4 mRNA, complete cds
U44975_at	7	Human DNA-binding protein CPBP (CPBP) mRNA, partial cds
U27655_at	7	Human RGP3 mRNA, complete cds
U27768_at	7	Human RGP4 mRNA, complete cds
U59752_at	7	Human Sec7p-like protein mRNA, partial cds
U28049_at	7	Human TBX2 (TXB2) mRNA, complete cds
X16416_at	7	Human c-abl mRNA encoding p150 protein
M92934_at	7	Human connective tissue growth factor, complete cds
U04636_at	7	Human cyclooxygenase-2 (hCox-2) gene, complete cds
U18300_at	7	Human damage-specific DNA binding protein p48 subunit (DDB2)
U81607_at	7	Human gravin mRNA, complete cds
X52142_at	7	Human mRNA for CTP synthetase (EC 6.3.4.2)
D90209_at	7	Human mRNA for DNA binding protein TAXREB67
X52599_at	7	Human mRNA for beta nerve growth factor
X58377_at	7	Human mRNa for adipogenesis inhibitory factor
L08246_at	7	Human myeloid cell differentiation protein (MCL1) mRNA
U08021_at	7	Human nicotinamide N-methyltransferase (NNMT) mRNA,
U40490_at	7	Human nicotinamide nucleotide transhydrogenase mRNA,
M12174_at	7	Human ras-related rho mRNA (clone 6), partial cds
M62831_at	7	Human transcription factor ETR101 mRNA, complete cds
J03764_at	7	Human, plasminogen activator inhibitor-1 gene, exons 2 to 9

(7) Key Research Accomplishments

A) Determination that IRES-GFP can be expressed from a retroviral vector in mouse cells

B) Determination that long cDNAs upstream of IRES-GFP constructs significantly reduces the amount of expression.

C) Identification of all known DNA damage inducible genes in the yeast S. cerevisiae.

(8) Reportable Outcomes.

None yet.

(9) Conclusions

So far our work has established that we can work with GFP expressing retroviruses and that the expression of an IRES-GFP fusion depends upon its proximity to the transcriptional start. We have decided to focus more on the translational fusion versions of the retrovirus system to get around these problems. We have also identified a number of DNA damage-inducible genes in a lower eukaryote, S. cerevisiae and these will serve as a background upon which to compare the human DNA damage inducible genes we obtain from our screen using the retroviruses proposed. The control experiment for looking at human cDNAs that differ between wild-type fibroblasts and fibroblasts expressing the E7 protein have been working and will help us identify genes whose expression is activated or repressed by Rb,

(10) References

1. M. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. Science 263: 802-5.

2. Ropp, J. D. Donahue, D., Wolfgang-Kimball, J. J., Hooley, J. Y., Chin, R. A., Hoffman, R. A. Cuthbertson, R.A. & Bauer, K. D. (1995) Aequorea green fluorescent protein analysis by flow cytometry.Cytometry 21: 309-317.

3. Peabody, D.S. and Berg, P. (1986) Termination-reinitiation occurs in the translation of mammalian mRNAs. Mol. Biol. Cell. 6: 2695-2703.

4. Levine, F., Yee, J.K., and Friedman, T. (1991) Efficient gene expression in mammalian cells from a dicistronic transcription unit in an improved retroviral vector. Gene 108:167-174.

5. Dirks, W., Wirth, M., and Hauser, H. (1993) Dicistronic transcription units for gene expression in mammalian cells. Gene 128: 247-249.

6. Heim, R., Cubitt, A.B., and Tsien, R.Y. (1995) Improved green fluorescence. Nature 373:663-664.

7. Krizman, D.B., Berget, S.M. Efficient selection of 3'-terminal exons from vertebrate DNA. Nucleic Acids Res. 1993 21:5198-202.

8. Zambrowicz, BP, Friedrich, GA, Buxton, EC, Lilleberg, SL, Person, C, Sands, AT. Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature 1998 392:608-11.

9. DeRisi, J.L., Iyer, V.R., Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 1997 278:680-6.

10. Cho, R.J., Campbell, M.J., Winzeler, E.A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T.G., Gabrielian, A.E., Landsman, D., Lockhart, D.J., Davis, R.W. A genome-wide transcriptional analysis of the mitotic cell cycle. Mol. Cell 1998 2:65-73.

(11) Appendices

None