

AD _____

Award Number: W81XWH-04-1-0082

TITLE: PARK2, a Large common Fragile Site Gene, Is Part of a Stress Response Network in Normal Cells that is Disrupted During the Development of Ovarian Cancer

PRINCIPAL INVESTIGATOR: David I Smith, Ph.D.
Yu Zhu, M.D., Ph.D.

CONTRACTING ORGANIZATION: Mayo Clinic
Rochester MN 55905

REPORT DATE: January 2007

TYPE OF REPORT: Final

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

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

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 01-01-2007		2. REPORT TYPE Final		3. DATES COVERED (From - To) 15 Dec 03 – 14 Dec 06	
4. TITLE AND SUBTITLE PARK2, a Large common Fragile Site Gene, Is Part of a Stress Response Network in Normal Cells that is Disrupted During the Development of Ovarian Cancer				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-04-1-0082	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) David I Smith, Ph.D. and Yu Zhu, M.D., Ph.D. E-Mail: smith.david@mayo.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mayo Clinic Rochester MN 55905				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. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES - Original contains colored plates: ALL DTIC reproductions will be in black and white.					
14. ABSTRACT Please see Abstract on page 4.					
15. SUBJECT TERMS Common fragile sites; Parkin; stress response					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			USAMRMC
			UU	20	19b. TELEPHONE NUMBER (include area code)

TABLE OF CONTENTS

Introduction.....	4
Body	5
Key Research Accomplishments	16
References.....	18
Publications Resulting from DOD Support	20

Abstract

PARK2 (Parkin) is an extremely large gene that spans greater than 1.3 megabases of genomic sequence within chromosomal band 6q26. This gene is derived from within the distal end of the highly unstable FRA6E (6q26) common fragile site (CFS). CFSs are large chromosomal regions that are highly unstable, found in all individuals, and prone to deletions and other alterations, especially in developing cancer cells. The central two questions that we want to address with this work are what role does inactivation of Parkin and other large CFS genes play in the development of ovarian cancer and whether these genes function as part of a stress response network. In order to address these two questions, we have analyzed the effect of re-introducing Parkin (and several other large CFS genes) into ovarian cancer cell lines that do not express it. We have now shown that the re-introduction of Parkin, and several other CFS genes, is associated with greater sensitivity to the induction of apoptosis. This is consistent with our hypothesis that the inactivation of these genes contributes to ovarian cancer development. We have now identified 20 extremely large genes like Parkin that reside within CFS regions. To determine if these genes are randomly inactivated during cancer development, we have utilized real-time RT-PCR analysis to measure the expression of a number of these genes including Parkin in panels of cancer cell lines and primary tumors for cancers of the prostate, **ovary**, breast, brain and liver. This analysis reveals a decidedly non-random inactivation of the expression of these genes in different cancers. In addition, we've found that there is a greater inactivation of expression of the large CFS genes (and greater numbers of these genes inactivated) in cancers that are generally more aggressive and have a poorer overall clinical prognosis. This may offer a prognostic test of individual ovarian cancers based upon the number of large CFS genes that are inactivated in each cancer. The second part of our studies was to examine Parkin and other large CFS genes as being part of a stress response system within cells. We have utilized the newly developed whole-genome tiling arrays which contain tiled oligonucleotides across the non-redundant portion of the genome to characterize transcripts within and around Parkin (and several other large CFS genes) and their response to two stresses, hypoxia and treatment with the carcinogen NNK. These studies reveal that there are stress-responsive non-coding transcripts within the large CFS genes and may begin to explain why these genes are so large in the first place. We've also now observed that the expression of many of these non-coding transcripts are altered in ovarian tumors. These studies thus support our overall hypothesis that the large CFS genes function as a stress response system within cells and this system, and the large genes, are uniquely susceptible to genomic instability.

Introduction

Parkin is a gene that spans an extremely large chromosomal region of 1.36 megabases (Mbs). This large gene spans the distal half within the highly unstable FRA6E CFS (6q26) and our novel hypothesis that received funding from the Department of Defense Ovarian Cancer Research Program funding was that Parkin and other large CFS genes were part of a stress response system that is disrupted during the development of ovarian cancer. There were two major goals to this proposal. The first was to determine if inactivation of expression of Parkin and other large CFS genes could contribute to the development of ovarian cancer. The second was to demonstrate that Parkin and other large CFS genes functioned as part of a stress response system within cells. We will now summarize the work that we've completed after our third year of Department of Defense funding. Our key findings are as follows: (1) We've demonstrated that Parkin is indeed inactivated in many ovarian and other cancers and that the re-introduction of Parkin results in growth inhibition of ovarian cancer cell lines; (2) We've now identified

an entire family of very large CFS genes that have similar structures to Parkin and that are also highly evolutionarily conserved (in spite of the fact that they reside within the highly unstable CFS regions; (3) We've shown that these genes are not randomly inactivated in ovarian, or other, cancers and that there is a correlation between inactivation of multiple large CFS genes and cancers that are generally associated with a poorer overall clinical outcome; (4) We've identified the retinoic acid receptor-related orphan receptor alpha (RORA) as a large CFS gene whose expression is abrogated in many different cancers including ovarian cancer. This nuclear transcription factor is extremely interesting because in addition to regulating many key cellular functions, it also appears to function as a stress regulated gene; (5) We've found that many of the large CFS genes appear to play key roles in cellular responses to stress and most particularly oxidative stress; and (6) We've initiated a very novel experiment to characterize the large CFS genes as part of a stress response system within cells utilizing the newly developed whole genome tiling arrays.

Body

We would again like to thank the Department of Defense for their support of our work. We continue to believe that the characterization of the common fragile sites (CFSs) is important because these large regions of genomic instability, that are found in all individuals, are chromosomal regions that are uniquely sensitive especially in developing ovarian cancers. In addition, many of them contain large genes which are particularly attractive tumor suppressor candidates which participate in ovarian cancer development. The first two genes identified within these unstable chromosomal regions were FHIT and WWOX. These two genes have a very unusual genomic organization as they both span extremely large genomic regions (1,2) greater than 1.0 Mb in size. In spite of this, the final processed transcripts of these genes are relatively small (1.0 Kb for FHIT and 2.0 Kb for WWOX), thus the majority of these genes (greater than 99.9%) are comprised of intronic sequences. These two genes have been demonstrated to function as tumor suppressors both *in vitro* and *in vivo* (3-5). In addition, inactivation of these genes is associated with a poorer clinical outcome (6-9). Finally, both genes are involved in cellular responses to stress (10-12).

The original goal of this proposal was to characterize another large CFS gene, the 1.36 Mb Parkin gene. This gene spans the third most unstable CFS region, FRA6E (6q26) (13). Our two main goals were: (A) to characterize the Parkin gene in ovarian and other cancers and to determine if the inactivation of this gene was a frequent event in ovarian cancer and to determine if this had any functional significance; and (B) To determine if Parkin and other CFS genes were involved in the cellular responses to stress. We completed the first Specific Aim in the first year of this work and we demonstrated that this gene was indeed frequently inactivated in ovarian cancers as well as in cancers of the prostate, breast and in hepatocellular carcinoma (13-14). We also showed that re-introduction of Parkin into an ovarian cancer cell line that did not produce any endogenous Parkin results in growth inhibition and also protected cells from mitochondria-independent apoptosis induced by ceramide. This work was published in *Oncogene* (13) and *Genes Chromosomes and Cancer* (14). Parkin has been found to also be a stress-responsive gene, similar to FHIT and WWOX, as it recently shown that there is induction of Parkin expression in the presence of oxidative stress (15).

In this report we summarize our work where we have now identified an entire family of extremely large CFS genes. One CFS gene that we've now identified is RORA, the retinoic acid orphan receptor alpha. This 730 Kb gene spans the center of the FRA15A CFS (15q22.2) and is an extremely interesting nuclear transcription factor involved in the regulation of a number of key cellular processes. In addition, this gene is a cellular stress response gene. We describe our work characterizing the RORA gene where we demonstrate that this gene is frequently inactivated in multiple cancers including ovarian cancer and

that it is involved in cellular stress response. The work on RORA was recently published in *Oncogene* (2).

When we originally wrote this proposal, we wanted to characterize how large genes like FHIT, WWOX, and Parkin could be responding to cellular stress. We could not imagine that a powerful technology like tiling arrays would be developed which in a single experiment would enable us to probe the entire genomes response to stress. We have been beta-testing the new 35 bp genome tiling arrays (which contain tiling oligonucleotides spaced 35 bp apart across the entire non-redundant portion of the human genome) and have devised an experiment to measure both coding and non-coding transcripts across the entire genome and their response to two types of stress, growth under hypoxic conditions and exposure to the carcinogen NNK. We have completed this experiment and are beginning to analyze the huge amount of data generated. In the next year, we will be able to determine whether these stresses cause changes in non-coding transcripts which are present within the large introns of CFS genes like Parkin. Our hypothesis is that the reason there are such large genes within the highly unstable CFS regions is that the CFS regions are able to somehow transduce different cellular stresses into the production of the appropriate non-coding transcripts which then regulate the expression of the large CFS genes.

In this report, we therefore summarize our work on the identification of an entire family of large CFS genes, the identification and analysis of the RORA gene, and our preliminary studies utilizing tiling arrays to characterize the entire genomes response to stress.

Large genes within many CFS regions

Our analyses of several CFSs revealed that there were large genes (genes >1.0 Mbs) located within several, but not all, of these regions including FHIT (1.5 Mbs), Parkin (1.36 Mbs), GRID2 (1.39 Mbs), and WWOX (1.0 Mbs). In addition, along with others we demonstrated that FHIT, WWOX, and GRID2 were highly evolutionarily conserved and that the chromosomal regions surrounding them were also CFSs in mice (16-18). This suggested that the large gene and the unstable chromosomal region might be co-conserved because together they serve some function within cells.

We became interested in whether other CFS regions might also include large genes. To address this question, we collaborated with Dr. Robert Kuhn, a researcher at the UCSC Genome Database. Dr. Kuhn provided a list of all genes larger than 500Kb, and we carefully examined the list to remove redundant clones. We generated a list of 240 distinct human genes that spanned greater than 500 Kb of genomic sequence. These 240 genes represent the largest 1% of human genes.

Many of the largest human genes are derived from within CFS regions

Examination of the large gene list revealed that a number of these were derived from chromosomal bands that contained CFSs; we were curious how many corresponded to CFS genes. Our laboratory had already localized 20 of the 89 known CFS regions, and a few other CFS regions have been defined by other groups (19-22). A detailed examination of the sequences surrounding these localized CFSs identified several other large genes as CFS genes, including CNTNAP2 [this is the largest human gene which spans 2.3 Mb within 7q35 (FRA7I)] and LRP1B (1.9 Mb in FRA2F).

We then decided to test several of the largest human genes derived from chromosomal regions known to contain a CFS to determine if they were also CFS genes. A BAC clone covering the approximate center of each large gene was selected, labeled, and used as a FISH probe against metaphases prepared from cells exposed to aphidicolin. This analysis identified several other large CFS genes. However, not every large gene was derived from within a CFS region. Out of the 10 largest genes, 6 were determined to be derived from within CFS regions. Closer examination of the genomic region surrounding many of the

localized CFS regions revealed that slightly less than half of the characterized CFS regions are associated with large genes. We can therefore estimate that there are approximately 40 large CFS genes distributed throughout the genome (23).

The Table below lists the 20 known large CFS genes that have been identified as of today, the size of the genomic region spanned by each gene, the number of exons and the size of the final processed transcripts (FPT), the chromosomal location, and the CFS that spans each gene.

Gene Name	Size	Exons/FPT	Location	Fragile Site
CNTNAP2	2304258	25/8107	7q35	FRA7I
DMD	2092287	79/13957	Xp21.1	FRAXC
LRP1B	1900275	91/16556	2q22.1	FRA2F
CTNNA3	1775996	18/3024	10q21.3	FRA10D
DAB1	1548827	21/2683	1p32.3	FRA1B
FHIT	1499181	9/1095	3p17.2	FRA3B
KIAA 1680	1474315	11/5803	4q22.1	FRA4D
GRID2	1467842	16/3024	4q22.3	FRA4D
Dlg2	1463760	23/3071	11q14.1	FRA11F
Parkin	1379130	12/2960	6q26	FRA6G
IL1RAPL1	1368739	11/2722	Xp21.2	FRAXC
WWOX	1113013	9/2264	16q23.2	FRA16D
PDGFFA	917434	24/2550	4q12	FRA4B
IMMPL2	899238	6/1540	7q31.1	FRA7K
RORA	731967	11/1816	15q22.2	FRA15A
PTPR6	731390	30/4707	3p14.2	FRA3B
Neurobeachin	730417	58/10812	13q13.2	FRA13A
LARGE	647480	16/4326	22q12.3	FRA22B
ARHGAP15	638958	14/1706	2q22.2	FRA2F
SCA1	462345	9/10601	6p22.3	FRA6C

Similarities between the known large CFS genes

The large CFS genes share a number of similarities. Each of these genes is predominantly intronic (greater than 99.7%) and span some of the most unstable chromosomal regions in the genome which are difficult regions to transcribe as well as replicate. In addition, several of the large genes such as FHIT, WWOX, and GRID2 have been found to be highly conserved and the chromosomal regions surrounding them are fragile sites in mice. When comparing what little is known about the function of some of these large genes, it appears that many of them have completely different functions. However, one interesting connection shared by many of the large CFS genes is an association with normal neurological development.

This is already quite clear with Parkin, which when inactivated results in specific cellular death of cells that make and respond to dopamine leading to early onset juvenile Parkinson's disease (24). A spontaneous mouse mutant was identified that had a 1.0 Mb deletion within the distal portion of FRA6E. This deletion removed coding sequences of both Parkin and the immediately adjacent large Parkin co-regulated gene product PACRG. The mice that are homozygous for this deletion have a neurological phenotype known as *Quaking*(viable) (25), which have quake-like tremors caused by improper myelination of the CNS.

The GRID2 gene is a very large CFS gene which was first identified because of spontaneous deletions in mice resulting in a neurological defect known as *Lurcher* (26). Heterozygous *Lurcher* mice display ataxia as a result of selective, cell autonomous and apoptotic death of cerebellar Purkinje cells during postnatal development (26). This gene is also highly conserved between humans and mice and the region surrounding this gene is a CFS in the mouse (22), identical to what was observed for FHIT and WWOX.

Yet another large gene identified by us as a CFS gene is DAB1. Dab1 is the human homolog of the *Drosophila* disabled locus and it interacts with Reelin. When this locus is mutated in mice it results in a neurological defect known as *scrambler* (27). These mice have cerebellar hyperplasia with Purkinje cell ectopia. The normal function of Dab1 is to promote normal positioning of upper layer cortical plate neurons (28).

We tested a number of the largest human genes that were involved in neurological development, which were also derived from chromosomal regions known to contain a CFS and identified a number of other large CFS genes. This includes the Duchene Muscular Dystrophy (DMD) gene in FRAXC, LARGE in FRA22B (which is associated with myodystrophy when deleted in the mouse), and the Sca1 gene (associated with spinocerebellar ataxia) in FRA6C. Thus, many CFS genes are large genes that are involved in normal neurological development.

RORA is another large CFS gene

In the second year of this work we began to characterize another very interesting gene, the 730 Kb retinoic acid related orphan receptor alpha (RORA) which was localized to a chromosomal band 15q22.2 that contained another CFS region, FRA15A. We demonstrated that this gene was localized within the middle of this CFS region. However, unlike our previous studies with Parkin, we did not completely characterize the FRA15A CFS region. Our reason for this was that FRA15A is a CFS region that is not frequently expressed (as are FRA3B, FRA16D and FRA6E) thus we would have had to analyze thousands of metaphase spreads to find sufficient metaphases with breakage/decondensation in the FRA15A region to fully characterize this CFS region. We then analyzed expression of RORA in ovarian cancers as well as in several additional cancers and found that this gene, like Parkin, is frequently inactivated in ovarian (and other) tumors. To determine if this had any functional significance we over-expressed RORA in a breast epithelial cell line and found that this resulted in growth inhibition of the cells (29).

RORA is a particularly interesting gene because it is a nuclear transcription factor that has already been linked to a number of critical cellular functions including control of circadian rhythm, control of cholesterol metabolism and it binds directly to hypoxia inducible factor 1 alpha (30). In addition, RORA is a stress-responsive gene whose expression increases in response to growth under hypoxic conditions, as well as exposure to aphidicolin and chemotherapeutic agents used in the treatment of ovarian cancer. Another interesting thing about RORA is that like Parkin is associated with a neurological defect. Mice that have spontaneous deletions in RORA have the neurological defect *staggerer*, which results from selective death of Purkinje cells (30). This turns out to be a common theme with a number of the large CFS genes, namely that defects in these genes are associated with neurological development and also with the development of cancer.

Because RORA shares similarities with other well studied large CFS genes and is located in a chromosomal band (15q22.2) known to contain the FRA15A CFS, we tested whether RORA was in fact a CFS gene. This analysis revealed that RORA spanned the middle of the FRA15A region. Figure 1 below shows representative FISH results with a BAC from the middle of the RORA gene demonstrating

that in one metaphase (A) the BAC hybridizes distal to the region of decondensation/breakage and in another metaphase (B) the BAC hybridizes proximal.

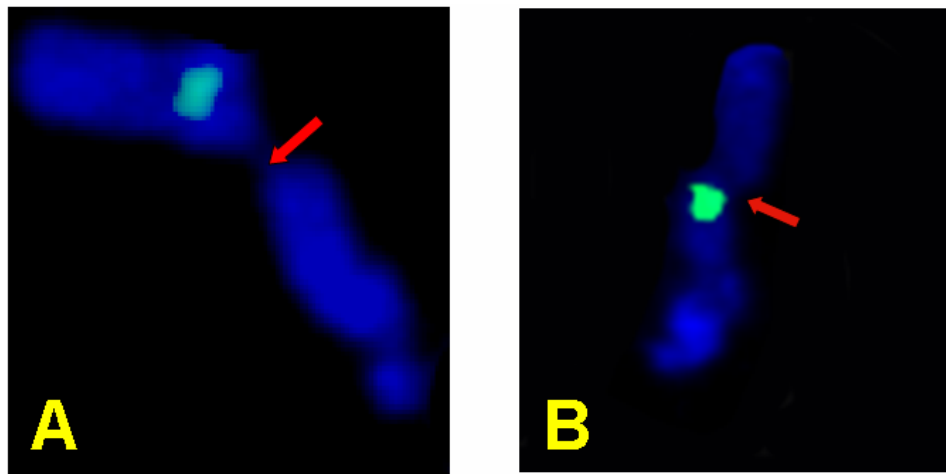


Figure 1 Depiction of FISH results obtained with a BAC clone crossing the middle of RORA and determined to be crossing FRA15A. BAC clone CTD-2034M3 was labeled with biotin and hybridized to normal human lymphocytes treated for 24 hours with 0.4 μ M aphidicolin. 20 metaphases with clear breakage/decondensation at 15q22.2 were scored. The hybridization signal appeared proximal to the break in 12 metaphases and distal in 8, showing that RORA is located in the approximate center of FRA15A. A. Representative metaphase with the hybridization signal appearing distal to the break. B. Representative metaphase with the hybridization signal appearing proximal to the break.

Out of 20 metaphases with good discernible breakage within 15q22.2, we found that this BAC hybridized proximal to the region of breakage 12 times and distal 8 times. This finding would place this BAC clone and the RORA gene itself within the middle and most unstable region of FRA15A.

According to one previously published study, there are four RORA isoforms (RORA 1, 2, 3, and 4) which are produced by alternative splicing. In our studies in various normal tissues, we found expression of only RORA 1 and 4. Figure 2 below shows the transcriptional level of RORA 1 and 4 in various normal tissues.

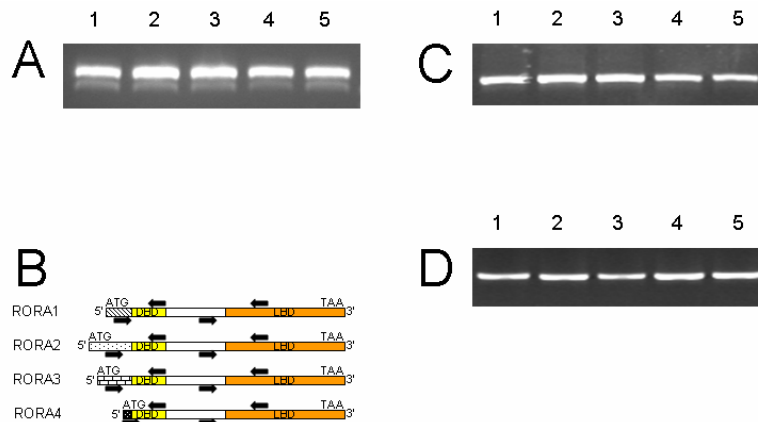


Figure 2 The transcriptional level of RORA and its isoforms in various normal tissues. Lane 1 brain; Lane 2 breast; Lane 3 liver; Lane 4 ovary; Lane 5 prostate. Total RNA was prepared from normal human tissues and cDNA was generated. Semi-quantitative RT-PCR was performed using the universal primers for all RORA isoforms and the specific primers for each isoform to measure the level of RORA. A. RORA universal primers; B. A schematic diagram showing the four different isoforms of RORA. The bold arrows show the positions of specific primers for each isoform and the universal primers for all isoforms. DBD DNA binding domain; LBD ligand binding domain. C. RORA1 (isoform 1) primers; D. RORA4 (isoform 4) primers.

We next examined the expression of RORA in several different types of human cancer samples, either in primary tumors or in tumor-derived cell lines using RT-PCR. This revealed that RORA was down-

regulated in breast, prostate, and **ovarian** cancers (see Figure 3 below). These results are consistent with those obtained in studies of other critical CFS genes including FHIT, WWOX, and Parkin.

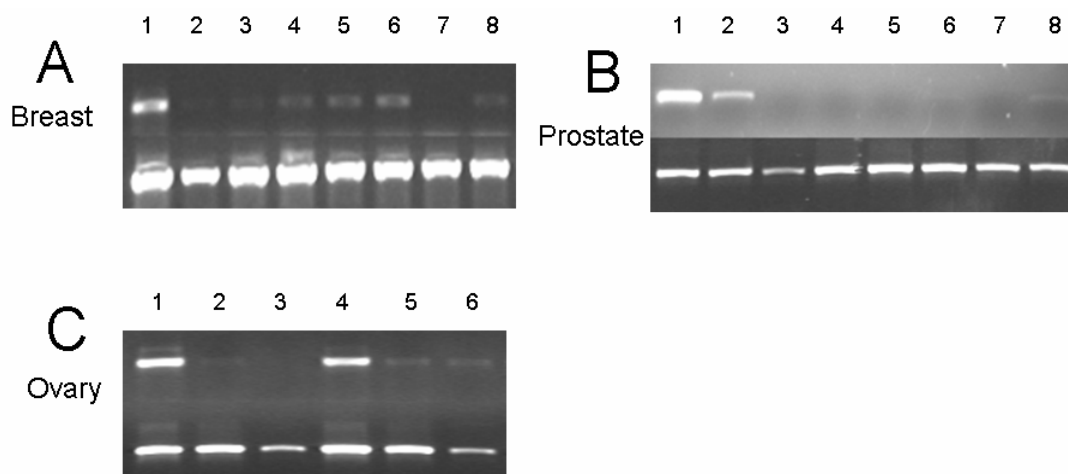


Figure 3. RORA is down-regulated in different types of cancer cell lines and human primary cancers. Total RNA was extracted and reverse transcribed into cDNA. PCR was performed with RORA universal primers to examine the transcriptional level of RORA. A. Breast cancer cell lines, Lane 1 MCF12F; Lane 2 MCF7; Lane 3 MDA157; Lane 4 UACC893; Lane 5 ZR75; Lane 6 MDA435; Lane 7 T47D; Lane 8 BT474. Top row RORA; Bottom row Actin. B. Prostate cancer cell lines and primary tumor samples, Lane 1 normal prostate control; Lane 2 DU145; Lane 3 PC3; Lane 4 LNCaP; Lane 5-8 primary prostate tumor tissues. Top row RORA; Bottom row Actin. C. Ovary cancer cell lines, Lane 1 normal ovarian epithelium control (OSE); Lane 2 OV167; Lane 3 OV177; Lane 4 OV202; Lane 5 OVCAR5; Lane 6 SKOV3. Top row RORA; Bottom row Actin.

We also examined whether RORA expression was modulated by different types of cellular stress other than hypoxia. We first demonstrated that RORA is activated by exposure to aphidicolin (Figure 4) and then subsequently showed similar activation by other types of cellular stress including exposure to ultraviolet radiation (UV), addition of the carcinogen MMS (methyl-methane sulfonate), and treatment with H₂O₂ (oxidative stress) (2). Figures 4 and 5 display the changes in RORA transcripts and RORA protein levels in response to some of these stresses.

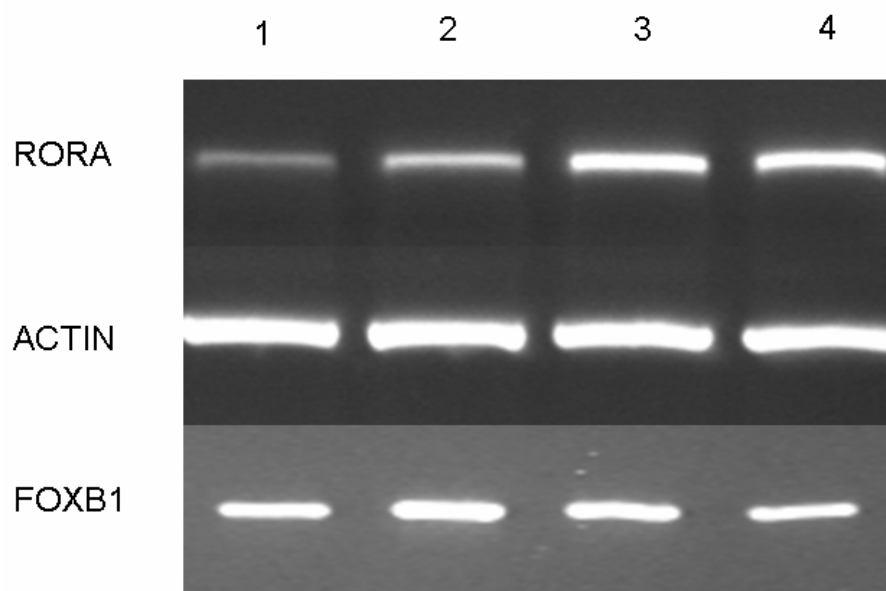


Figure 4 The effect of aphidicolin (APC) on the transcription of RORA in MCF12F cells. MCF12F cells were treated with various doses of APC for 24 hours before total RNA was extracted and cDNA was prepared. PCR was set up to check the transcriptional level of RORA, FOXB1 (a gene within FRA15A right next to RORA) and Actin, using the universal primers for RORA, primers for FOXB1 and the control primers for Actin. Lane 1 cell without APC treatment; Lane 2 with APC 0.2 μM; Lane 3 with APC 0.4 μM; Lane 4 with APC 0.8 μM.

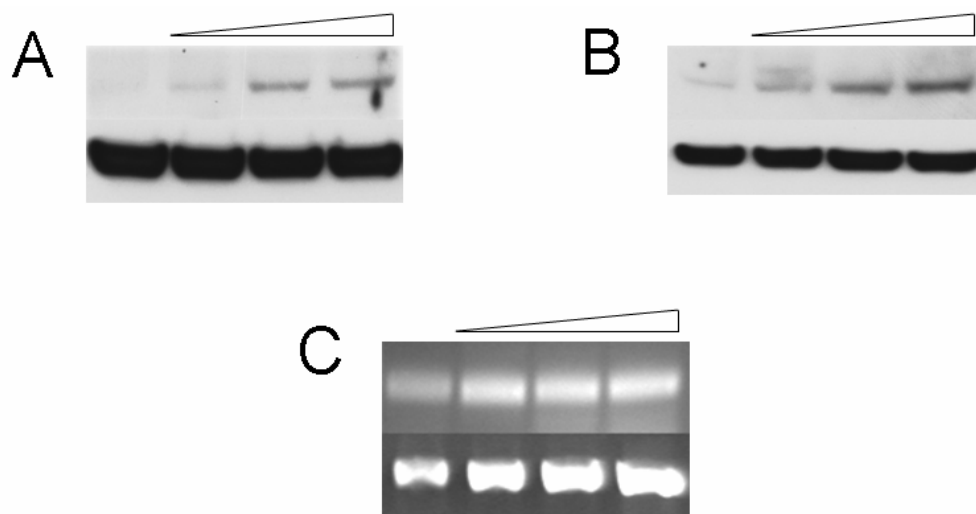


Figure 5. The expression of RORA in MCF12F cells is activated by different types of stress treatments. A. The effect of UV on the protein level of RORA. MCF12F cells were treated with UV at 10, 20 and 50 J/m² and total protein was prepared. B. The effect of MMS on the protein level of RORA. MCF12F cells were treated with MMS at 0.001%, 0.005% and 0.01% for 24 hours. The level of Rora was examined with anti-Rora antibody. C. The effect of H₂O₂ on the transcriptional level of RORA. MCF12F cells were treated with H₂O₂ at 100, 200 and 500 μM for 24 hours and then total RNA was extracted and cDNA was prepared. RT-PCR was performed using universal primers for RORA.

An important question is what role alterations in expression of this large CFS gene play in the development of breast cancer? Indeed all of the CFS genes could be frequent targets of alterations in unstable cancer cells because of the unstable regions that surround them. We transfected RORA into the breast cell line MCF12F and found that increased RORA expression resulted in decreased growth of MCF12F cells (see Figure 6 below). These results are similar to those obtained with FHIT, WWOX, and Parkin.

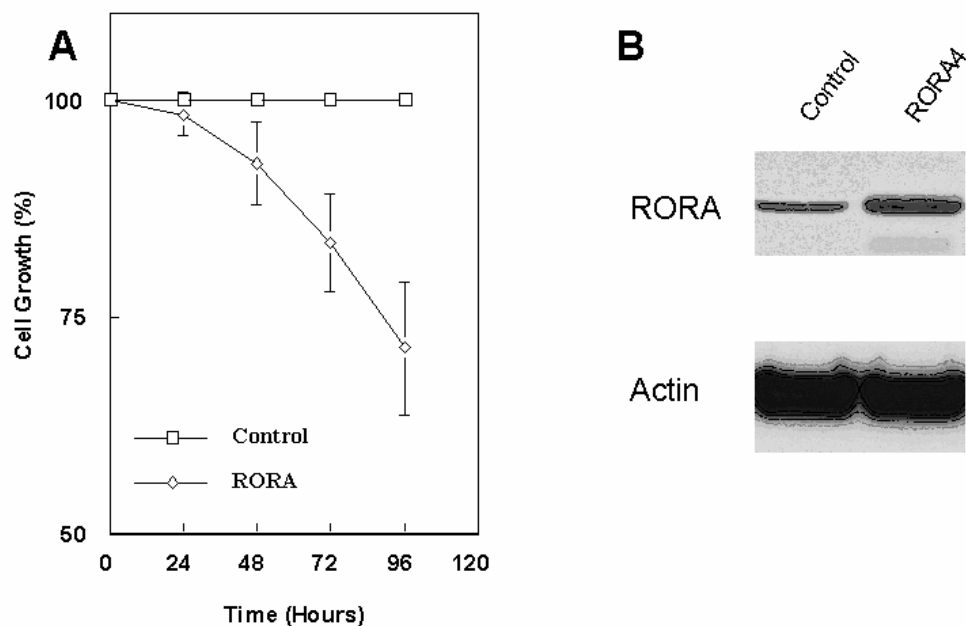


Figure 6 The effect of RORA over-expression on cellular growth. A. MCF12F cells were plated and incubated overnight before transfection. The plasmids (pcDNA3 as control and pcDNA3-RORA4) were transfected into cells using a Lipofectamine 2000 transfection kit following the manufacturer's protocol. The cell number was counted 24, 48, 72 and 96 hours later. All results are the average of at least three independent experiments with standard deviations shown by bars. B. The level of RORA in pcDNA3-RORA4 transfectants detected by the Western blotting assay. Top row RORA; Bottom row Actin.

Thus, changes in RORA expression are associated with readily observable changes in growth rates of MCF12F cells, which supports our contention that inactivation of RORA expression could provide a significant growth advantage to cells thus participating in breast cancer development. All of this work is summarized in our recent paper that was published in *Oncogene* (29).

Expression of the large CFS genes in cancers and cancer-derived cell lines

Our hypothesis is that the large CFS genes are part of a stress response system within cells that is uniquely susceptible to genomic instability and that in cancers with considerable genomic instability, there will be inactivation (alterations) of expression of multiple CFS genes. We have already demonstrated observable phenotypic changes associated with alterations in the expression of these genes. Next we sought to determine whether these genes were randomly inactivated or whether there might be some selection for inactivation of specific CFS genes in different cancers.

To address this question, we used real-time RT-PCR to precisely measure the expression of seven representative large CFS genes (FHIT, WWOX, **Parkin**, GRID2, DLG2, DAB1 and the two expressed RORA isoforms 1 and 4) in panels of primary tumors and cancer-derived cell lines for cancers of the prostate, breast, ovary, liver, and brain. PCR primers were constructed to be optimal for real-time RT-PCR analysis (100-125 bp products derived from the 3' end of the final processed transcripts from these genes) and then we performed real-time RT-PCR in the ABI 7900 real-time PCR machine. To quantify the expression of each of these genes (we constructed primers to differentiate between the two RORA isoforms), we compared the C_t measurements obtained with each gene to that of the β -actin gene and used the delta C_t measurements to quantify message amounts for the large CFS genes. We obtained several normal tissues for comparison for each tissue/tumor type and compared the expression of β -actin and the CFS genes in those normal tissues to panels of cancer-derived cell line, as well as primary tumors of that same type. We considered any gene to be aberrantly regulated if its expression was more than 4-fold up or down relative to the range of expression determined for the normal samples after each sample was run in triplicate.

We found that the expression of these genes was frequently abrogated in different cancers and there appeared to be a very non-random pattern of gene inactivation. We also observed that many cancers had inactivation of multiple large CFS genes. Those cancers with a great deal of genomic instability will have inactivation of many of these genes simultaneously which could have a profound phenotypic effect on those cells. For each of the CFS genes tested, the Table below indicates the number of primary tumors/cell lines that had decreased expression compared to normal samples divided by the total number of tumors/cell lines tested.

	FHIT	WWOX	Parkin	Grid2	Dlg2	Dab1	RORA1	RORA4
Prostate	0/17	1/17	1/17	0/17	2/17	0/17	10/17	1/17
Ovary	2/18	1/18	3/18	1/18	12/18	0/18	2/18	5/18
Breast	4/16	3/16	5/16	0/16	8/16	3/16	3/16	8/16
Brain	7/17	10/17	5/17	9/17	17/18	11/17	5/17	0/17
Liver	4/15	11/15	14/15	14/15	12/15	12/15	15/15	0/15

It is important to note that there was a very interesting preliminary correlation between the frequency of inactivation of these CFS genes and cancers that have very poor clinical outcomes. We found the least inactivation of CFS gene expression in cancers of the prostate, which of the various cancers examined has the best clinical outcome. There was much greater loss of expression of these genes in cancers of the breast and ovary, and many of these cancers tend to be more aggressive than prostate cancers. However,

the cancers with the greatest inactivation of these genes were cancers of the brain and liver. These cancers are highly aggressive, and there is a high probability that patients who develop these tumors will succumb to them.

Disabled-1 (Dab1) is another large CFS gene

Another large gene that we began to examine was the human homolog of the disabled locus from *Drosophila*, Dab-1. This gene spans 1.25 Mbs within chromosomal band 1p32.2, which is also the band that contains the FRA1B CFS. We were therefore interested in whether Dab1 was another large CFS gene that could be important in cancer development. We found that the large Dab1 gene was derived from within the middle of the FRA1B CFS region. We then used real-time RT-PCR analysis to find that Dab1 was inactivated in several different cancers, including ovarian cancer. However, it was more frequently down-regulated in brain and endometrial cancers. We then found that over-expression of Dab1 resulted in decreased cell growth in two different cell lines. Interestingly, when this gene is inactivated in mice it results in the neurological mutant *scrambler* (27) which has abnormal cortical lamination that is indistinguishable from another mouse neurological mutant *Reeler*, caused by mutations in the *reelin* gene.

Monitoring the entire genomes response to stress using Tiling Arrays

Much of the focus in cancer genetics has been on the identification of important alterations during cancer development which could contribute to that process. This focus has been primarily upon the genes, and more importantly the coding portions of those genes. However, the coding portions of genes comprises less than 2% of the entire genome. If you subtract the 45% of the genome that is comprised of repetitive DNA, that still leaves 53% that is neither coding or repetitive. Nowhere is this discrepancy between genome size and apparent coding potential more evident than in some of the large CFS genes. The 1.36 Mb Parkin gene produces a 2960 bp final processed transcript, making this gene over 99.8% intronic. Similarly, the 730 Kb RORA gene has a 1095 bp final processed transcript, hence this gene is also over 99.8% intronic. Why produce such a large initial transcript only to process it down to such a small final processed transcript? One possibility is that are regulatory non-coding RNAs encoded within the intronic sequences which regulate the expression of such large genes like Parkin. However, with no knowledge of where within such large genes these transcripts are derived, there used to be no feasible way to identify these stress-regulated transcripts.

We were considering the construction of sufficient oligonucleotides to probe across the large Parkin introns in order to identify putative non-coding transcripts at a cost of many thousands of dollars when Affymetrix began to beta-test their new whole-genome tiling arrays. The Human Tiling 1.0R Array set contains tiling oligonucleotides spaced 35 bp apart (from the center of one oligonucleotide to the center of the next adjacent oligonucleotide across the entire non-redundant portion of the genome. We obtained multiple sets of these chips as part of the beta-testing of these new arrays from Affymetrix and decided to set up an experiment that would enable us to not only identify transcripts (and non-coding transcripts) across the entire genome but also to ask which of those transcripts were stress-responsive (especially since we believe that the large CFS genes are actually part of a stress-response system within cells).

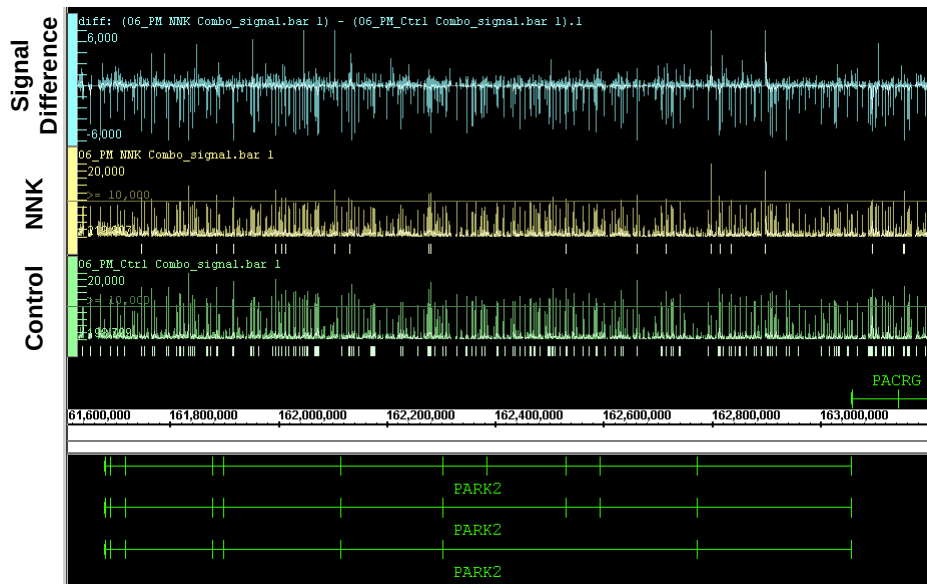
The work with tiling arrays has been pioneered by Dr. Tom Gingeras and co-workers at Affymetrix. Using 5 bp tiling arrays (which have a much higher density of tiled oligonucleotides that are 5 bp apart from center to center of adjacent oligonucleotides, they demonstrated that unannotated, non-polyadenylated transcripts comprise the majority of the transcriptional output for the human genome. This provides additional support for our hypothesis that the large introns in the large CFS genes might produce important stress-regulated transcripts which are involved in the stress-response that the large CFS genes are involved with.

Our microarray experiment was set up to measure the entire genomes response to two different types of cellular stress: growth under hypoxic conditions and exposure to the carcinogen from cigarette smoke-4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK) (24). We cultured normal ovarian surface epithelial cells and exposed them to hypoxia or NNK. Hypoxia is a physiologically important endoplasmic reticulum (ER) stress that is present in all solid tumors. In addition, both Parkin and RORA respond to oxidative stress. We chose NNK because it induces DNA damage, thus we felt that the comparison between the two different types of stress would yield different types of stress-responsive transcripts.

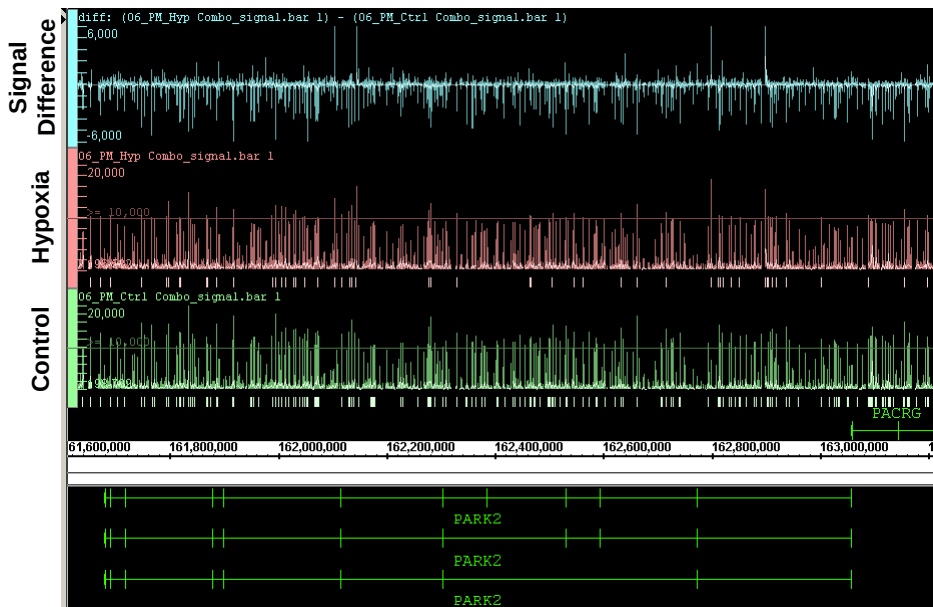
This work was done in collaboration with the Microarray Core of the Mayo Clinic and together, we went through the protocol for producing cDNA from all the RNA species present (not just poly A⁺ RNA) using random oligonucleotide primers. One potential problem with this experiment is that total RNA contains a vast excess of ribosomal RNA which could potentially swamp out signals from both coding genes and non-coding transcripts. There is a commercially available kit using magnetic beads which can purify away the ribosomal RNA (the RiboMinus Kit), and we compared the hybridization of cDNA produced from total RNA to cDNA produced after the use of the RiboMinus kit. The RiboMinus protocol resulted in significant enrichment of non-ribosomal RNA but also resulted in some degradation of the remaining RNA species. We hybridized equivalent amounts of labeled total RNA and RiboMinus purified RNA (-rRNA) to the tiling array chips so that we could compare the hybridization signals to help determine whether the additional expense of the RiboMinus purification kits was worthwhile. We found almost identical hybridization with both sets of RNAs (after being converted to cDNA), thus we did the entire experiment without the RiboMinus purification. We hybridized the three sets of labeled probes (control ovarian surface epithelial cells, cells exposed to hypoxia, or cells exposed to NNK) to two sets of the whole genome 35 bp tiling array each.

In our last years report to the Department of Defense we described some of our preliminary work where we compared total RNA to -rRNA as well as our results when we examined the hybridization to a single chip (one containing chromosome 7 and the largest known human gene CNTNAP2). In the past year we have been analyzing the results with the whole genome hybridization (to a total of 14 chips each run in duplicate). We performed two types of analyses which will be described here. The first was to focus on some of the largest human genes and for this we focused on how oligonucleotides tiled across Parkin and RORA hybridized to the three sources of RNA, thus we could examine both transcripts that hybridized as well as transcripts that were stress responsive.

The Figure below shows the results obtained when we examined the Parkin gene. The bottom Figure (in each case) was the control hybridization. The middle Figure is the hybridization to one of the two stresses utilized, and the top Figure is the subtraction of transcripts seen after stress by the control transcripts. The top Figure is the signals obtained comparing NNK exposure to the control. The bottom of the Figure shows the different Parkin transcripts that are produced (hatched lines represent exons).

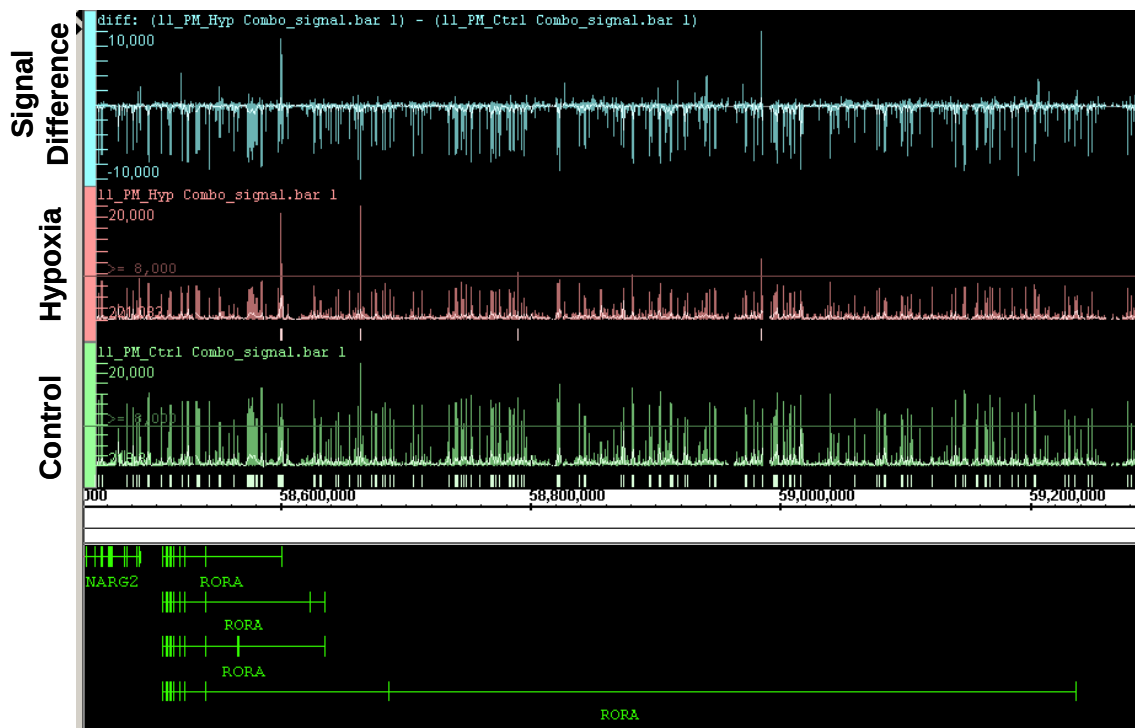
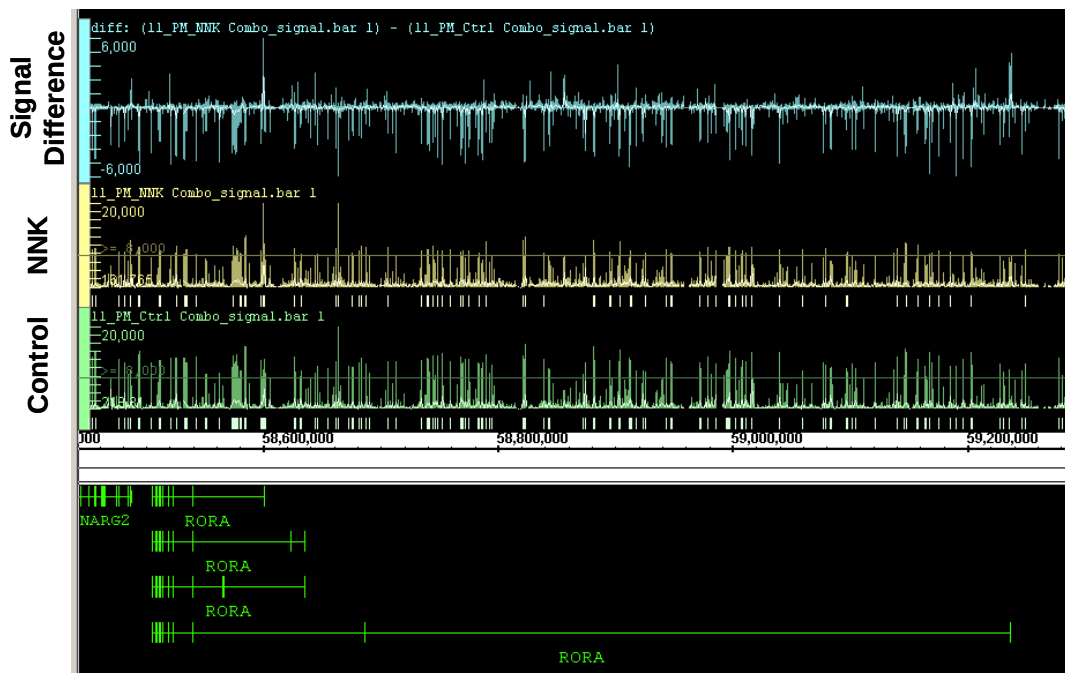


The Figure below shows Parkin comparing the control to hypoxia-induced stress.



Any of the signals that increase in response to the stress are shown as signals above the graph and signals that decrease are shown as signals below the graph. What is very clear from this experiment is that there are multiple transcripts that are not derived from the exon sequences and that there were many more transcripts which decreased in response to one or both of the stresses applied.

The graphs on the following page show the results that we obtained when we examined the RORA gene.



We have asked for a one year no-cost extension to this grant and plan on focusing on characterizing the stress-regulated non-coding transcripts within Parkin and RORA to determine: (1) if they are have altered expression in ovarian cancer; (2) If they are mutational targets in ovarian cancer; and (3) What role they play in normal ovarian surface epithelial cells.

Key Research Accomplishments

We have made several major research accomplishments during the course of this Department of Defense grant. The most important is the identification of an entire family of extremely large CFS genes. There

are now 20 known large CFS genes and we estimate that there may be 45-50 all together. We have found and studied a very interesting stress responsive gene, RORA. This important nuclear transcription factor is involved in the regulation of a number of key cellular processes and we've shown that the expression of this gene is frequently inactivated in many ovarian cancers. Interestingly, this gene, similar to Parkin, is involved in cellular responses to oxidative stress. We have also shown that the expression of the large CFS genes is frequently inactivated in multiple tumor types and that there is more inactivation of their expression in cancers associated with a poorer overall clinical outcome (such as brain or liver cancer). Finally, we've identified a family of new non-coding transcripts which reside within the two large stress-responsive CFS genes we're currently studying. Our next studies will be to characterize these non-coding transcripts to determine what role, if any, they play in the development of ovarian cancer and how they are involved in the oxidative stress response pathways that Parkin and RORA participate in.

REFERENCES

- 1) Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K. The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell* 1996; 84: 587-597.
- 2) Paige AJ, Taylor KJ, Taylor C, Hillier SG, Farrington S, Scott D, Porteous DJ, Smyth JF, Gabra H, Watson JE. WWOX: a candidate tumor suppressor gene involved in multiple tumor types. *Proc Natl Acad Sci USA* 2001; 98: 11417-11422.
- 3) Siprashvili Z, Sozzi G, Barnes LD, McCue P, Robinson AK, Eryomin V, Sard L, Taliabue E, Greco A, Fusetti L, Schwartz G, Pierotti MA, Croce CM, Huebner K. Replacement of Fhit in cancer cells suppresses tumorigenicity. *Proc Natl Acad Sci USA* 1997; 94: 13771-13776.
- 4) Dumon KR, Ishii H, Fong LY, Zanesi N, Pidanza V, Mancini R, Vecchione A, Baffa R, Trapasso F, During MJ, Huebner K, Croce CM. FHIT gene therapy prevents tumor development in Fhit-deficient mice. *Proc Natl Acad Sci USA* 2001; 98: 3346-3351.
- 5) Fabbri M, Iliopoulos D, Trapasso F, Aqeilan RI, Cimmino A, Zanesi N, Yendamuri S, Han SY, Amadori D, Huebner K, Croce CM. WWOX gene restoration prevents lung cancer growth in vitro and in vivo. *Proc Natl Acad Sci USA* 2005; 102: 15611-15616.
- 6) Arun B, Kilic G, Yen C, Foster B, Yardley DA, Gaynor R, Ashfazz R. Loss of FHIT expression in breast cancer is correlated with poor prognostic markers. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 1681-1685.
- 7) Ginestier C, Bardou VJ, Popvici C, Charafe-Jauffret E, Bertucci F, Geneix J, Adelaide J, Chaffenet M, Hassoun J, Viens P, Jacquemier J, Birnbaum D. Loss of FHIT protein expression is a marker of adverse evolution in good prognosis localized breast cancer. *Int J Cancer* 2003; 107: 854-862.
- 8) Pluciennik E, Kusinska R, Potemski P, Kubiak R, Kordek R, Bednarek AK. WWOX- the FRA16D cancer gene: expression correlation with breast cancer progression and prognosis. *Eur J Surg Oncol*, 2006; 32: 153-157.
- 9) Guler G, Uner A, Guler N, Han SY, Iliopoulos D, Hauck WW, McCue P, Huebner K. The fragile genes FHIT and WWOX are inactivated coordinately in invasive breast carcinoma. *Cancer* 2004; 100: 1605-1614.
- 10) Ishii H, Mimori K, Inoue H, Inageta T, Ishikawa K, Semba S, Druck T, Trapasso F, Tani K, Vecchione A, Croce CM, Mori M, Huebner K. Fhit modulates the DNA damage checkpoint response. *Cancer Res* 2006; 66: 11287-11292.
- 11) Thavathiru E, Ludes-Meyers JH, MacLeod MC, Aldaz CM. Expression of the common fragile site genes, WWOX/FRA16D and FHIT/FRA3B is downregulated by exposure to environmental carcinogens, UV, and BPDE but not by IR. *Mol Carcinog.* 2005; 44: 174-182.
- 12) O'Keefe LV, Liu Y, Perkins A, Dayan S, Saint R, Richards RI. FRA16D common chromosomal fragile site oxido-reductase (FOR/WWOX) protects against the effects of ionizing radiation in *Drosophila*. *Oncogene* 2005; 24: 6590-6596.
- 13) Denison SF, Callahan G, Becker NA, Phillips LA, Smith DI. Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. *Genes, Chromosomes Cancer* 2003; 38: 40-52.
- 14) Wang F, Denison S, Lai JP, Phillips LA, Montoya D, Kock N, Schule B, Klein C, Shridhar V, Roberts LR, Smith DI. Parkin gene alterations in hepatocellular carcinoma. *Genes, Chromosomes Cancer* 2004; 40: 85-96.
- 15) Yang YX, Muqit MM, Latchman DS. Induction of parkin expression in the presence of oxidative stress. *Eur J Neurosci* 2006; 24: 1366-1372.
- 16) Shiraishi T, Druck T, Mimori K, Flomenberg J, Berk L, Alder H, Miller W, Huebner K, Croce CM. Sequence conservation at human and mouse orthologous common fragile regions, FRA3B/FHIT and Fra14A2/Fhit. *Proc Natl Acad Sci USA* 2001; 98: 5722-5727.

- 17) Krummel KA, Denison SR, Calhoun E, Phillips LA, Smith DI. The common fragile site FRA16D and its associated gene WWOX are highly conserved in the mouse at Fra8E1.
- 18) Matsusuyama A, Shiraishi T, Trapasso F, Kuroki T, Alder H, Mori M, Huebner K, Corce C. Fragile site orthologs FHIT/FRA3B and Fhit/Fra14A2: evolutionarily conserved but highly recombinogenic. *Proc Natl Acad Sci USA* 2003; 100: 14988-14992.
- 19) Savelyeva L, Sagulenko E, Schmitt JG, Schwab M. The neurobeachin gene spans the common fragile site FRA13A. *Hum Genet* 2006; 118: 551-558.
- 20) Helmich A, Stout-Weibder K, Matthei A, Hermann K, Heiden T, Schrock E. Identification of the human/mouse syntenic common fragile site FRA7K/Fra12C1—relation of FRA7K and other common fragile sites to evolutionary breakpoints. *Int J Cancer* 2007; 120: 48-54
- 21) Helmich A, Stout-Weider K, Hermann K, Schrock E, Heiden T. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res* 2006; 16: 1222-1230.
- 22) Rozier L, El-Achkar E, Apiou F, Debatisse M. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. *Oncogene* 2004; 23: 6872-6880.
- 23) Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett* 2006; 232: 48-57.
- 24) Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 1998; 392: 605-608.
- 25) Lorenzetti D, Antalffy B, Vogel H, Noveroske J, Armstrong D, Justice M. The neurological mutant quaking(viable) is Parkin deficient. *Mamm Genome* 2004; 15: 210-217.
- 26) Lalouette A, Guenet JL, Vriz S. Hotfoot mouse mutations affect the delta 2 glutamate receptor gene and are allelic to lurcher. *Genomics* 1998; 50: 9-13.
- 27) Ware ML, Fox JW, Gonzalez JL, Davis NM, Lambert de Rouvroit C, Russo CJ, Chua SC Jr, Goffinet AM, Walsh CA. Aberrant splicing of a mouse disabled homolog, mdab1, in the scrambler mouse. *Neuron* 1997; 19: 239-249.
- 28) Rice DS, Sheldon M, D'Arcangelo G, Nakajima K, Goldowitz D, Curran T. Disabled-1 acts downstream of Reelin in a signaling pathway that control laminar organization in the mammalian brain. *Development* 1998; 125: 3719-3729.
- 29) Zhu Y, McAvoy S, Kuhn R, Smith DI. RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene* 2006; 25: 2901-2908.
- 30) Dussault I, Fawcett D, Matthysen A, Bader JA, Giguere V. Orphan nuclear receptor RORA alpha-deficient mice display the cerebellar defects of staggerer. *Mech Dev* 1998; 70: 147-153.

Publications Resulting from Department of Defense Support

- 1) Denison SR, Wang F, Becker NA, Schule B, Kock N, Phillips LA, Klein C, Smith DI. Alterations in the common fragile site gene Parkin in ovarian and other cancers. *Oncogene* 2003; 22: 8370-8378.
- 2) Wang F, Denison S, Lai JP, Phillips LA, Montoya D, Kock N, Schule B, Klein C, Shridhar V, Roberts LR, Smith DI. Parkin gene alterations in hepatocellular carcinoma. *Genes Chromosom Cancer* 2004; 40: 85-96.
- 3) Zhu Y, McAvoy S, Kuhn R, Smith DI. RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene* 2006; 25: 2901-2908.
- 4) Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett* 2006; 232: 48-57.
- 5) Smith DI, McAvoy S, Zhu Y, Perez D. Large common fragile site genes and cancer. *Semin Cancer Biol* 2007; 17: 31-41.
- 6) McAvoy S, Zhu Y, Perez DS, James CD, Smith DI. Disabled-1 (Dab1) is a large common fragile site gene- which has decreased expression in endometrial and brain cancer. *Submitted to Genes, Chromosomes and Cancer* 2/1/07.
- 7) Perez DS, Hoage TR, Streng PS, Smith DI. Genomic mapping reveals abundantly expressed, highly conserved non-coding transcripts that are aberrantly expressed in ovarian and breast cancer. *Submitted to Genome Research* 2/11/07.