Thomas Jefferson University

Annual Progress Report: 2006 Formula Grant

Reporting Period

July 1, 2009 – June 30, 2010

Formula Grant Overview

The Thomas Jefferson University received $3,674,642 in formula funds for the grant award period January 1, 2007 through December 31, 2010. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Identification of a New Class of Genes that Contributes to the Development of Breast Cancer - The proposed research will identify a new class of genes that contributes to the development of breast cancer. The researcher's studies of over 2,000 human breast cancer samples have shown the loss of the gene Dachshund predicts a substantially higher risk of death. Breast cancer cellular migration and metastasis are blocked by Dachshund and breast cancer cells are reverted to a precancerous phase. The researcher will characterize the function of this new class of genes to regulate breast cancer development. The researcher has developed a new mouse model which will allow the researcher to remove the Dach1 gene in breast cancer. These mice will allow the researcher to prove Dachshund can block the development of breast cancer. These mice will be valuable to test new therapies for breast cancer.

Anticipated Duration of Project

7/26/2007 – 12/31/2010

Project Overview

Human breast cancer resistant to current forms of therapy kills women each year. Resistance of ErbB2-mediated breast cancer to current therapies arises in women treated with current therapy. New molecular targets governing breast tumor onset and progression are urgently needed. DACH1 represents a novel tumor suppressor pathway. Reduced DACH1 expression correlates with poor prognosis in human breast cancer. Reintroduction of DACH1 into human mammary epithelial cells transformed with distinct oncogenes (Ras, ErbB2, d-Myc) reverts the transformed phenotype in 3D cultures. Revision of the tumor phenotype and the associated genotype led to the DACH1 gene being referred to as a "commandeering" tumor suppressor. DACH1 is structurally distinct to the current known tumor suppressors. This research has identified key functional domains of DACH1 required for inhibition of tumor growth. These residues may represent excellent new targets for breast tumor therapy. As DACH1 functions as a cell fate determination factor, understanding the mechanisms by which DACH1 inhibits breast

tumorigenesis may have important implications for the role of cell fate in tumorigenesis.

Aim 1. Determine the mechanisms governing DACH1//Eya/Six inhibition of breast tumor cell proliferation. The researcher will determine the DACH1 binding proteins governing breast tumor growth. The researcher will define endogenous gene targets of DACH1 using ChiP on chip assays. This will include: 1) determining the components of DACH1/Eya complex inhibiting breast tumor growth; 2) determining the mechanisms by which the DACH1DS domain regulates cyclin D1 and $p21^{CIP1}$; 3) determining the mechanisms by which the DACH1 DS domain is recruited in local chromatin to the cyclin D1 promoter; and 4) identification of chromatin regions occupied by DACH1 and its associated proteins.

Aim 2. Determine the role of DACH1 as a suppressor of mammary tumor onset, progression and metastasis using $DACH1^{fl/fl}$ mice. To determine the role of Dach1 in mammary tumorigenesis; 1) the researcher will deploy inducible DACH1 SIRNA to access DACH1 function suppressor in oncogene collaboration assay. 2) The researcher will determine Dach1 tumor suppressor function using mammary epithelial cell-targeted $Dach1^{fl/fl}$ transgenic mice crossed with MMTV-ErbB2/MMTV-Cre transgenic mice. 3) In order to link the findings of these studies to human breast cancer the research will conduct analysis of human breast cancer samples.

Principal Investigator

Richard G. Pestell, MD, PhD
Thomas Jefferson University
Kimmel Cancer Center
Bluemle Life Sciences Building
233 South 10th Street, Room 1032
Philadelphia, PA 19107

Other Participating Researchers

None

Expected Research Outcomes and Benefits

Completion of these studies will define for the first time a comprehensive analysis of a completely novel model of tumor suppression. The *in vivo* significance will be established using a state-of-the-art molecular genetic approach. The $Dach1^{fl/fl}$ mice and the inducible siRNA viral vectors are innovative and will establish the importance of endogenous Dach1 in tumor suppression. The mechanistic studies, using ChIP on chip, provide the highest level of resolution known, to establish the mechanisms by which DACH1 regulates gene expression. The ChIP on chip analysis are not open ended but are highly focused on a specific question using tiled arrays which are available. The studies of human breast cancer use tissue arrays that are well annotated and previously published by us. Collectively the studies of human and murine breast cancer will provide an integrated understanding of the function of DACH1 in human breast cancer and potentially a new avenue for targeted therapies of human breast cancer.

Summary of Research Completed

Aim 1. Determine the mechanisms governing DACH1//Eya/Six inhibition of breast tumor cell proliferation. The researcher will determine the DACH1 binding proteins governing breast tumor growth. The researcher will define endogenous gene targets of DACH1 using ChiP on chip assays.

Aim 1.4: Identification of chromatin regions occupied by DACH1 and its associated proteins

Summary: The Drosophila Dachshund (Dac) gene, cloned as a dominant inhibitor of the hyperactive growth factor mutant ellipse, encodes a key component of the retinal determination gene network that governs cell fate. Herein, cyclic amplification and selection of targets identified a DACH1 DNA-binding sequence that resembles the FOX (Forkhead box-containing protein) binding site. Genome-wide in silico promoter analysis of DACH1 binding sites identified gene clusters populating cellular pathways associated with the cell cycle and growth factor signaling. ChIP coupled with high-throughput sequencing mapped DACH1 binding sites to corresponding gene clusters predicted in silico and identified as weight matrix resembling the cyclic amplification and selection of targets-defined sequence. DACH1 antagonized FOXM1 target gene expression, promoter occupancy in the context of local chromatin, and contact-independent growth. Attenuation of FOX function by the cell fate determination pathway has broad implications given the diverse role of FOX proteins in cellular biology and tumorigenesis.

Detailed results: Identification of a DACH1 DNA Sequence Specific Binding Site
In order to determine whether *DACH1* encodes a DNA sequence-specific binding protein, CASTing analysis was conducted. An N-terminal FLAG epitope-tagged DACH1 (FLAG-DACH1) expression vector was used to transfect HEK293T cells. Nuclearextracts were prepared and DACH1 was immunoprecipitated using antibodies directed to the FLAG epitope. 75-base oligonucleotideswith a random group of 35 bases were synthesized as double-stranded DNA, incubated with immunoprecipitated DACH1 and PCR reactions conducted after washing. The control IgG was included to monitor the binding specificity. The PCR products were cloned and 40 colonies sequenced (Fig. 1A). The DACH1 consensus-binding sequence (Table S1) encodes a core sequence identified from all 40 sequenced colonies, with 33 colonies sharing the consensus-binding sequence (Fig.1A). Electrophoretic mobility shift assays (EMSA) were used to further characterize the DNA protein interaction. The full-length probe was incubated with nuclear extracts prepared from HEK293T cells transfected with an expression vector encoding FLAG-DACH1.

The DNA/protein complex detected (Fig. 1B, lane 2) was competed by 100-fold molar excess cold probe cognate competitor (Fig.1B, lane 3). Pre-incubation of nuclear extracts with anti-FLAG antibody (M2) delayed the migration of the DNA/DACH1 complex (Fig.1B, lanes 4 *vs.* 5). Immunodepletion using an anti-FLAG antibody abolished the DNA-protein binding complex (Fig.1B, lane 6). FLAG-DACH1 was expressed in HEK293T cells and purified to homogeneity using an anti-FLAG-agarose affinity columnDACH1 purified in this manner was incubated with the DACH1 DNA binding element (DRE). The DNA/protein complex was abolished by the competitor (Fig.1C, lanes 2*vs*. 1). The DACH1/DNA complex was supershifted by the anti-FLAG antibody (Fig.1C, lane 4). Full-length DACH1 produced through *in vitro*

translation also formed a DNA/protein complex that was competed by competitor and was supershifted by the anti-FLAG antibody (Fig.1D lanes 4 and 6). Immunopurified and *in vitro* translated DACH1 protein was detectable by Western Blot directed to the FLAG epitope (data not shown). DACHBox-N is predicted to form the winged/Forkhead class of DNA binding protein (Kim et al., 2002). We therefore determined whether deletion of this domain abrogated DNA binding in EMSA. The deletion mutant of this domain (ΔDS)was unable to bind the DNA probe (Fig. 1B, lanes 7-9, Fig.1D, lane 7). Thus the DNA sequences selected by CASTing serve as a DNA target site for both *in vivo* and *in vitro* synthesized full-length DACH1 protein and DNA binding requires the DACHBox-N domain predicted to form the Forkhead class of DNA binding protein.

DACH1 Encodes a Heterologous Transcriptional Repressor – A multimeric DRE sequence was cloned up-stream of the E4 TATA minimal or β-globin promoter to determine whether the transcriptional repression by DACH1 was promoter selective. DACH1 repressed DRE-Luc transcriptional activity when linked to the E4 TATA box or the β-globin TATA box (Fig. 1E and data not shown). In order to determine whether the DACH1 DNA binding sequence functions asa repressor sequence, the DACH1 binding sequence (DRE) was cloned in either a forward or reverse orientation upstream of heterologous promoters (Fig. 1F and G) to create luciferase reporters (DRE-luc). Transfection of DRE-luc with DACH1 demonstrated DACH1 conveys transcriptional repression through the DNA binding sequence, independently of sequence orientation. In order to determine whether DACH1 conveyed heterologous transcriptional repression function, the DACH1 cDNA was linked to the Gal4 DNA binding sequence and analyzed using the upstream activator sequence linked to the minimal TATA box. DACH1 expression conveyed *trans*-repression function independently of its DNA binding sequence (Fig. 1H).Thus, DACH1 is capable of repressing transcriptional activity in a DNA-sequence specific manner.

Genome-wide Screening of Putative DACH1 Targets **–** In order to determine candidate DACH1 target genes at the whole genome level, the 2000 bp upstream region of human genes in PAINT were scrutinized using the DACH1DNA binding templates (Table S1). A linear sequence search revealed a total of 51 exact matches for the six DNA templates and 2887 unique genes with exact or single (0/1) mismatches(Table S2, S3), presenting a sequence logo for the DACH1 binding site shown in Figure 2A. The Gene Ontology terms that are statistically enriched ($p<0.05$) for the sum of the 0/1 mismatch gene lists are shown in Table S5. As indicated in this table, proteins coded by genes with DACH1 binding sites preferentially populate biological processes essential for signaling such as protein modification, phosphorylation, and translation, as well as metabolic processes associated with nucleosides, RNA and serine family amino acids. The cell compartment and molecular function Gene Ontology categories listed in the Table S5 attest to DACH1's potential for forming complexes with other proteins. KEGG cell cycle pathway and the glioma pathway are populated with genes that have DACH1 binding sequences (Fig. S1, S2) providing a context for the possible roles of these genes in the development and progression of cancer.

In order to identify potential transcription factor partners of DACH1, we further analyzed the gene sets with upstream DACH1 binding sites using gene expression profile associated with MDA-MB-231 cellsat 18 hours after induction of DACH1. Specifically, we used hierarchal

clustering to cluster genes with DACH1 binding sites as well as their sum into four groups:low, medium low, medium high and high expression (Fig.2B). The choice of four groups was optimal as they partitioned the set of genes on the array into approximately equal size subgroups. The median expression values for the four clusters as log2 of the RMA units are as follows: 1.2016, 2.2016; 4.2901; and 8.0081. Figure 2C shows the list of statistically significant transcription factors associated with DACH1 for each of the four gene clusters ($p<0.05$). This set contained a number of Forkhead family transcription factors such as FOXD3, XFD-1, FOXJ2, HFH-3, and XFD2. Next we considered a subset of genes with the DACH1 binding set as determined from microarray data on MDA-MB-231 cells. Specifically, we selected those genes that had undergone at least 1.5 fold change in gene expression between 0 and 18 hours after induction of DACH1 (Fig.2D and Table S4) and supplied this list to PAINT for enriched transcription factors analysis (Fig.2E). Again,the Forkhead family of transcription factors was prominent on the list. These results implicate Forkhead family transcriptional factors as partners in DACH1 associated gene regulation.

Modeling Studies of DACH1 and FOXM1 Binding with DNA – Given the evidence that a subset of genes induced by DACH1 in MDA-MB-231 cells were also enriched in their promoter region for the Forkhead family of transcription factors,we proposed a competition/cooperation model of DACH1 andForkheadproteinsin regulating gene expression.The human *DACH1* gene encodes a protein resembling the conserved winged/helix-turn-helix motif found in FOX proteins. Comparison of the CASTing deduced DACH1 binding sequence identified homology with the consensus Forkhead family protein binding site, in particular the FOXM1 binding site (Fig. 3A). Taken together, these data strongly indicate that DACH1 may blockForkhead protein function in regulating gene expression through direct competing for promoter DNA binding. A prediction model of the DNA binding domain of DACH1 and FOXM1 was generated (Fig. 3C). For DACH1 and FOXM1, the HLH motifs are shown by cyan ribbon spirals and the other helices by red ribbon spirals. Theβ-strands are shown by magenta ribbon with arrows. The loops are shown as gray ribbons. The interacting HLH motif is similar in orientations although the secondary structures shown on the top and the left-side of the proteins between the DACH1 and FOXM1 are different.

DACH1 Attenuates Forkhead-dependent Transcriptional Activity – To investigate the possibility that DACH1 may regulate Forkhead transactivation, luciferase reporter assays were conducted with several Forkhead family proteins including FOXO1 (FKHR), FOXM1, and FOXC2. DACH1 repressed activity of the DACH1-responsive reporter and deletion of the DNA-binding domain (ΔDS) abrogated transcriptional repression (Fig. 3D). Expression of the FOX protein, FOXO1, activated the DACH1 response element ~11-fold (Fig. 3D). The oncoprotein FOXM1 DNA binding sequence resembled the DACH1 DNA binding site. FOXM1 induced DACH1 reporter activity ~12-fold (Fig. 3E). DACH1 repressed activity of a reporter containing six copies of the FOXM1-binding consensus sequence $(FoxA)_6$-Luc by ~60% (Fig. 3F). FOXC2 induction of DACH1 reporter activity was also observed (Fig. 3G). Collectively these observations provided evidence that DACH1 and FOX protein share the same DNA binding sequences and could regulate expression of the same group of genes, raising the possibility that DACH1 may regulate gene expression by attenuating FOX protein function. This was proven by the fact that FOXM1 induction of DACH1 reporter activity was repressed by DACH1 in a dose-

dependent manner (Fig. 3H). Thus, DACH1 attenuates transcriptional activity of FOX protein at DACH1 consensus binding sites.

FIGURE 1 (Wang)

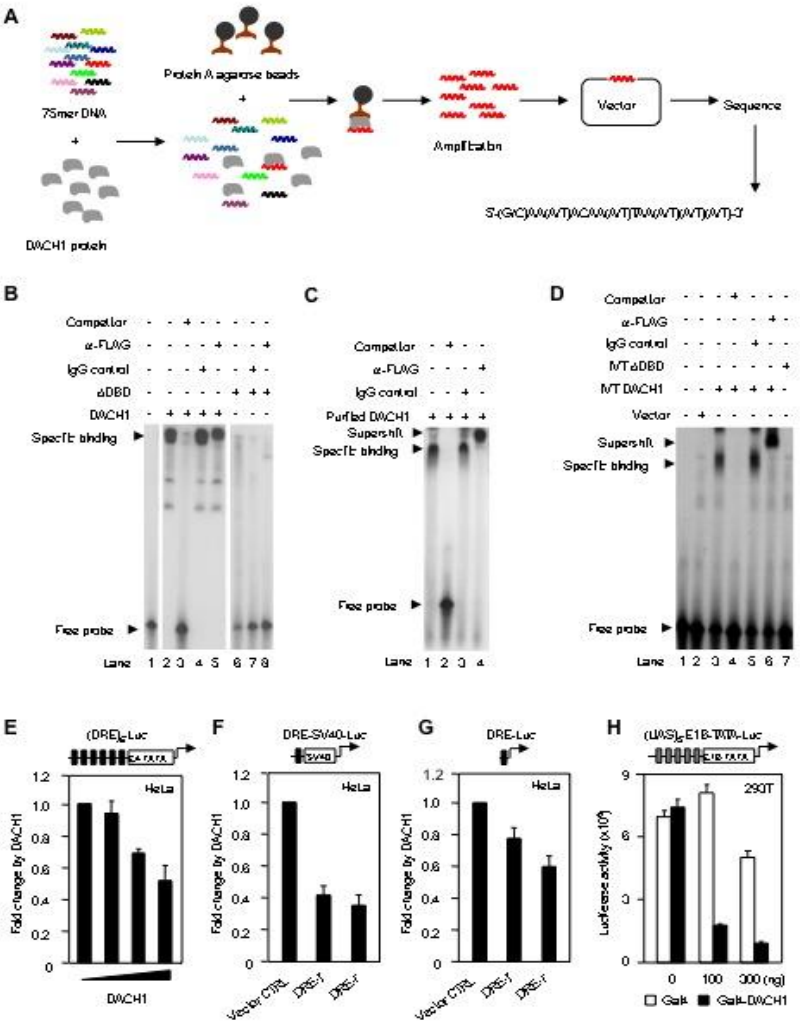
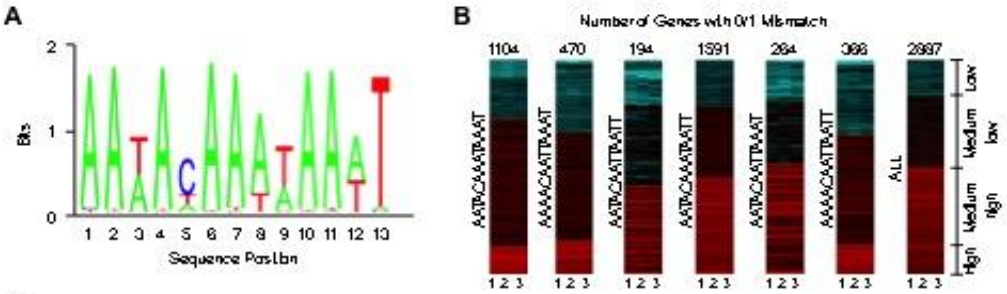


Figure 1. Identification of DACH1-response element. (A) The DACH1 binding sequence was deduced by CASTing. (B) The ^{32}P-labeled oligonucleotide encoding the DACH1 consensus binding site was used in electrophoretic mobility shift assays (EMSA). The probe migration was retarded by DACH1 binding (lane 2). The protein/DNA complex was competed by 100-fold cognate cold-competitor (lane 3) and abrogated by immunodepletion of DACH1 (lane 6). The DACH1 DNA binding domain (DBD) deletion (ΔDS) failed to bind the DNA probe (lane 7). (C) Immunopurified DACH1 was incubated with double-stranded DACH1 probe. The anti-FLAG antibody delayed mobility of the DNA-protein complex. (D) *In vitro* translated DACH1 or ΔDS were incubated with the double-stranded DACH1 DNA binding sequence. Deletion of DBD abolished DNA binding. (E) Cells were transfected with an expression vector encoding DACH1 and luciferase reporter construct containing 6 copies of DRE in the upstream of E4 TATA Box. (F) HeLa cells were transfected with an expression vector encoding the DACH1-response element in either the forward or reverse orientation upstream of either *SV40* promoter or (G) the promoter-less pGL3-basic. (H) DACH1 linked to the Gal4 DNA binding domain was assessed using the UAS-Gal4-DBD binding site linked to the minimal TATA box. Transfection with increasing amounts of Gal4-DACH1 repressed transcription.The data throughout are shown as mean±standard deviation (SD).

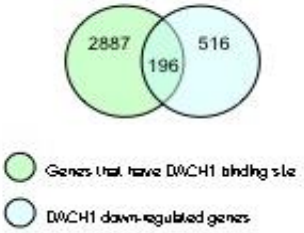
FIGURE 2 (Wang)



C

Medium Low			Medium High			High Expression		
TF	Homology	p-value	TF	Homology	p-value	TF	Homology	p-value
Nkx2-5	72.96%	0.0046	Pax-4	85.42%	0.0004	Pax-4	85.63%	0.0000
FOXD3	**53.86%**	**0.0000**	Nkx2-5	74.41%	0.0000	Nkx2-5	73.57%	0.0000
RFX1	7.73%	0.0068	**FOXD3**	**52.80%**	**0.0000**	**FOXD3**	**47.42%**	**0.0000**
XFD-1	**4.08%**	**0.0012**	**FOXJ2**	**31.09%**	**0.0000**	Pax-6	43.65%	0.0001
NRF-2	3.00%	0.0103	**HFH-3**	**6.01%**	**0.0143**	**FOXJ2**	**28.59%**	**0.0003**
			Arnt	4.79%	0.0499	v-Myb	17.85%	0.0168
			XFD-2	**4.38%**	**0.0011**	NF-Y	10.32%	0.0000
			XFD-1	**2.65%**	**0.0389**	**HFH-3**	**5.65%**	**0.0193**
			MEIS1A/ HOXA9	1.94%	0.0021	E2F	4.67%	0.0209
			MEIS1B/ HOXA9	1.83%	0.0014	SREBP-1	4.67%	0.0169
			RREB-1	1.63%	0.0122	GATA-2	3.35%	0.0000
			Sox-5	1.43%	0.0101	E2	1.88%	0.0286
			Ik-3	1.33%	0.0454	MEIS1A/ HOXA9	1.46%	0.0280
			MEF-2	1.12%	0.0417	MEIS1B/ HOXA9	1.12%	0.0160
			Clox	1.02%	0.0463	HEN1	0.84%	0.0111

D



E

TF	%	p-value
Pax-4	85.71%	0.0024
Nkx2-5	75.24%	0.0000
FOXD3	**50.95%**	**0.0000**
FOXJ2	**31.90%**	**0.0000**
NF-E2	3.81%	0.0093
HFH-1	**3.49%**	**0.0049**
XFD-1	**3.33%**	**0.0064**
GATA-2	2.86%	0.0311
MEIS1A/HOXA9	2.06%	0.0087
MEIS1B/HOXA9	1.59%	0.0221

Figure 2. Genome-wide *In-silico* promoter analysis. (A) A sequence logo depicting the sequence conservation of the template sequence for all single mis-match hits. (B)The clustered gene expression data for the exact and single mismatch genes with minimal expression in cyan and high expression in red. The yellow bars indicate the cutoffs between minimal, low, medium and high expression subsets. (C) The transcription factors that are enriched for the minimal expression, low expression, medium expression and high expression subsets of the genes with exact match or single mismatch DACH1 binding site templates. Homology refers to the percentage of genes within the subset that have the transcription factor. The p-values were calculated using the hypergeometric test ($p < 0.05$). (D)Overlap of DACH1-regulated genes and genes whose promoter has DACH1 binding sequence. (E) The transcription factors that are enriched for DACH1-regulated genes.

FIGURE 3 (Wang)

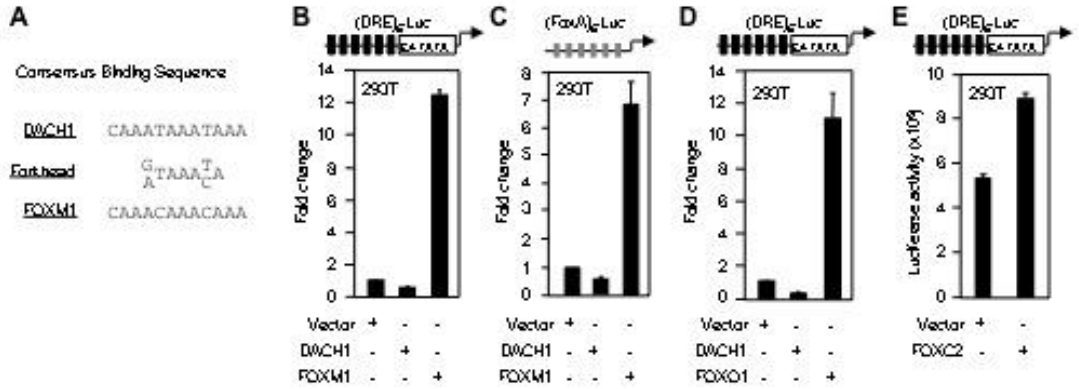


Figure 3. DACH1 inhibits transcription through the Forkhead protein-response element. (A) Human consensus Forkhead binding site and FOXM1 binding site. (B)Schematic representation of DACH1 and FOXM1 expression vectors used for luciferase reporter assay. (C) A ribbon model representation of the FOXM1 protein with a DNA binding site is shown (unpublished). The Helix-Loop-Helix (HLH) domain of DACH1 resembles the FOXM1 DNA binding domain. The DNA is shown as a stick model. The color coding for bases are A: yellow, T:green, G:cyan and C:red . The x-ray structure of the DNA-binding motif of Dachshund is shown in ribbon representation. For both proteins, the segments HLH are shown by blue colored ribbon. α-helices are colored red and the β-strand by purple. The picture was generated by the UCSF Chimera visualization program (Pettersen et al., 2004). (D) The DACH1 response element luciferase reporter gene was transfected into HEK293T cells, showing DACH1, but not DNA binding domain deletion mutant (ΔDBD), repressed reporter activity. (E, F) 293T cells were transfected with $(DRE)_6$-E4-TATA-Luc or $(FoxA)_6$-Luc, showing the DACH1 repressed while FOXM1 induced the reporter activity.The data throughout are shown as mean±standard deviation (SD) from at least two separate experiments with triplicate samples each.

Table S1. Nucleotide frequency matrix for DACH1 binding site

Consensus	A/T	A/T	T	T	A	A/T	T	T	G	T	A/T	T	T
Profile (%)	94	100	100	100	100	100	100	100	100	100	100	97	100
A	31	11			33	18					24	1	0
T	2	22	33	33		15	33	33		33	9	32	33
G									33				
C													

Research Project 2: Project Title and Purpose

Role of the MYC Oncoprotein in Human Cancer - Human cancer is caused by genetic mutations that occur in otherwise normal genes. The most commonly mutated gene in cancer is the MYC oncogene. MYC is mutated so frequently because it is a master regulator of the cell division. MYC accomplished this by controlling the activity of ~3000 other genes. While this number is too large to provide useful information about new drug targets or therapeutic strategies, recent data suggests that only a small subset of these 3000 genes play a direct role in cancer. Using new technology that allows the examination of all genes in a cell at once, the MYC targets actually required for transformation are being defined. The researcher has begun to decipher how these few critical targets participate in converting a normal cell to malignancy. Studies such as these should allow the rational design of new therapeutic strategies.

Anticipated Duration of Project

1/1/2007 - 12/31/2010

Project Overview

Genetic alterations targeting the *myc* oncogene are the most common transforming events in human cancer. These alterations cause overexpression of the transcription factor MYC. A thorough understanding of the MYC transcription network is critical to progress against cancer. Despite intensive study, the essential genes lying downstream of MYC remain poorly understood. Of the 629 reported MYC target genes, only 6 have documented roles in MYC-mediated transformation. An expression-profiling screen that utilized a discrete mutant of MYC was used to selectively identify those targets relevant to transformation. This screen identified 40 novel MYC targets and studies have begun to validate their importance in MYC-driven tumorigenesis. For example, shRNA-mediated depletion of the novel target MTA1 blocks the increased invasiveness of tumor cells caused by MYC. MTA1 reportedly controls the epithelial-to-mesenchymal transition; however its link to MYC was previously unknown. Additional targets from the screen (BAG-1, POLRMT and CD30), were characterized at the molecular level and are now being characterized to determine the biological function they contribute during MYC-mediated transformation. *The aim of this study is to understand the biological roles played by MTA1, BAG-1, POLRMT and CD30 in MYC-driven tumors.*

Together, these studies are likely to identify new cellular functions for MYC, including regulating invasiveness, blocking the anti-tumor immune response and inhibiting the apoptotic stimuli generally associated with oncogene overexpression. Using well-defined systems in which human cells can be transformed by the MYC oncoprotein *in vitro*, each of the targets identified will be dissected in order to understand the precise advantage they provide to MYC-driven tumor cells.

Principal Investigator

Steven B. McMahon, PhD
Thomas Jefferson University
Room 609 Bluemle Bldg.
Kimmel Cancer Center
233 South 10th Street
Philadelphia, PA 19107

Other Participating Researchers

Xiaoyong Zhang, PhD, Maya Varthia, BS - employed by Thomas Jefferson University

Expected Research Outcomes and Benefits

An overwhelming number of deaths from human cancer can be directly attributed to MYC overexpression. As the design of rationally targeted therapies becomes more commonplace, a complete knowledge of the MYC pathway will be critical. The recent identification of target genes activated by MYC has yielded a number of exciting candidates for therapeutic intervention. The best characterized of these targets, MTA1, displays the property that its induction is absolutely required for MYC mediated transformation. Thus the screen significantly enriched the pool of transformation-relevant MYC targets. The identification of MTA1 itself also provides novel insight in the MYC pathway. In general, MYC activation has been considered to be an early, initiating event in human cancer. The demonstration that MYC regulates a step later in the process (i.e., the acquisition of invasive growth capacity) may have important clinical ramifications. For example, this knowledge may impact decisions about when to administer some of the anti-MYC therapies that are now being developed. The other targets from this original screen appear to have an equal likelihood of providing fundamental knowledge critical to understanding and treating cancer. BAG-1 for instance may help explain the lingering enigma of why MYC can induce transformation in some settings and cell death in others. This specific knowledge may be useful for reactivating the cell death pathway in the many cases of cancer where MYC is overexpressed. The ability to selectively activate the intrinsic cell death pathway in MYC expressing tumor cells but not normal cells would be of obvious therapeutic advantage.

Summary of Research Completed

This laboratory previously demonstrated that induction of BAG1 is a critical event that blocks apoptosis in cells expressing high levels of the human oncoprotein MYC. These studies were

conducted using a panel of shRNA molecules to deplete cells of endogenous BAG1 mRNA. In order to advance these studies to pre-clinical relevance, we performed experiments to assess whether small molecule inhibitors of BAG1 could phenocopy the effects observed with genetic depletion.

The chaperone function of BAG1 depends largely on its role in a multi-protein complex that includes HSP70. HSP70 expression and function can be inhibited by treatment of cells with the small molecule flavinoid quercetin. Quercetin, a naturally occurring plant derivative, has long been suspected of having tumor preventive activity. This activity may be in part due to the ability of quercetin to block the pro-survival activity of the HSP70/BAG1 complex in tumor cells that are then poised to undergo apoptosis. For the observation that BAG1 inhibition induces MYC-dependent apoptosis to be clinically relevant, small molecules that mimic the function of BAG1 shRNA treatment must be available. As an initial test of whether this strategy might be worthwhile, the BAG1/HSP70 complex was inhibited by treatment of the U2OS MYC/ER cells with quercetin.

Quercetin treatment had little to no effect on viability in the absence of MYC activation (Figure 1). However, like RNAi-mediated depletion of BAG1, treatment with quercetin resulted in robust cell death when the conditional allele of MYC was activated, even in the presence of growth factor. Quercetin treatment sensitized cells to MYC-induced apoptosis in cells displaying either of two different levels of MYC activity (Figure 1B). These observations support the idea that small molecules that specifically inhibit components of the BAG1 pathway might be particularly useful in sensitizing tumors with high MYC levels to apoptosis.
Since blocking BAG1 function with either shRNA or small molecule treatment resulted in heightened sensitivity to apoptosis, and because ectopic MYC expression can also sensitize cells to apoptosis induced by DNA damaging agents, experiments were conducted to assess whether the simultaneous activation of MYC and depletion of BAG1 can make tumor cells more sensitive to damage induced apoptosis. Treatment of U2OS cells with low doses of UV irradiation failed to induce significant apoptosis, even when the MYC/ER allele was activated (Figure 2). In contrast, when MYC was activated in BAG1 depleted cells, the low dose UV irradiation resulted in substantial cell death. Coupled with the quercetin data, these observations support a model in which blocking BAG1 induction may increase the efficacy of chemotherapeutic agents and lower the doses required.

Collectively, the studies conducted during this funding period have established that targeting BAG1 may be a viable therapeutic strategy in the many forms of human cancer which overexpress the MYC oncoprotein. The data with quercetin suggests that small molecules may be of efficacy as single agents and the studies with radiation suggest that blocking BAG1 function may provide a benefit to cancer patients receiving therapies that are based on the induction of DNA damage.

Publications during this funding period

S.M. Sykes, T.J. Stanek, A. Frank, M.E. Murphy and S.B. McMahon. 2009. Acetylation of the DNA Binding Domain Regulates Transcription-Independent Apoptosis By p53. *J. Biol. Chem.* 284:20197-205

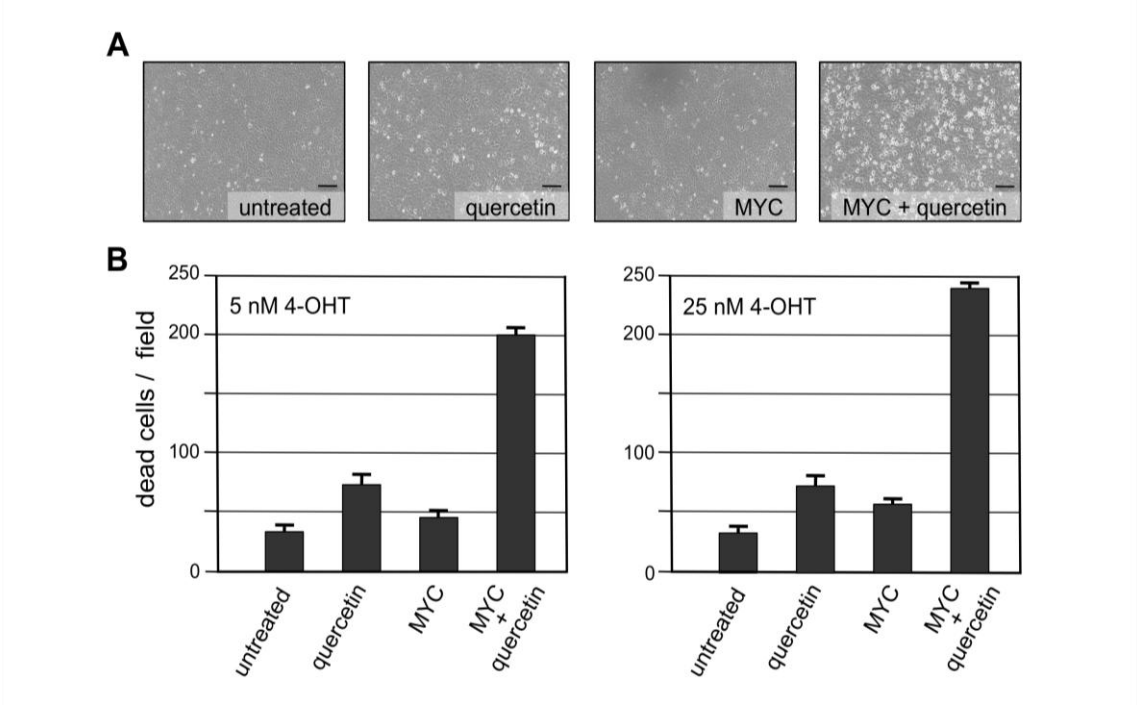


Figure 1 Treatment with a small molecule inhibitor of the BAG1 partner HSP70, sensitizes cells to MYC-dependent apoptosis. MYC was activated in U20S MYC/ER cells by the addition of either 5 or 25 nM 4-0HT, as indicated, for 24 hours. This was followed by treatment with 7 uM quercetin for 72 hours. Cells were photographed **(A)** with scale bar representing 200 µm, and non-adherent, trypan blue positive cells quantitated **(B)**. Bars represent standard error of multiple fields.

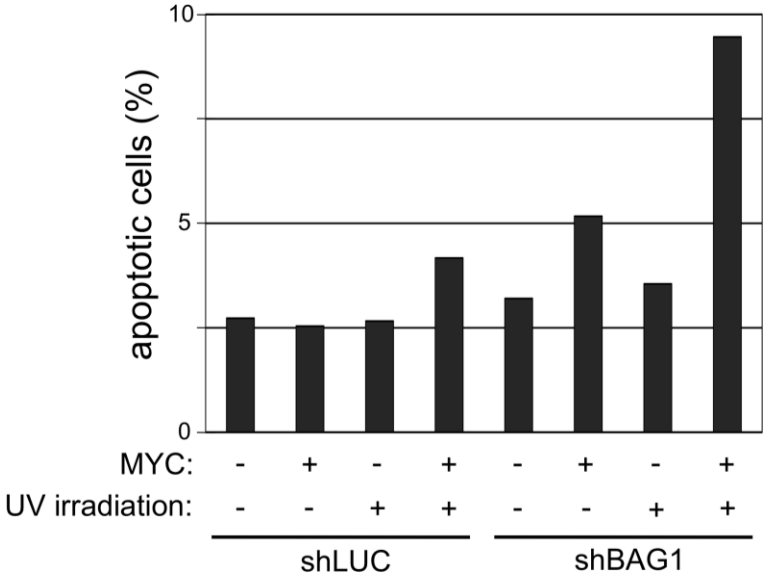


Figure 2 BAG1 depletion sensitizes U2OS osteosarcoma cells to DNA damage, when MYC is activated. U20S MYC/ER cells expressing shRNA targeting luciferase or BAG1 were treated with 5 nM 4-OHT for 24 hours. Cells were then irradiated with 7.5 J/m^2 UV for 18 hours and harvested for apoptosis.

Research Project 3: Project Title and Purpose

Stat5 Interaction with Androgen Receptor in Growth Promotion of Prostate Cancer - This project may validate Stat5a/b as a novel therapeutic target for androgen-independent prostate cancer. The outcome of the studies will increase the knowledge about the molecular mechanisms underlying androgen-independent growth of prostate cancer. The proposed studies may validate active Stat5a/b in prostate cancer as a biomarker for poor response to androgen deprivation therapy. This project has a high impact on prostate cancer patients and for prostate cancer research. This is because there are currently no effective pharmacological therapies for hormone-refractory prostate cancer. Stat5a/b may provide a molecular target for new pharmacological therapy development. Specifically, pharmacological inhibition of Stat5a/b can be achieved by small-molecule inhibitors for dimerization or DNA binding of Stat5a/b.

Anticipated Duration of Project

1/1/2007 – 12/31/2010

Project Overview

This project aims to determine whether transcription factor Stat5a/b promotes androgen-independent growth of prostate cancer through functional interaction with androgen receptor (AR). The rationale behind this proposal centers on the researcher's finding of a critical role of Stat5a/b for cell viability of human prostate cancer cells. Also, inhibition of Stat5a/b blocks prostate xenograft tumor growth in mice. Androgen-receptor signaling, in turn, is known to be necessary in prostate cancer development, and androgen receptor signaling promotes proliferation and survival of prostate cancer cells.

The researcher hypothesizes that Stat5a/b increases transcriptional activity of AR, and that AR stimulates transcriptional activity of Stat5a/b in prostate cancer cells. The researcher proposes that the functional interaction of Stat5a/b and AR promotes prostate cancer cell viability *in vitro* and prostate tumor growth *in vivo.* The researcher will test these hypotheses in human prostate cancer cells in culture and in human prostate cancer tumor models in nude mice using the researcher's special molecular tools to manipulate Stat5a/b and that active Stat5a/b predicts poor response of primary prostate cancer to androgen deprivation therapy. The hypothesis of this project is supported by strong preliminary data. The researcher's new data support a novel concept that Stat5a/b and androgen receptor functionally synergize in prostate cancer cells. Specifically, the researcher's preliminary data indicate that: 1) Stat5a/b promotes transcriptional activity of androgen receptor, and 2) androgen receptor increases transcriptional activity of Stat5a/b. No previous published work exists on functional interaction of Stat5a/b and androgen receptor in prostate cancer.

The project is innovative since the concept of functional interaction between Stat5a/b and AR in promotion of androgen-independent growth of human prostate cancer is entirely novel. Specifically, the researcher proposes that transcription factor Stat5a/b is a co-activator for AR in

human prostate cancer cells. These concepts are new and no published work exists on these topics.

Principal Investigator

Marja T. Nevalainen, MD, PhD
Thomas Jefferson University
Bluemle Life Science Bldg, Suite 309
233 South 10th Street
Philadelphia, PA 19107

Other Participating Researchers

Feng Shen, PhD - employed by Thomas Jefferson University

Expected Research Outcomes and Benefits

The researcher has recently established transcription factor Stat5a/b as a critical protein for survival of human prostate cancer cells. Stat5a/b is constitutively activated in human prostate cancer but not in adjacent normal epithelium, and activation of Stat5a/b is associated with high histological grade of prostate cancer. The researcher has shown that active Stat5a/b in primary prostate cancers predicts early disease recurrence. The researcher's preliminary data show that Stat5a/b is highly active in the majority of recurrent clinical human prostate cancers.

This project is important for several reasons: 1) The outcome of the studies may validate Stat5a/b as a novel therapeutic target for androgen-independent prostate cancer. 2) The outcome of the studies will increase the knowledge about the molecular mechanisms underlying androgen-independent growth of prostate cancer. This project has high impact on prostate cancer patients and for prostate cancer research. This is because there are currently no effective pharmacological therapies for hormone-refractory prostate cancer. Stat5a/b may provide a molecular target for new pharmacological therapy development. Specifically, pharmacological inhibition of Stat5a/b can be achieved by small-molecule inhibitors for dimerization or DNA binding of Stat5a/b.

Summary of Research Completed

In the current funding period, additional molecular mechanisms underlying the transcriptional synergy between Stat5a/b and androgen receptor in human prostate cancer cells were identified.

<u>Prl (prolactin) stimulation is not able to bring androgen receptor (AR) to the nucleus of prostate cancer cells without the presence of Stat5a/b</u>. In addition, DHT (dihydrotestosterone)-stimulation is not able to bring Stat5a/b to the nucleus without the presence of AR. Finally, AR does not translocate to the nucleus in response to Prl-stimulation in Stat5-expressing prostate cancer cells in the absence of Prl-receptor.

To investigate the effect of AR on nuclear translocation of Stat5a/b, double immunostaining of Stat5a/b and AR by indirect immunofluorescence in prostate cancer cells was established. Stat5a, AR and PrlR were expressed in PC-3 cells using replication-deficient adenovirus as an expression vector, as previously demonstrated. After serum starvation for 12 h, the cells were stimulated with hPrl (10 nM) for 30 min and/or DHT (1 nM) for 60 min.

In the absence of added hPrl, Stat5a was predominantly found in the cytoplasm of PC-3 cells transfected with PrlR and Stat5a (Fig. 1. (i)). As expected, addition of hPrl (10 nM) led to nuclear translocation of Stat5a at 10 min with maximal immunostaining of Stat5a in the nucleus at 60 min. However, if the cells co-expressed liganded AR in addition to Stat5a and PrlR, stimulation of the cells with Prl resulted in remarkably more intense nuclear immunostaining for Stat5a at 10 min and at 60 min, compared to the cells stimulated with Prl but expressing only Stat5a and PrlR. Importantly, active Stat5 was located in the nuclei in the absence of hPrl-stimulation, if the cells co-expressed liganded AR. These results suggest that nuclear translocation of Stat5a/b is enhanced by liganded AR.

In the absence of added DHT, AR was predominantly found in the cytoplasm of PC-3 cells transfected with PrlR and Stat5a (Fig. 1. (i)). As expected, addition of DHT (1 nM) led to nuclear translocation of AR at 30 min with maximal immunostaining of Stat5a in the nucleus at 60 min. However, if the cells co-expressed active Stat5a/b in addition to AR, stimulation of the cells with DHT resulted in remarkably more intense nuclear immunostaining for AR at 10 min and at 60 min, compared to the cells stimulated with DHT but expressing only AR. Importantly, AR was located in the nuclei in the absence of DHT-stimulation, if the cells co-expressed active Stat5a/b. These results suggest that nuclear translocation of AR is enhanced by active Stat5a/b.

<u>Prl stimulation is not able to bring AR to the nucleus of prostate cancer cells without the presence of Stat5a/b (Fig. 1 (ii))</u>. AR and PrlR were expressed in PC-3 cells using replication-deficient adenovirus as an expression vector. After serum starvation for 12 h, the cells were stimulated with hPrl (10 nM) for 30 min and/or DHT (1 nM) for 60 min. As shown in Fig. 1 (ii*)*, when PC-3 cells not expressing Stat5a/b were stimulated with hPrl, AR does not become translocated to the nucleus.

<u>DHT-stimulation is not able to bring Stat5a/b to the nucleus without the presence of AR (Fig. 1 (iii))</u>. PC-3 cells were infected with adenoviruses expressing PrlR and Stat5a/b, but not AR, serum-starved for 20 h, treated with DHT (1 nM) for 60 min. As shown in Fig. 1 (iii*)*, when PC-3 cells not expressing AR were stimulated with DHT, Stat5a/b does not become translocated to the nucleus.

<u>AR does not translocate to the nucleus in response to Prl-stimulation in Stat5-expressing prostate cancer cells in the absence of Prl-receptor (Fig. 1(iv)</u>. PC-3 cells were infected with adenoviruses expressing AR and Stat5a/b, serum-starved for 20 h, treated with hPrl (10 nM) for 30 min. As shown in Fig. 1 (iv*)*, when PC-3 cells not expressing PrlR were stimulated with hPrl, Stat5a/b nor AR do not become translocated to the nucleus.

To further investigate the molecular mechanisms underlying the mutually stimulatory interaction between Stat5a/b and AR in prostate cancer cells, the physical interaction between Stat5a/b and AR was evaluated in a greater detail. Truncation mutants of Stat5a/b were generated by standard cloning techniques (Fig. 2). Full length (FL), N-terminal domain deleted (NTD), N-terminal domain and coiled-coiled domain deleted (NTCCD), SH2 and transactivation domains deleted (N-ter) and N-terminal domain, coiled-coiled domain and part of the DNA-binding domain deleted (C-ter) were all Flag-tagged. PC-3 cells were co-transfected with prolactin receptor (PrlR), Myc-AR, the truncation mutants of Stat5a, as indicated, serum-starved for 20 h, and then stimulated with hPrl (10 nM) and the androgen (DHT) (1 nM) for 16 h. AR was immunoprecipitated from the cell lysates by anti-Myc mAb and blotted with a mAb recognizing the Flag-tagged Stat5a/b (anti-Flag mAb). Stat5a (94 kDa) and Stat5b (92 kDa) formed a complex with AR in prostate cancer cells (Fig. 2, Lane 1). NTD, NTCCD and N-ter all immunoprecipitated with the AR while C-ter did not. These results indicate that the DNA binding domain of Stat5 mediates the physical interaction with the AR in human prostate cancer cells.

Next, truncation mutants of the AR were generated by standard cloning techniques (Fig. 3). Full length (FL), Activation domain (1-565), DNA-binding domain and the ligand binding domain (566-919), ligand-binding domain deleted (1-670) and ligand-binding domain and DNA-binding domain deleted (1-608) were all Myc-tagged. PC-3 cells were co-transfected with prolactin receptor (PrlR), Myc-AR truncation mutants and Flag-tagged Stat5a, as indicated, serum-starved for 20 h, and then stimulated with hPrl (10 nM) and the androgen (DHT) (1 nM) for 16 h. Stat5a was immunoprecipitated from the cell lysates by anti-Flag mAb and blotted with a mAb recognizing the Myc-tagged AR (anti-Myc mAb). The results of these experiments indicate that the N-terminal domain of the AR (1-565) is the domain mediating the physical interaction of AR with Stat5a/b.

<u>To identify Stat5-regulated genes in human prostate cancer</u>, Stat5 expression was inhibited 48 h prior to the gene expression analysis by Stat5a/b siRNA using scramble siRNA as control (Ctrl) (Fig.4) in DU145 cells. Overall, 777 genes were differentially expressed between Ctrl siRNA and Stat5 siRNA groups in DU145 cells using false-discovery rate (FDR) < 0.01 on the full dataset of 15,992 genes. To define functional groups within the Stat5-regulated genes, we used the descriptions from the Gene Ontology annotations as our tool. The Gene Ontology is constructed in a hierarchical manner with categories corresponding to each gene ontology identifier (GO ID) being potentially subdivided into more precise sub-categories, each with its own GO ID. “The metastasis group” of genes was defined as a set of 88 GO IDs and their subcategories (a total of 1975 GO IDs) corresponding to genes encoding proteins related to metastases processes. This resulted in 3,234 unique metastases related genes out of 15,992 unique annotated genes on the chip. The apoptosis-related genes as well as proliferation-related genes were defined similarly. Out of the 777 genes differentially expressed between Ctrl siRNA and Stat5 siRNA groups in DU145 prostate cancer cells, 163 (21%) were from the “metastases group” of genes, 30 (3.9%) were from the “apoptosis group” of genes and 61 (7.9%) were from “the proliferation group” of genes (Fig. 3).

A heatmap based on a hierarchical clustering of the expression values of 50 most differentially expressed genes (as determined by the smallest p-values) related to metastases processes (Fig. 4).

Red represents higher expression and green represents lower expression in the heatmap, and the genes (rows) are re-ordered based on hierarchical clustering using the correlation metric. Cytokines and cytokine receptors (interferon alpha, beta and omega receptor-1, interleukin 31 receptor A and erythropoietin), VEGF-C, collagen types VI and XII were identified among the metastases-related genes regulated by Stat5 in prostate cancer cells. In summary, the majority of Stat5-regulated genes in DU145 prostate cancer cells were metastases-related.

Fig. 1

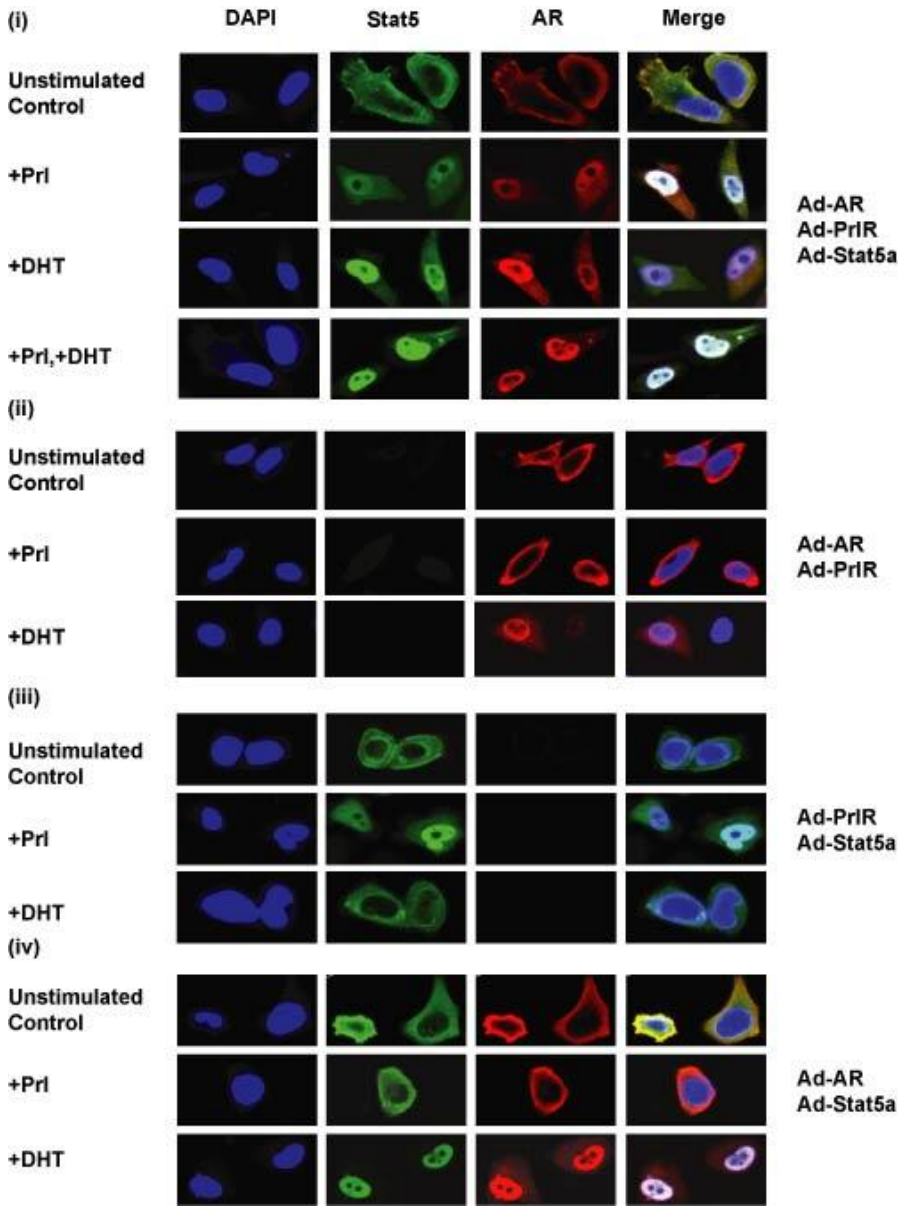


Fig. 2

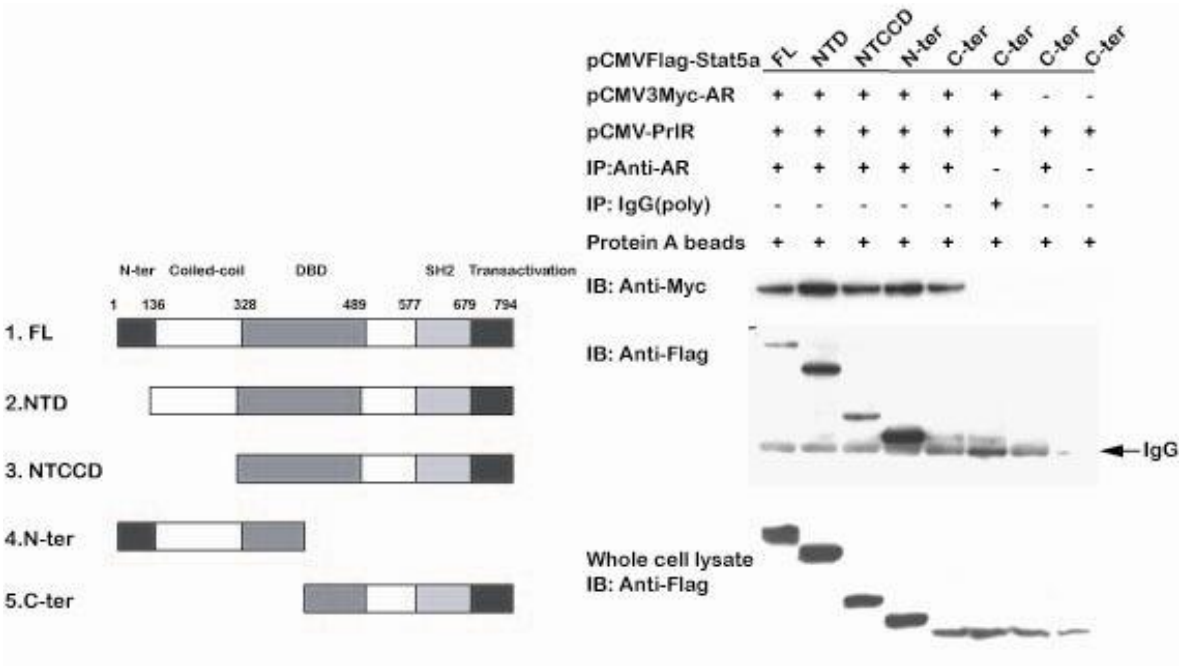

pCMVFlag-Stat5a FL NTD NTCCD N-ter C-ter C-ter C-ter C-ter
pCMV3Myc-AR + + + + + + - -
pCMV-PrIR + + + + + + + +
IP:Anti-AR + + + + + - + -
IP: IgG(poly) - - - - - + - -
Protein A beads + + + + + + + +
N-ter Coiled-coil DBD SH2 Transactivation
1 136 328 489 577 679 794
1. FL
2.NTD
3. NTCCD
4.N-ter
5.C-ter
IB: Anti-Myc
IB: Anti-Flag
IgG
Whole cell lysate
IB: Anti-Flag

Fig. 3

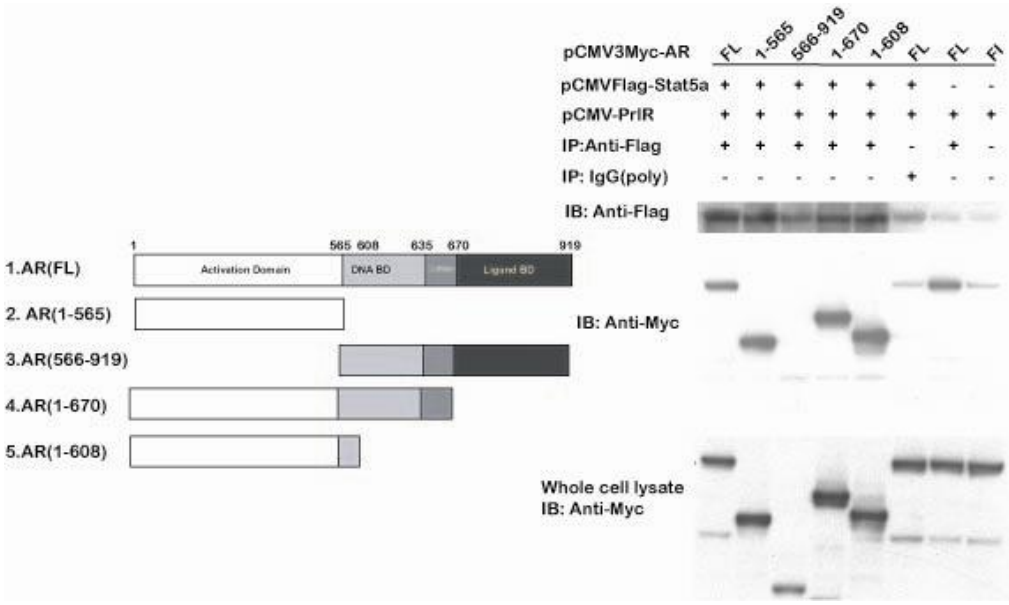
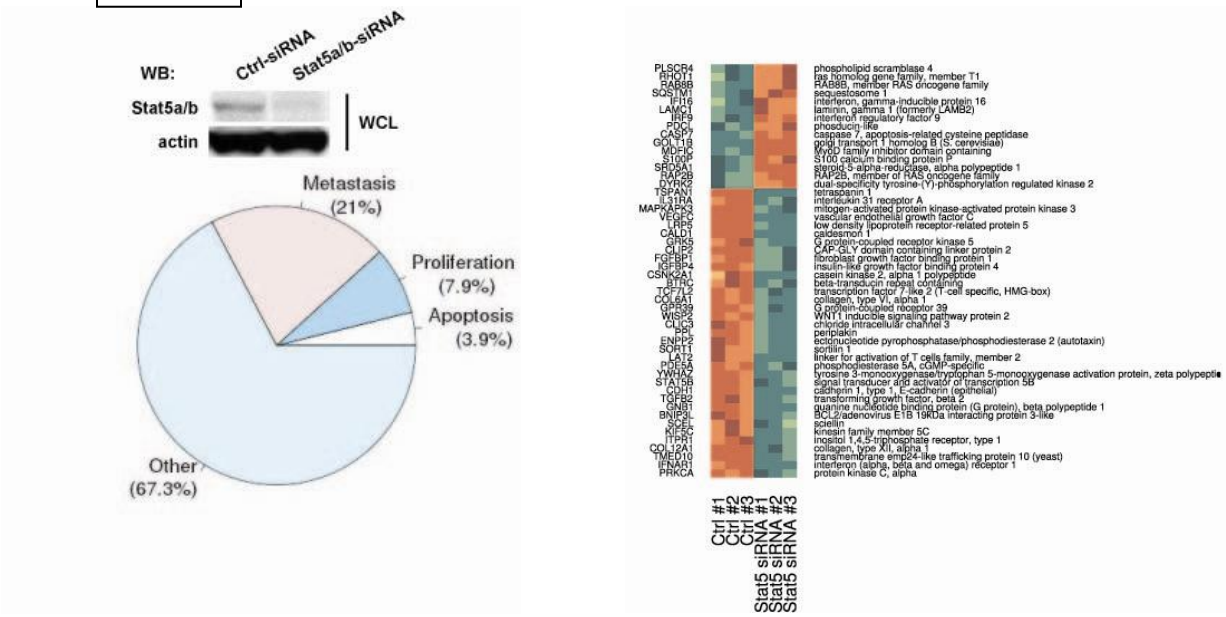

pCMV3Myc-AR FL 1-565 566-919 1-670 1-608 FL FL Fl
pCMVFlag-Stat5a + + + + + + - -
pCMV-PrIR + + + + + + + +
IP:Anti-Flag + + + + + - + -
IP: IgG(poly) - - - - - + - -
IB: Anti-Flag
1 565 608 635 670 919
1.AR(FL)
Activation Domain
DNA BD
Ligand BD
2. AR(1-565)
3.AR(566-919)
4.AR(1-670)
5.AR(1-608)
IB: Anti-Myc
Whole cell lysate
IB: Anti-Myc

Fig. 4



Research Project 4: Project Title and Purpose

Multidisciplinary Systems Approach to Predicting Response to Combination Therapy - This project is focused on developing the mathematical and experimental framework to determine the molecular mechanisms for the efficacy of cancer therapies based upon the combination of existing drugs. This model is important for not only predicting whether two drugs in combination will have synergy, but will also be used for determining the protein biomarkers of response for patients. The aim of this project is to find a mathematical model based on fuzzy logic that could accurately predict a patient's response to drugs in a clinical setting as it would facilitate the individualization of patient treatment and help to find some prognostic and predictive markers of breast cancer.

Anticipated Duration of Project

1/1/2007 - 12/31/2010

Project Overview

Objective: to develop a mathematical model that will accurately predict the response of chemotherapeutic drugs alone and in combination in breast cancer cell lines and xenograft mouse models.

Develop a fuzzy-logic based signaling model that describes the complex interactions of the altered pathways due to combination therapy using HDAC inhibitors and anti HER2/neu therapies. The researcher will use a combination of these drugs at different doses as well as siRNA to perturb the molecular pathways in the SK-BR-3 cell line. Protein-protein interactions

will be determined by co-immunoprecipitation and mass spectrometry and the researcher will utilize a combination of proteomics techniques to determine changes in protein expression. The researcher will pay particular attention to the differential expression of the phosphorylated proteins. Recently developed computational techniques will be used to describe signaling pathways.

Specific Aim 1 Define response as the phenotypic change in cell death and growth. Determine response upon treatment of the HDAC inhibitors TSA, valproic acid, and MS-275 in combination with anti-HER2/neu therapy in the SK-BR3 breast cancer cell line.

Specific Aim 2 Determine global changes in protein expression resulting from drug treatments described in Specific Aim 1.

Specific Aim 3A Use the data from Specific Aims 1-2 to develop a fuzzy logic based signaling model of cellular response; 3B Use the model developed in Specific Aim 3A to identify key interacting proteins of interest, modify their expression using siRNA, and determine response to drugs to validate and refine the model.

Specific Aim 4 Test the protein interaction model developed using a mouse xenograft model. Determine the differentially expressed proteins in the tumors and serum during treatment.

Principal Investigator

Andrew A. Quong, PhD
Thomas Jefferson University
815 Bluemle Life Sciences Building
233 South 10th Street
Philadelphia, PA 19107

Other Participating Researchers

Wen-Shuz Yeow, PhD - employed by Thomas Jefferson University

Expected Research Outcomes and Benefits

The proposed research aims at establishing a research program aimed at the rational design of combination therapy for breast cancer. Although this program is initially focused on breast cancer, the computational and experimental tools that will be developed will be useful for describing signaling pathways underlying not only other diseases but cellular processes in general. In the long run, the benefits of having this toolset are the ability to examine the molecular signature of a patient's disease and help with the design of an individualized therapy. As part of this, the model would also provide a molecular signature of response that can be used to more quickly monitor a patient to see if they are responding to the treatment. Such an approach is needed in cancer therapy where commonly used agents are ineffective in many patients.

Summary of Research Completed

During the current reporting period we have made progress in developing computational tools, which utilize molecular data such as gene or protein expression, and cellular phenotype data,to produce a comprehensive model that relates the molecular pathways to the phenotype. Identification of regulatory molecules in biological signaling pathways is critical for understanding cellular behavior. Given the complexity of the transcriptional gene network, the relationship between molecular expression and phenotype is difficult to determine using reductionist experimental methods. Computational models provide the means to characterize complex regulatory mechanisms in the context of gene networks, and to dissect the relationship between molecular expression and phenotype. This knowledge enables a guided search to identify perturbations to the system that influence cellular behavior. Our present study describes an approach incorporating gene expression values and quantitative measures of cellular phenotype, as interconnected nodes in a network model of budding in Saccharomyces cerevisiae.

A simplified model of the cell cycle in *Saccharomyces cerevisiae* was generated using a data-driven model to identify putative gene-gene and gene-phenotype interactions on the basis of gene expression microarray time series data. The set of seventeen genes included in the model have been identified as central regulators of cell cycle progression: cyclins (*CLN1-3*, *CLB1-2*, *CLB5-6*), cyclin-dependent kinases (*CDC28*), transcription factors (*MCM1*, *MBP1*, *Swi4*, *Swi5*, *Swi6*), as well as cell cycle activating (*CDC6*) and inhibiting (*SIC1*, *CDH1*, *CDC20*) genes. In order to describe the effect of gene expression on phenotype, we also included data describing the fraction of budding cells observed throughout the time series. Budding begins at the G1/S phase transition and continues through mitosis. As such, the fraction of budding cells is a morphologically observable indicator of the number of cells in a population that is actively undergoing cell division.

A fuzzy logic framework was implemented to derive regulatory relationships between genes and the fraction of budding cells from data collected at different time points throughout the cell cycle. Gene expression data used for model generation and evaluation were generated by synchronizing cells by one of three different methods: α factor, elutriation and *CDC15*-based synchronization. Data obtained from the *CDC15*-based synchronization method were used in this study, as this is the only dataset for which phenotype data was available. Within the fuzzy logic framework, an exhaustive search was used to identify the best-fitting relationships from the set of all potential interactions (i.e. gene A interacts with gene B) and all possible fuzzy rules that relate the state of input nodes to output nodes. The coefficient of determination (COD) was used to calculate the fitness of the data generated models by comparing the predicted gene expression values or fraction of budding cells with experimentally observed measurements (Table 1).

High values indicate that the identified models accurately predict the magnitude of expression values over the measured time points. Figures 1A and 1B depict the best-fit models for each node and the integrated network model respectively. Blue edges indicate activation, in which input and output are positively related. Red edges indicate inhibition, in which input and output are negatively related. Dashed edges indicate biphasic interactions (i.e. {1,3,1}; {3,1,3}; etc.). Figure 2 is a plot of the observed and predicted fraction of budding cells over time derived from the best-fit phenotype model. The legend indicating the phase of the cell cycle throughout the

time series was generated by observing the expression values of genes that are associated with cell cycle progression (e.g. G1, *CLN2*; S, *HTA1*; G2, *CLB4*; M, *Swi5;* M/G1, *ASH1*). The maximum and slope of the predicted fraction of budding cells is highly similar to the observed fraction of budding cells throughout the time series. Noticeably, the troughs of the predicted fraction of budding cells do not reach the minimum observed value. High correlation between predicted values and observed measurements suggest that the model built on simplified relationships in fuzzy space sufficiently captures the complex, multivariate relationships between variables.

A number of the genes included in the cell cycle model are known to be functionally homologous. CLN1 and CLN2, as well as CLB1/CLB2 and CLB5/CLB6, have a high degree of sequence similarity (57 percent identity). Additionally, yeast strains with deletions of either CLN1 or CLN2 remain viable and divide normally, suggesting that these proteins are interchangeable parts in the cell cycle machinery. In order to accurately model redundancy, we manually replaced the nodes representing these functionally homologous genes with a node corresponding to the maximum expression value of the homologous pair. The maximum operator is a common representation of a logical union and a reasonable approximation of functionally homologous genes in the fuzzy logic framework described above. Comparison of model fitness with and without the homologous gene representation supports that this is an appropriate method to account for functional homology.

Gene knockout is a powerful method for characterization of gene function and identification of essential genes necessary for cell viability. Gene and phenotype-specific fuzzy models were integrated into a composite model to investigate the dynamic behavior of the system as an interdependent network and predict the identity of essential genes *in silico* (Figure 1B). Physiologically relevant initial conditions for each node were derived from expression values at each time point and the observed fraction of budding cells at each time point in the *CDC15* dataset. The state of each node was iteratively updated based on the state of its immediate input genes and the inferred fuzzy rules describing the relationships between nodes. In effect, expression data from each time point serves as a separate set of initial conditions to investigate the dynamic behavior of the model. Iteration continued until the state of the model remained relatively constant between iterations.

Figure 2 shows the predicted fraction of budding cells at each time point at convergence of the dynamic model. The model converges to one of several local minima associated with high, medium or low fraction of budding cells. Importantly, the timing of the transitions between local minima is similar to the observed changes in the fraction of budding cells. Using this dynamic model, we simulated the effect of gene knockout on the fraction of budding cells by fixing the state of the corresponding node to a low expression profile. Viability of *in silico* gene knockout models was determined by observing the predicted fraction of budding cells. The phenotype of viable models closely approximates the oscillatory behavior of the experimentally observed fraction of budding cells over time (Figure 2).

Inviable models were identified by convergence to a constant fraction of budding cells over time, indicating cell cycle arrest. The predicted viability of *in silico* gene knockout models was compared to experimental observations from a systematic investigation of single gene deletions

in yeast under normal growth conditions. In this manner, in the absence of a true validation set of independent gene expression and phenotype data, independent experimental observations were used to evaluate the validity of our model. Table 2 summarizes the correlation between predicted and observed viability. Overall, fourteen of seventeen genes were correctly predicted. Accurate identification of essential genes with simulation of gene knockouts *in silico* indicates the validity of the inferred yeast cell cycle model and demonstrates the value of fuzzy logic models in hypothesis generation.

Double gene knockouts can be used to discover genes with physical or genetic interactions that are essential for proper cellular function. Synthetic lethals are observed when a double knockout induces lethality and both single gene knockouts are viable. Conversely, synthetic rescue is observed when a double knockout of an essential and non-essential gene alleviates the lethality of the essential gene knockout. As an additional challenge to our dynamic yeast cell cycle model, we evaluated the predicted viability of double gene knockouts against experimentally identified synthetic phenotypes. Experimentally identified synthetic lethal and synthetic rescue phenotypes were obtained from publications compiled in the Saccharomyces Genome Database. We limited the analysis to the fourteen genes for which viability is correctly predicted with *in silico* single gene knockouts. Among these genes, thirteen gene pairs have been observed to induce a synthetic lethal phenotype and five gene pairs have been observed to induce a synthetic rescue phenotype. Nine of eighteen total synthetic phenotypes were correctly predicted including seven of thirteen synthetic lethal and two of five synthetic rescue phenotypes. The larger fraction of incorrect predictions from the analysis *in silico* double knockout is likely a result of the size of the model. It can be assumed that many of the redundant mechanisms that preserve cell cycle regulation in yeast are not represented in our seventeen gene model.

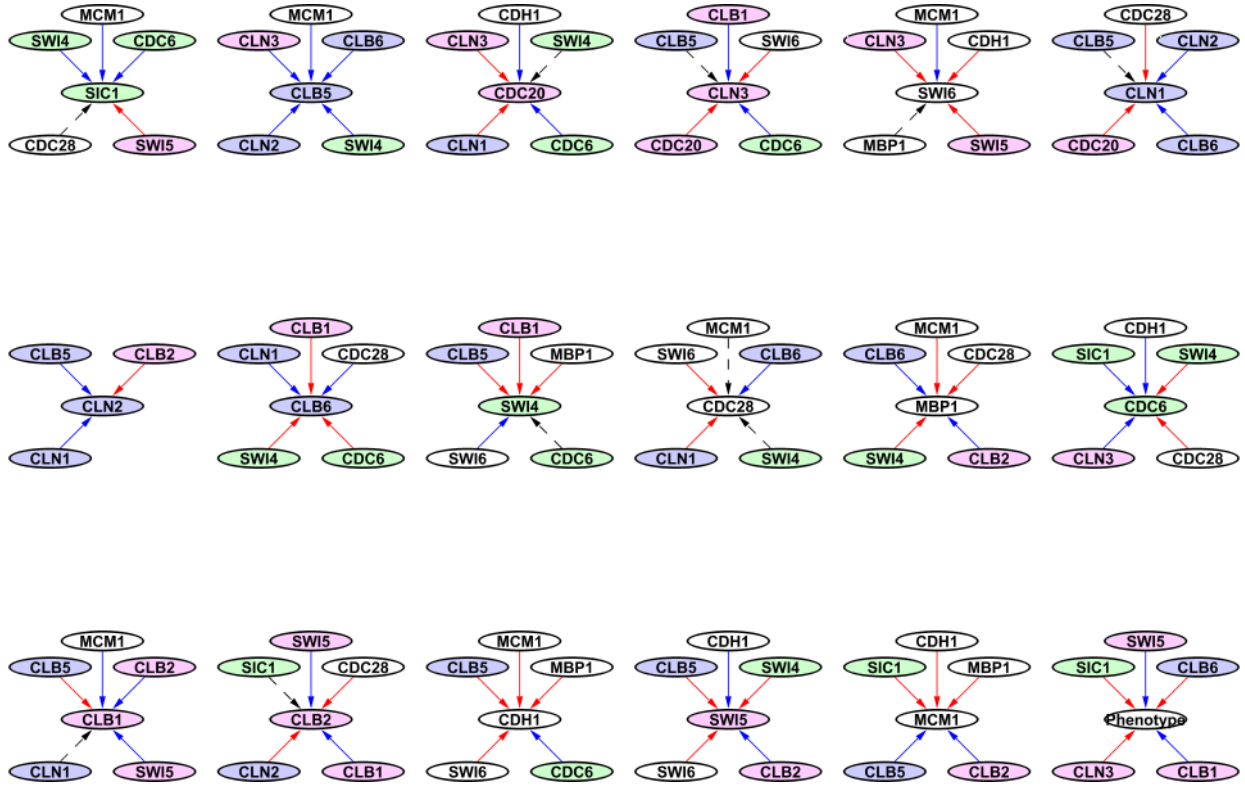
Molecular profiling technologies such as gene expression microarrays have enabled the quantification of molecular abundance at a genome-wide scale in high throughput. As technologies have matured, goals of data analysis have grown from the identification of conserved expression patterns across samples or conditions to the comprehension of gene function in the context of complex regulatory and functional networks. In this study, we have extended a fuzzy-logic based modeling approach to derive a transcriptional network consisting of seventeen genes known to be important for cell cycle regulation in yeast and a network element representing a phenotypic observation (the fraction of budding cells). Both the topology (i.e. interactions) and regulatory traits (e.g. stimulation, inhibition) of relationships between genes and between genes and phenotype are derived from publicly available gene microarray expression data and phenotype data using a bounded exhaustive search of the potential interaction space.

Comparison of inferred gene regulatory interactions with known interactions in the literature provides confidence in the biological relevance of model predictions. Through the analysis of best-fit fuzzy logic signaling models, genes with direct and indirect effects on phenotype expression were identified. In addition, we developed a methodology to predict the effects of gene knockdown on cellular viability. In this manner, the methodology we have developed provides a direct link between computational analysis of molecular profiling data and experimental observations. We envision coupling this computational modeling method with experimentation in an iterative fashion to incrementally build understanding of regulatory

mechanisms that control poorly understood cellular functions. This approach allows for systematic characterization of gene function in the context of disease-related, functional mechanisms that will facilitate rational design of targeted therapies.

Figure 1 - Graphical depictions of best-fit models identified using the fuzzy logic model fitting procedure.

A



B

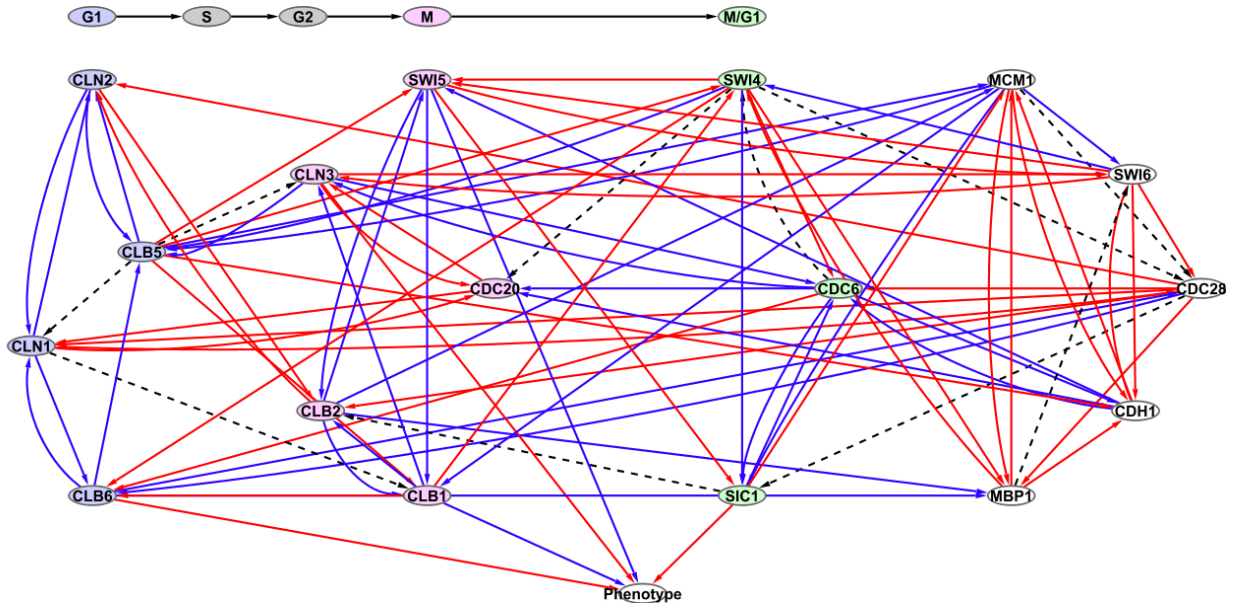


Figure 2 - Observed and predicted fraction of budding cells at different time points in cell cycle progression.

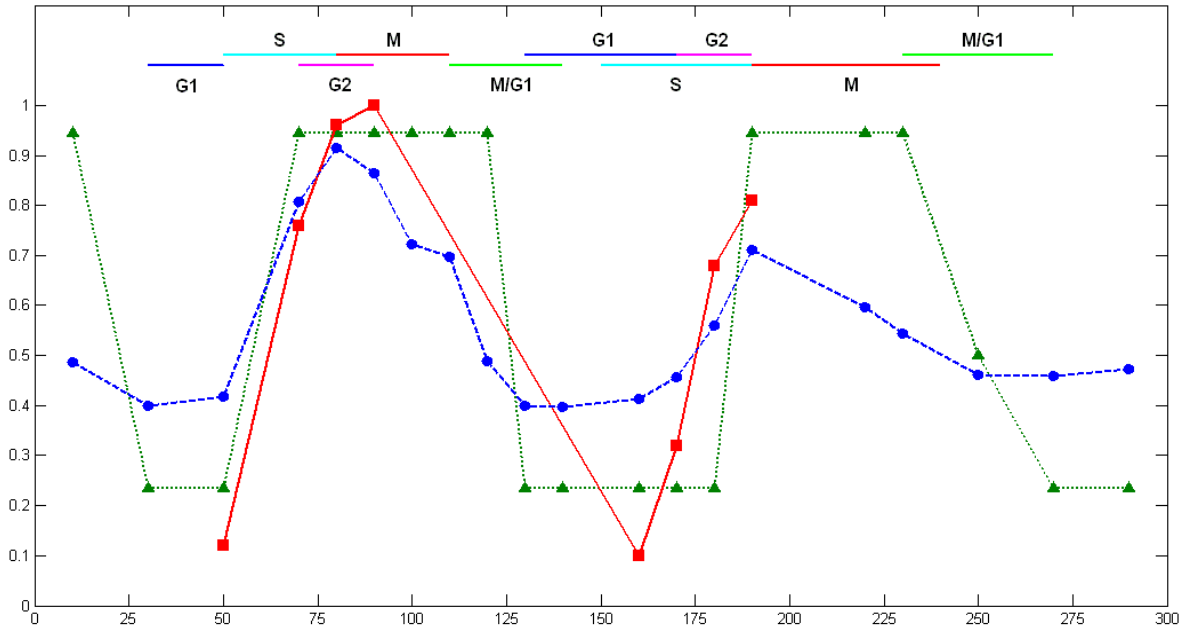


Table 1 – Goodness of fit of predicted vs. observed expression for the best fitting models. The fuzzy models that produce these metrics are displayed in Figure 1.

Node	ORF	COD
SIC1	YLR079W	0.8524
CLB5	YPR120C	0.9484
CDC20	YGL116W	0.8270
CLN3	YAL040C	0.8296
SWI6	YLR182W	0.8128
CLN1	YMR199W	0.9643
CLN2	YPL256C	0.9314
CLB6	YGR109C	0.8431
SWI4	YER111C	0.8659
CDC28	YBR160W	0.6770
MBP1	YDL056W	0.7485
CDC6	YJL194W	0.8485
CLB1	YGR108W	0.9322
CLB2	YPR119W	0.9316
CDH1	YGL003C	0.7956
SWI5	YDR146C	0.9262
MCM1	YMR043W	0.7745
Phenotype	NA	0.7264

Table 2 – In silico gene knockdown models predict the viability of experimental deletion of yeast cell cycle genes.

Node	ORF	Experimental Outcome	Model Prediction
SIC1	YLR079W	Viable	Viable
CLB5	YPR120C	Viable	Viable
CDC20	YGL116W	Inviable	Inviable
CLN3	YAL040C	Viable	Viable
SWI6	YLR182W	Viable	Viable
CLN1	YMR199W	Viable	Viable
CLN2	YPL256C	Viable	Viable
CLB6	YGR109C	Viable	Viable
SWI4	YER111C	Viable	Viable
CDC28	YBR160W	Inviable	Inviable
MBP1	YDL056W	Viable	Viable
CDC6	YJL194W	Inviable	Viable
CLB1	YGR108W	Viable	Viable
CLB2	YPR119W	Viable	Viable
CDH1	YGL003C	Viable	Inviable
SWI5	YDR146C	Viable	Inviable
MCM1	YMR043W	Inviable	Inviable

Research Project 5: Project Title and Purpose

Cyclin D1 Regulation of Cellular Migration and Invasion - These studies aim to further understanding of how cyclin D1 collaborates in cellular invasion, migration and transformation. The abundance of the cyclin D1 gene product is rate-limiting in progression through the G_1 phase of the cell-cycle and is required for transformation by Ras, ErbB2, Src, Stats, Wnt and mutations of the *Apc* gene. The researcher's laboratory has shown that cyclin D1 promotes DNA synthesis and cellular survival. Recent studies from our laboratory using *cyclin D1*$^{-/-}$ mice and cyclin D1 siRNA have demonstrated that cyclin D1 regulates cellular migration and invasion. The proposed studies will determine the mechanism by which cyclin D1 regulates migration in cells and *in vivo*. Several novel technologies developed in this laboratory will be used in the proposed studies to provide high-resolution analysis of the mechanisms governing this process.

Duration of Project

1/1/2007 – 12/31/2009

Project Overview

Clinical observations demonstrated a correlation between cyclin D1 over-expression and cellular metastases in several tumor types. Analyses of cells derived from *cyclin D1*$^{-/-}$ mice showed a defect in cellular migration and reintroduction of cyclin D1 rescued the cellular migratory defect. Scanning EM demonstrated reduced membrane ruffling in *cyclin D1*$^{-/-}$ bone marrow macrophages. Studies of human endothelial cells showed *cyclin D1*$^{-/-}$ endothelial cells are

defective in matrigel invasion. Subsequent studies of *cyclin D1*$^{-/-}$ MEFs and mammary epithelial cells demonstrate a similar defect in cellular migration, which was restored through the reintroduction of cyclin D1. Together, these studies suggest cyclin D1 may contribute to the metastatic cellular phenotype observed in cyclin D1 over-expressing cells. The molecular mechanisms by which cyclin D1 regulates cellular migration are not known but represent a useful potential target for inhibiting tumor metastases. In the researcher's studies investigating the mechanism by which cyclin D1 regulates migration, the researcher observed the supernatant of *cyclin D1*$^{+/+}$ cells rescued the migration defect of cyclin D1$^{-/-}$ cells.

The researcher has observed increased production of TSP1 in *cyclin D1*$^{-/-}$ supernatant. Immunoneutralizing antibody and TSP1 partially reversed the defect in migration of *cyclin D1*$^{-/-}$ cells. Analysis of the supernatant identified reduced scription of SCF. Finally, in the researcher's studies of *cyclin D1*$^{-/-}$ MEFs, the researcher has observed increased expression of ACRP30 an inhibitor of cellular migration. TSP1 is known to inhibit cell migration. Aim 1: Identification of cyclin D1-mediated gene expression profile in promoting cellular migration. High-throughput microarray analysis will be used to identify gene expression profile. Aim 2: To define the mechanism by which cyclin D1-regulated target genes governing migration. Cell migration, chemotaxis in a modified Boyden chamber, adhesion on fibronectin coated plates, interference reflection microscopy (Irm), PMD, wound healing, and Rho GTPase assays (Rac, Rho, Cdc42) and metastasis assays are well established in the researcher's laboratory. Aim 3: To determine the role of cyclin D1-mediated secretory factors in cellular migration. The researcher will carry out cytokine antibody array analysis to screen the cyclin D1-mediated cytokines. Aim 4: To determine the requirement of pRb in cyclin D1 regulation of secreted proteins and metabolism. *Cyclin D1* and *Rb* gene knockout mice required for the proposed experiments are available to perform assays listed in Aim 2.

Principal Investigator

Chenguang Wang, PhD
Thomas Jefferson University
Room 1032A BLSB
233 South 10th Street
Philadelphia, PA 19107

Other Participating Researchers

Jie Zhou, PhD - employed by Thomas Jefferson University

Expected Research Outcomes and Benefits

Determining the mechanisms by which cyclin D1 promotes cellular migration and invasion may provide important insights into the mechanisms governing tumor growth/metastasis and potential avenues for new therapeutic interventions for metastatic cancer.

Summary of Research Completed

1. We have established inducible cyclin D1 wild-type, kinase-deficient mutant (KE), and alternative splicing form (cyclin D1b) transgenic mice.

Cyclin D1a, cyclin D1KE and cyclin D1b transgenics have been recently generated under control of the tetracycline enhancer are available to my laboratory (Figure 1). Three founder lines for cyclin D1a and KE each and 14 founder lines for cyclin D1b have been established (referred to *D1aTg, KETg, D1bTg* respectively). The transgenic lines have be characterized by PCR and analyzed for tetracycline-inducible expression *in vivo*. Animals received tetracycline in the drinking water as previously described to assess transgene induction and after 1 week, biopsy of the mammary gland was performed. This material was analyzed for transgene expression and expression of the cyclin D1 transgene was determined by FLAG immunohistochemistry and Western blot. There was an approximately 30-to 50-fold acceleration of the tumorigenesis rate using the Tet-Op driven transgene expression system compared with the same transgene driven by the MMTV promoter.

These transgenic lines were crossed to MMTV-*rtTA* (a generous gift from Dr. L. Chodosh) to generate the cyclin D1-TetOp transgenics (*D1aTg/TetOp, KE/TetOp,* and *D1b/TetOp*) (Figure 2). The expression of cyclin D1a and KE under control of the tetracycline operon in the mammary epithelium has been tested (Figure 1C and D). The cyclin D1-TetOp lines will be crossed to *cyclin D1*$^{-/-}$ to generate *cyclin D1*$^{-/-}$*/D1a/TetOp* and *cyclin D1*$^{-/-}$*/KE/TetOp* lines. This system conveys more robust induction of transgene expression than MMTV-*cyclin D1* (Figure 1 and data not shown).

Mammary gland targeted tetracycline-inducible transgenics will be used to determine the requirement for cyclin D1 in tumor commitment. Cyclin D1 binds several proteins that serve as substrates for cyclin D1. Mutants of the cyclin D1 transgenics will be used to determine the domains of cyclin D1 required for mammary tumorigenesis.

2. Comparative study among cyclin D1a, KE, and cyclin D1b.

We and others previously identified the distinct functions represented by the kinase-defective mutant KE, and cyclin D1 alternative splicing form cyclin D1b *in vitro*. Cyclin D1 conveys cdk-dependent and independent functions. The cdk-independent functions of cyclin D1 include binding and regulating nuclear hormone receptors and their coactivators. This laboratory has shown that cyclin D1 regulates the function of several proteins that in turn regulate tumorigenesis, ERa, AR and PPARg, BRCA1, p300 and HDACs. The mammary gland expansion during pregnancy requires cyclin D1 function, but not the kinase-dependent function and pRb pathway. ErbB2 requires cyclin D1 to induce mammary tumorigenesis, and the mice expressing kinase-deficient cyclin D1 were resistant to ErbB-2-induced breast cancers.The proposed studies are designed to formally establish the role of cyclin D1 kinase function in mammary tumorigenesis *in vivo* using inducible transgenic mice.

Cyclin D1b has an enhanced capability to transform fibroblasts and it associates with poor prognosis in older patients with ER-negative breast cancer. Our preliminary data, conducted to

understand the enhanced transforming capability of the cyclin D1b isoform, showed that cyclin D1b conveyed enhanced cellular apoptosis and reduced binding to cdk4, $p21^{CIP1}$ and $p27^{KIP1}$. Cyclin D1a promotes cellular migration, requiring residues in the carboxyl terminus and $p27^{KIP1}$ binding. Together these studies suggest the cyclin D1a and cyclin D1b forms have distinct protein binding partners and distinct biological effects on cultured cells. The role of cyclin D1b in mammary gland tumorigenesis is not known. In the proposed studies, for the first time, we are able to examine the role of cyclin D1b in mammary tumorigenesis *in vivo* using inducible, mammary gland targeting transgenic mice.

Materials have been prepared from these transgenic lines to conduct further experiments as continuous research as illustrated in Figure 3.

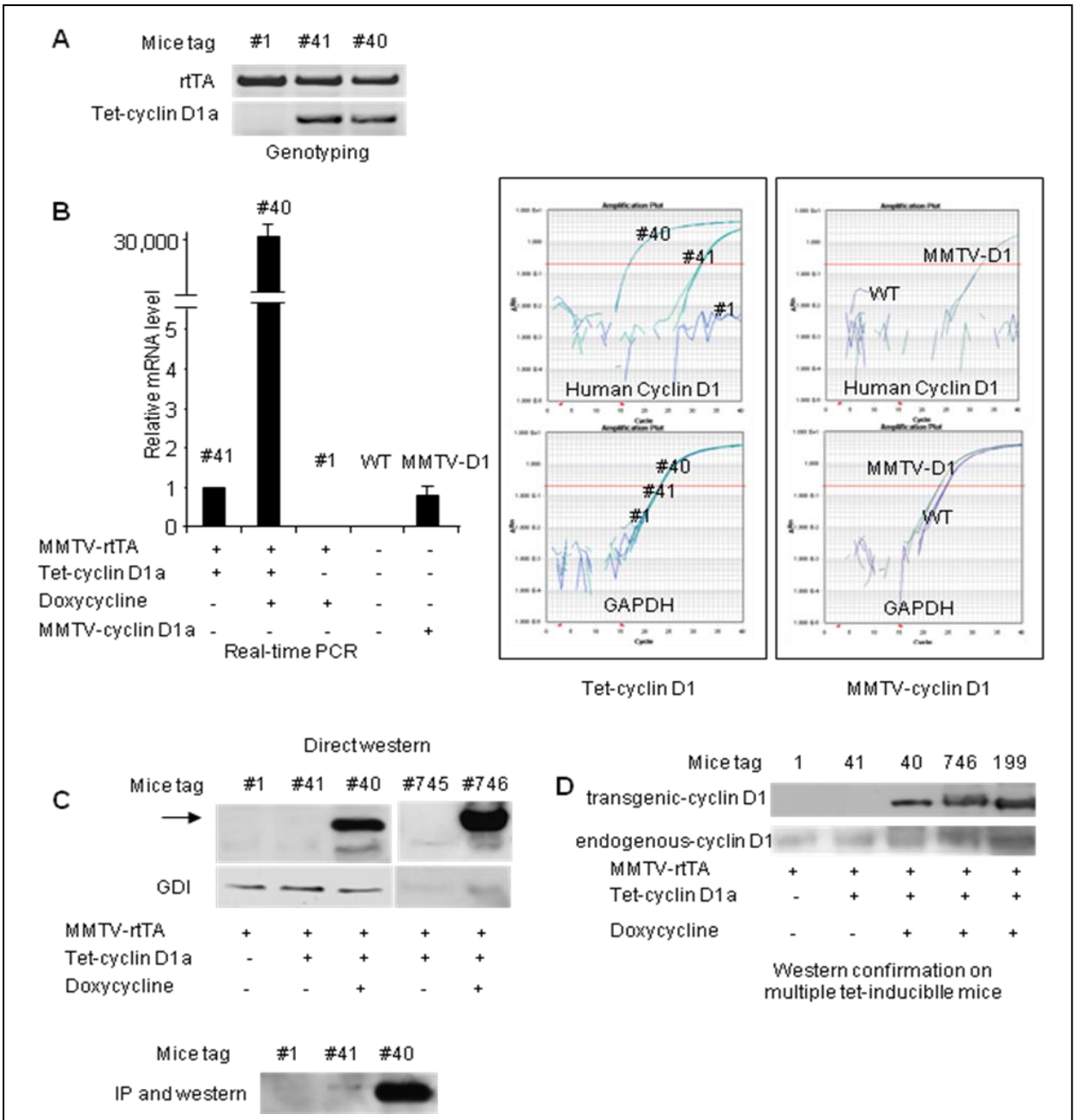


Figure 1. Tetracycline-inducible cyclin D1 transgene expression. A). Genotyping. Mouse line #1 is MMTV-*rtTA* single transgene. #41 and #40 are bitransgenics and mouse #40 was fed with doxycycline supplemental water. B). Real-time PCR analysis of transgene expression showing robust of induction of cyclin D1 transcript. C). Western blot detection of transgene expression at protein level, showing cyclin D1 was induced by doxycycline.

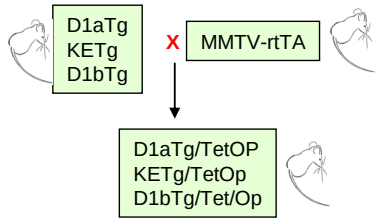


Figure 2. Mating strategy for genarating *D1aTg/TetOp, KE/TetOp* and *D1b/TetOp* mice.

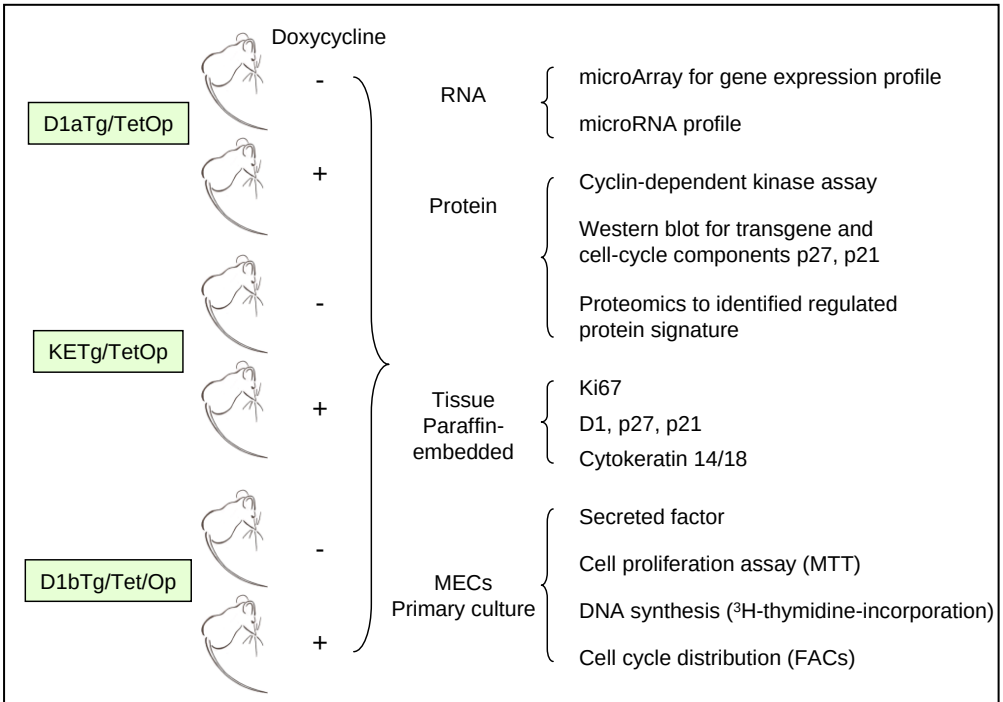


Figure 3. Schematic representation of proposed experiments.