Distribution Agreement

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant Emory University and its agents the non-exclusive license to archive, make accessible, and display my dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this dissertation. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation

_____   _____
Christopher D. Scharer   Date

INTEGRATING GENOMICS AND MOLECULAR BIOLOGY:
IDENTIFYING TRANSCRIPTIONAL TARGETS FOR THE PROSTATE
CANCER ONCOGENE *SOX4* AND EVALUATING THE EFFICACY OF
AURORA KINASE INHIBITION IN CHEMOTHERAPY-RESISTANT
OVARIAN CANCER

By

Christopher D. Scharer
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

Carlos S. Moreno, Ph.D.
Adviser

Jeremy M. Boss, Ph.D.
Committee Member

Tamara Caspary, Ph.D.
Committee Member

Kenneth H. Moberg, Ph.D.
Committee Member

Paula M. Vertino, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

Date

INTEGRATING GENOMICS AND MOLECULAR BIOLOGY:
IDENTIFYING TRANSCRIPTIONAL TARGETS FOR THE PROSTATE
CANCER ONCOGENE *SOX4* AND EVALUATING THE EFFICACY OF
AURORA KINASE INHIBITION IN CHEMOTHERAPY-RESISTANT
OVARIAN CANCER

By

Christopher D. Scharer

B.S., Emory University, 2004

Adviser: Carlos S. Moreno, Ph.D.

An Abstract of A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2009

Abstract

INTEGRATING GENOMICS AND MOLECULAR BIOLOGY: IDENTIFYING
TRANSCRIPTIONAL TARGETS FOR THE PROSTATE CANCER ONCOGENE *SOX4* AND
EVALUATING THE EFFICACY OF AURORA KINASE INHIBITION IN
CHEMOTHERAPY-RESISTANT OVARIAN CANCER

By

Christopher D. Scharer

The advent of genomic microarray technology has transformed molecular biology and

genetics from single gene studies, allowing for the first time, experiments that analyze global

regulatory networks and functions, giving birth to the fields of transcriptomics and genomics.

Here, we have applied DNA expression-microarrays to discover global gene expression changes

in cancer and transfected cell lines, tiling-microarrays to identify genomic binding sites for the

SOX4 DNA binding protein, and unique double stranded DNA microarrays to determine

sequence binding specificities for SOX4. In two independent studies, we built on knowledge

gathered from genomic techniques to evaluate the efficacy of Aurora kinase inhibition in ovarian

cancer and gain novel insights into the role of the transcription factor SOX4 in prostate cancer. In

the first study, we investigated whether the Aurora family inhibitor VE-465 could induce

apoptosis in a taxol-resistant and taxol-sensitive ovarian cancer cell line. We find that at Aurora-

A specific doses, VE-465 synergized with taxol to induce apoptosis in ovarian cancer cell lines.

These studies provide preliminary evidence that Aurora family inhibition may be beneficial in

ovarian cancer patients whose tumors express high levels of Aurora-A. In the second study we

applied two different array technologies to identify, for the first time, 282 high-confidence target

genes for the SOX4 transcription factor. We also determined the sequence binding specificities

for SOX4 using a novel double-stranded DNA microarray, and show that SOX4 binding sites are

enriched in our ChIP-chip peaks. SOX4 influences key signaling pathways at the transmembrane

and transcriptional level, as well as the small RNA pathway. These data provide novel insights

into SOX4's role in development and cancer progression.

INTEGRATING GENOMICS AND MOLECULAR BIOLOGY:
IDENTIFYING TRANSCRIPTIONAL TARGETS FOR THE PROSTATE
CANCER ONCOGENE *SOX4* AND EVALUATING THE EFFICACY OF
AURORA KINASE INHIBITION IN CHEMOTHERAPY-RESISTANT
OVARIAN CANCER

By

Christopher D. Scharer

B.S., Emory University, 2004

Adviser: Carlos S. Moreno, Ph.D.

A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2009

ACKNOWLEDGEMENTS

I would like to thank the members of my thesis committee, Drs. Jeremy Boss, Ken Moberg, Paula Vertino, and Tamara Caspary, as well as other GDBBS faculty for thoughtful input and advice during the course of this research project.

Additionally, I would like to thank the members of the Moreno Lab, past and present, for their friendship, scientific support, and general 'good times' in the lab. In particular I would like to single out Dr. Colleen McCabe for her programming expertise that was instrumental in deciphering the data NimbleGen sent us, and Noelani Anderson who laid the ground work for both of the projects presented here and especially her constant drive to enjoy life. To my fellow graduate students, Veronica Henderson and Yu-heng Lai, thank you for truly becoming my friends and co-workers during the past 5 years, keep up the hard work wherever life leads you.

I especially would like to thank my friends in graduate school and in Whitehead Room 135. Graduate school can be a frustrating experience and you help me come to work with a smile and were there to remind me of life outside the lab. The friendships and experiences I have had during the past 5 years I will never forget.

I would like to thank my advisor Dr. Carlos Moreno. Your passion and drive for science motivated me throughout graduate school. Importantly, you trusted and allowed me a unique freedom to develop my analytical and technical skills, never shying away from new techniques or fields, and were always there to guide me when it was needed.

Finally, I would like to thank my family whose encouragement and constant support have been instrumental in my development as a scientist and a person.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter

I

Aurora Kinases:

Mitotic Regulators and Cancer Agents

Eukaryotic cells have developed stringent cell cycle controls to ensure mitosis consistently occurs error free. Cell cycle checkpoints have evolved to guarantee the inheritance of the correct complement of chromosomes and to ensure that each chromosome is faithfully passed on undamaged. Misregulation or aberrant expression of cell cycle regulatory proteins can lead to genomic instability and contribute to cancer progression. The Aurora family of serine/threonine kinases directly regulates mitosis, controlling critical cellular processes such as the G2-M transition, chromosome condensation, centrosome maturation, the mitotic spindle checkpoint and cytokinesis. The taxane class of chemotherapeutics directly binds and stabilizes microtubules causing cell cycle arrest and apoptosis due to the activation of the mitotic spindle checkpoint, a process regulated by the Aurora kinases. Due to the high incidence of taxane resistance in advanced, recurrent ovarian tumors, and the recent finding that Aurora-A is over expressed in ovarian tumors and could contribute to taxane resistance, we investigated whether a pan-Aurora small molecule inhibitor can synergize with the taxane paclitaxel to cause apoptosis in paclitaxel sensitive and resistant ovarian cancer cells.

1.1 Aurora-A and Family

Eukaryotic cells have evolved an assortment of proteins to accurately regulate

mitosis and ensure each daughter cell receives the full complement of genetic material.

The Aurora family of evolutionarily conserved serine/threonine kinases are critical

eukaryotic regulators of the mitotic (M) phase of the cell cycle. The founding member of

the Aurora family, Aurora-A (AurA) was first identified in *Drosophila melanogaster* in a

screen looking for mutants that exhibited defective spindle-pole assembly (60).

Appropriately named, mutations in AurA caused monopolar spindles that resembled the

northern light displays at each cellular pole. The Aurora family is highly conserved and

Organism	Name	Alternative Names	Localization
Saccharomyces cerevisiae	Ipl1	-	Kinetochore
Schizosaccharomyces pombe	Ark1	-	Kinetochore
Caenorhabditis elegans	Aurora-A	AIR-1	Centrosome
	Aurora-B	AIR-2	Chromosome Passenger
Drosophila melanogaster	Aurora-A	DmAurora	Centrosome
	Aurora-B	IAL	Chromosome Passenger
Xenopus lavies	Aurora-A	Eg2	Centrosome
	Aurora-B	XAIRK2	Chromosome Passenger
Mammals	Aurora-A	Aurora 2, AIRK1, ARK1, BTAK, STK6, STK15, AYK1, IAK1	Centrosome
	Aurora-B	Aurora 1, AIRK2, ARK2, IAL1, AIK2 STK12, AIM1	Chromosome Passenger
	Aurora-C	Aurora 3, AIRK3, AIE2, STK13, AIE1, AIK3	Spindle Pole or chromosome passenger

Table 1: Eukaryotic Aurora kinase family members.

homologs exist in all eukaryotes (**TABLE 1**). Mammalian genomes contain three members of the Aurora family, AurA, B, and C, while *Drosophila*, *Xenopus*, and *C. elegans* contain two members and *Saccharomyces cerevisiae* have one Aurora kinase respectively. Nevertheless, cell-cycle regulatory function has been conserved throughout evolution with mutants showing similar defects in lower and higher eukaryotes (56, 120).

AurA is universally expressed and regulates cell cycle events from late G2 through the completion of M phase. Structurally, AurA contains two domains, a regulatory region at the N-terminus containing the A-box/D-box activating Domain and a C-terminal catalytic domain (**FIGURE 1**). AurA is activated during late G2 phase by the LIM protein Ajuba (76). Ajuba physically binds AurA, inducing conformational changes that result in catalytic activation and autophosphorylation on Threonine 288 of the catalytic domain (76, 134). Activation of AurA is essential for the G2-M transition as
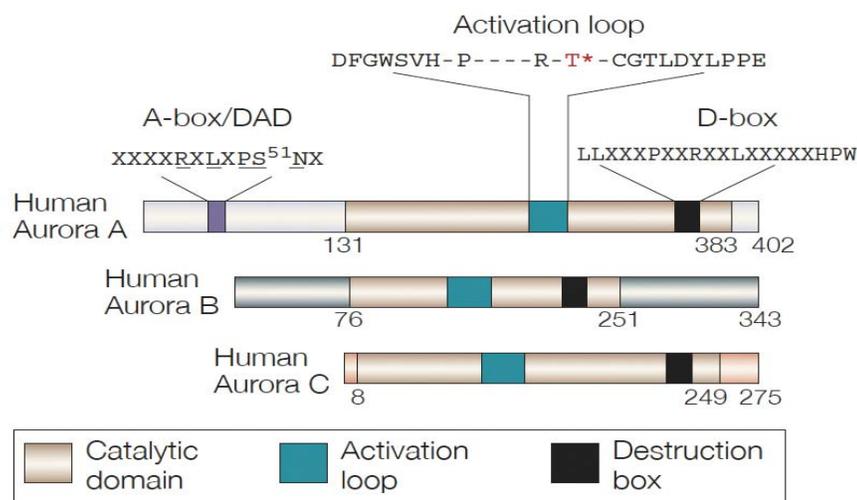
Figure 1: Domain structure of the Human Aurora kinases. The A-box/D-box-activating domain and activation loop are shown for the best characterized Aurora kinase, Aurora A

synchronized HeLa cells depleted of AurA fail to progress into M phase of the cell cycle (76). Following mitotic entry, AurA kinase activity can be further stimulated by Targeting protein for Xenopus 2 (TPX2) (49, 96), which also is an AurA phosphorylation substrate along with other key cellular regulators such as tumor protein p53 (p53) (113), Breast cancer 1, early onset (BRCA1) (137), and Cyclin dependent kinase 1 (CDK1) (76). During mitosis AurA primarily regulates processes involved in spindle assembly, including centrosome maturation (18, 70), bipolar-spindle assembly (49, 96), and metaphase chromosome alignment (97, 119). AurA also has roles in cytokinesis and mitotic exit (119).

If AurA is considered the polar Aurora, AurB is the equatorial Aurora, regulating events at the metaphase plate. AurB kinase is a chromosomal passenger protein which is critically involved in mitotic steps governing chromosome segregation and cytokinesis (30). AurB expression peaks at the G2-M transition, but kinase activity is restricted and peaks during mitosis (23, 185). Less is known about the proteins that activate AurB, but the prevailing hypotheses speculate that members of the chromosomal passenger protein complex such as Inner centromere protein antigen (INCENP) and Survivin are involved (25, 34, 77). Similar to AurA, binding of protein partners stimulates conformational changes that result in autophosphorylation on Thr232 and kinase activation (134). Once activated, AurB phosphorylates target proteins including histone H3 (38), the histone H3 variant Centromere protein A (CENP-A) (205), INCENP and Survivin (77) as well as the spindle-checkpoint proteins Budding uninhibited by benzimidazoles 1 (BUBR1) and Mitotic arrest deficient-like 2 (MAD2) (131). AurB primarily presides over the correct attachment of sister kinetochores to microtubules emanating from opposite poles. When

kinetochores align and attach incorrectly, AurB recruits BUBR1 and MAD2 to the kinetochore, halting mitosis until alignment is corrected (51). Following mitosis, AurB is ubiquitinated by the Anaphase-promoting complex/cyclosome (APC/C) and quickly degraded, allowing for mitotic exit (169). The APC/C is a multiprotein complex that ubiquitinates protein substrates, tagging them for proteasomal degradation. Whereas AurA and B have been extensively characterized, AurC shows distinct meiotic functions, paralogous to the known activity of AurB and is specifically expressed in the testis (93, 182).

Proteins that control the cell cycle must be tightly regulated as defects can lead to genomic instability and cancer and the Aurora family of mitotic regulators are no exception. Human AurB was first cloned as a gene amplified in colorectal tumors (23), and AurA was identified as the Breast tumor amplified kinase (BTAK) from chromosome 20q13; a region commonly amplified in breast tumors and other cancer cell lines (23, 165, 208). Microarray profiling of patient tumor samples has shown AurA is over expressed in ovarian (66, 81, 128), breast (181), colorectal (178), and metastatic prostate cancer (190) and it is up regulated in response to simian virus 40 (SV40) small tumor (ST) antigen (129). The lack of over expression data for AurB may be explained by conflicting studies showing that both over expression, or loss of AurB function results in a genomic instability in human cells (11, 183).

AurA over expression can induce centrosome amplification (9, 61, 126) and override the mitotic spindle checkpoint (9), thereby promoting genomic instability and transformation. Initially in NIH3T3 cells and rat fibroblasts, AurA over expression was demonstrated to cause transformation and tumors when injected into immunodeficient

mice (23, 111, 208). However, recent data from a mouse mammary over expression model revealed that AurA, by itself, was insufficient for transformation (207). AurA can phosphorylate p53 on two independent residues resulting in proteasome-dependent degradation and cell cycle progression (88, 113). Therefore, second mutations in tumor suppressor genes responsible for maintaining cell cycle fidelity, along with over expression of AurA may be enough to drive tumorigenesis in particular tumor classes. Nevertheless, loss of Aurora kinase family regulation causes chaotic chromosomal abnormalities in cells and can perhaps set the stage for the subsequent steps needed for complete transformation.

1.2 Ovarian Cancer: Therapeutics and the Aurora kinases

Ovarian cancer, which represents only 3% of all cancers affecting American females, is a deadly disease with almost 22,000 new cases diagnosed in 2008 resulting in over 15,000 deaths, a mortality rate of 72% (3). The high mortality of ovarian cancer patients is primarily due to the stage of disease at diagnosis. Patients with early ovarian tumors are asymptomatic and, as with other abdominal tumors such as those of the pancreas, patients present with seemingly unrelated abdominal symptoms, only to be diagnosed with advanced stage disease (29). Currently, the standard of care for advanced ovarian cancer is debulking surgery followed by the postoperative combination chemotherapy of the platinum analog carboplatin, and the taxane paclitaxel (43). Carboplatin is an alkylating agent that forms intrastrand DNA cross-links, resulting in cell death from extensive global DNA damage (94). The taxanes, including paclitaxel, bind the β-tubulin subunit of microtubules resulting in stabilization and inhibition of microtubule depolymerization (135). Microtubule dynamics are essential for completion

of mitosis. Taxane treatment results in activation of the spindle assembly checkpoint and subsequent cell death (135).

While some ovarian cancer patients remain disease free, unfortunately, the majority of patients relapse within 18 months of first-line therapy, and 25-59% of these patients develop resistance specifically to paclitaxel (75). Mutations in β-tubulin that remove the paclitaxel binding site or isotype switching can lead to resistant disease. Recent evidence suggests that AurA kinase may also play a role in taxane resistance (135). One study demonstrated that over expression of AurA in HeLa cells induced resistance to paclitaxel (9) while another reported sensitization of pancreatic cancer cells to paclitaxel following AurA knockdown with siRNA (73). Interestingly, a recent study in ovarian cancer cells reported that over expression of AurA can increase cell survival in the presence of paclitaxel (202). From what is known of AurA function, it seems the mechanism of resistance resides in AurA's ability to override the mitotic spindle checkpoint, forcing cells through mitosis and promoting cell survival. Chromosomal abnormalities resulting from a 'forced' mitosis contribute to cancer progression and overall genomic instability.

Due to the apparent role of the Aurora kinases in cancer, a number of next-generation, small-molecule inhibitors have been developed for *in vivo* inhibition of kinase activity. Supporting the Aurora family's value as a drug target, siRNA towards AurA can enhance sensitivity to chemotherapeutics and irradiation, causing apoptosis in human cancer cells (73, 193). To date, six separate Aurora inhibitors have been developed for cancer therapy: MLN8054 (Millennium Pharmaceuticals), Hesperadin (Boehringer Ingleheim), ZM447439 and AZD1152 (AstraZeneca), VX-680 (Merck/Vertex

Pharmaceuticals), and PHA-680632 (Nerviano Medical Sciences)(31, 90). Hesperadin (AurB) and MLN8054 (AurA) show specific inhibition of an Aurora family member, where as VX-680 is a potent and selective pan-Aurora family inhibitor introduced in 2004 (71). VX-680 is a diaminopyrimidine molecule that competes with ATP for the active site of each of the three Aurora kinases. It has shown to be a powerful and specific antitumor activity in animal xenograft models of leukemia, pancreatic and colon cancer, and importantly, ovarian cancer (71, 110). While the efficacy of Aurora inhibitors are currently being tested in human clinical trials in advanced stage tumors, these novel drugs offer hope to patients whose disease has progressed to resistance and for whom few to no treatment options are available.

The following chapter will describe experiments aimed at uncovering the relationship between AurA and paclitaxel-resistant ovarian cancer. Through microarray profiling of primary ovarian cancer samples, we observed that AurA was significantly over expressed in ovarian carcinomas compared to adenomas. We confirmed AurA expression at the protein level by staining tissue microarrays from the same patient samples. To determine if the Aurora kinases are an effective therapeutic target for ovarian tumors that have acquired resistance to paclitaxel, we tested the ability of VE-465, an Aurora kinase family inhibitor and VX-680 analog, to induce apoptosis in the presence and absence of paclitaxel in taxol-sensitive 1A9, and taxol-resistant PTX10 ovarian cancer cells (59). VE-465 potently induced apoptosis in both paclitaxel-sensitive and resistant ovarian cancer cell lines. In addition, we discovered that VE-465 synergistically enhanced apoptosis in combination with paclitaxel in taxol-sensitive cells at AurA specific doses (1-10 nM). Our data indicate that VE-465 is effective at inducing

apoptosis in both taxol-sensitive and taxol-resistant ovarian cancer cell lines, and thus

may be an effective therapy for patients with ovarian cancer, including those with

paclitaxel-resistant disease.

Chapter

II

Results:

AURORA KINASE INHIBITORS SYNERGIZE WITH PACLITAXEL TO INDUCE APOPTOSIS IN OVARIAN CANCER CELLS

Portions of this chapter reproduced from Scharer *et al*. 2008 (160)

2.1 Expression Profiling Reveals Aurora-A Over Expression in Ovarian Cancer Patients

We established gene expression profiles of ovarian cancer patients in order to determine the genes whose expression was significantly different between carcinoma, adenoma, and tumors pretreated with chemotherapy. Expression profiling of 9 carcinoma, 10 adenoma and 24 neoadjuvant chemotherapy-treated ovarian cancer patients was performed using the Affymetrix U95A gene chip, and a comprehensive analysis of these results has been published (128). Z-score normalization followed by Significance Analysis of Microarray (SAM) revealed 962 probe sets significantly up-regulated and 565 probe sets significantly down-regulated at least two fold (**FIGURE 2A**). We observed AurA to be significantly over expressed 5-fold in ovarian carcinoma patients relative to adenoma patients (**FIGURE 2B**), which is consistent with previous reports (66). We observed by SAM analysis, AurA to be over expressed 2.3-fold in carcinomas pretreated with chemotherapy compared to adenomas. However, SAM analysis did not reveal AurB or C to be significantly over or under expressed in these patients. Consistent with microarray results, we confirmed over expression of AurA by qPCR (**TABLE 2**). Interestingly, Ingenuity Pathway Assist (IPA) analysis (www.ingenuity.com) of significantly altered genes revealed that seven genes downstream of AurA in the AurA network were also up regulated at least two fold (**FIGURE 2C AND TABLE 3**). The IPA knowledge database is a curated data set based on PubMed published protein-protein interactions that are used to build gene interaction networks. IPA software can search large gene lists, such as those generated from microarray data, to find over represented gene networks. The AurA network is based on

Normalized Expression (Z score)

Low   -2   -1   0   1   2   High
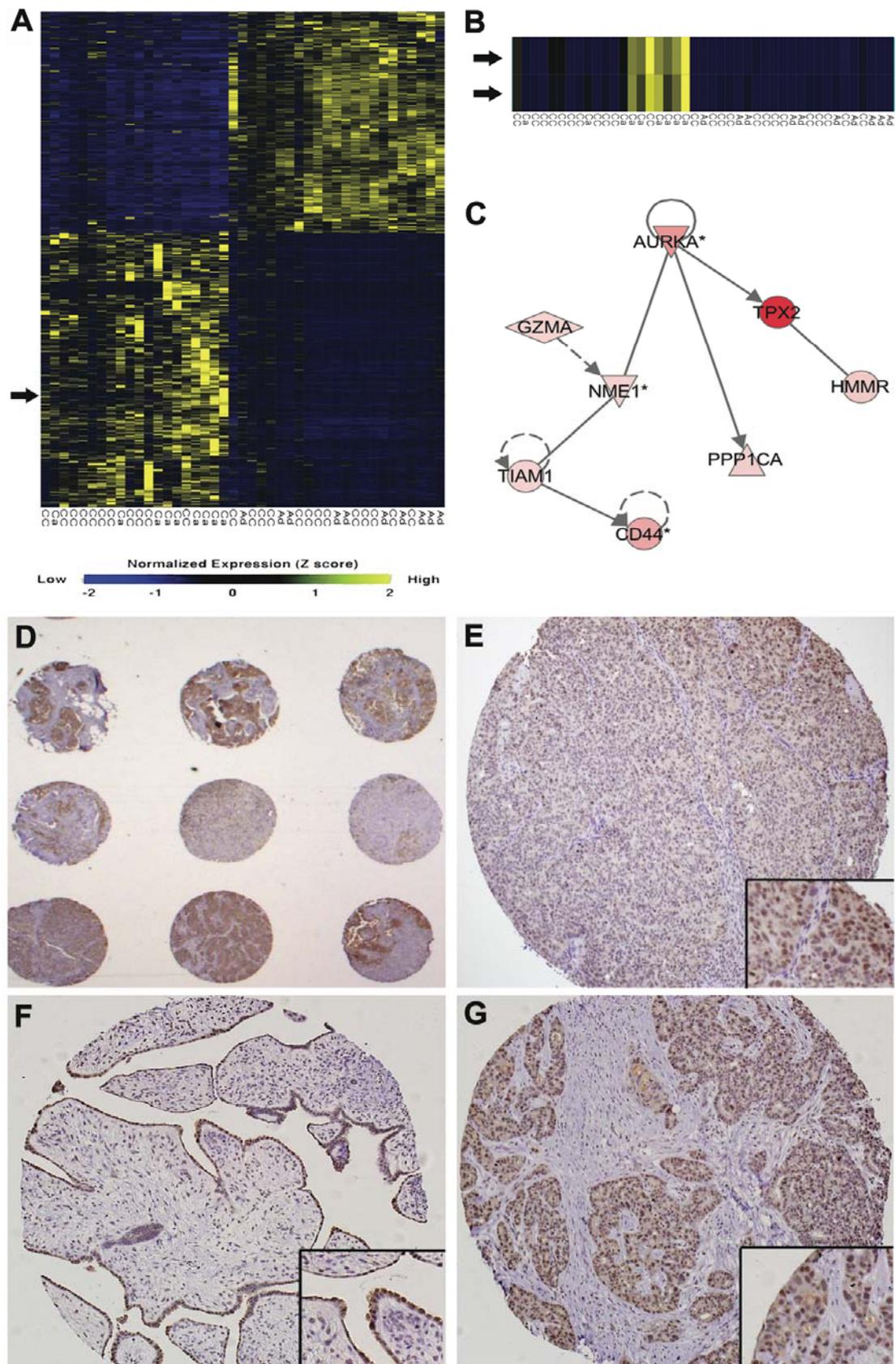
Figure 2: Aurora-A is over expressed in carcinomas. **(A)** Heat map image of Z-score normalized microarray expression data from Affymetrix U95A gene chips. Genes with lower expression compared to normal tissue are shown in blue and yellow indicates genes that are over expressed. Arrow indicates AurA. **(B)** AurA is over expressed 5 fold in carcinomas compared to adenomas. Both AurA probes are shown. Ca – carcinoma, Ad – adenoma, CC – cancers pre-treated with chemotherapy. **(C)** Ingenuity Pathway Assist analysis of significantly over expressed genes. Diagram represents an interaction network of the 7 genes and AurA kinase. **(D)** Low power (2x) image of ovarian tissue microarray stained for AurA by immunohistochemistry. **(E)** AurA staining of TMA core of ovarian carcinoma without adjuvant chemotherapy (20x). **(F)** AurA staining of TMA core of benign ovarian tissue (20x). **(G)** AurA staining of TMA core of ovarian carcinoma with adjuvant chemotherapy (20x).

Gene	*Fold Change (qPCR)*	*Fold Change (Microarray)*
TPX2	27.6 (2.1)	15.4
AURA	1.7 (0.2)	5.1
NME1	3.0 (0.7)	2.1

Table 2: Confirmation of increased mRNA by qPCR. RNA from eight patient samples (four carcinoma-like and four adenoma-like) was analyzed by qPCR, confirming increased expression levels measured by microarray analysis. Brackets represent standard deviations.

Affymetrix Probe ID	Gene	Name	Fold Change
39109_at	TPX2	TPX2, microtubule associated, homolog	15.42
1125_s_at; 1126_s_at	CD44	CD44 molecule	4.51
36863_at	HMMR	Hyaluronan-mediated motility receptor	2.73
32157_at	PPP1CA	Protein phosphatase 1, catalytic subunit, alpha isoform	2.46
40757_at	GZMA	Granzyme A	2.26
1985_s_at	NME1	Non-metastatic cells 1	2.24
38370_at	TIAM1	T-cell lymphoma invasion and metastasis	2.18

Table 3: Ingenuity Pathway Assist analysis of genes involved in the Aurora-A kinase pathway. Data represents fold enrichment in carcinoma patients versus adenoma patients. * False Discovery Rate for all genes is 0.

the following publishd interactions present in the Ingenuity Pathway knowledge database (14, 16, 26, 33, 37, 44, 50, 89, 96, 102, 122, 136, 187). Among the most highly expressed is the known AurA activator TPX2, which was over expressed 15-fold in carcinoma patients. To confirm these observed changes in gene expression by an independent method, we measured the mRNA levels of AurA, TPX2 and NME1 by quantitative real-time PCR (qPCR), verifying the microarray results (**TABLE 3**).

2.2 Ovarian Cancer Tissue Microarray Analysis of Aurora-A

To characterize the changes of expression of AurA at the protein level in ovarian cancers and benign tissues, we stained two ovarian cancer Tissue Microarrays (TMA) with antibodies to AurA. The TMA contained 212 cores from 35 patients (7 benign, 7 carcinoma without chemotherapy, and 21 carcinoma with adjuvant chemotherapy). Each

core was scored for intensity of staining (1 = weak, 2 = moderate, 3 = strong), as well as

the percentage of total cells positive for AurA, and data averaged for each patient's cores.

The TMA staining data, including detailed patient information is summarized in **TABLE

4**. On average, the benign tumors contained the highest percentage of cells staining

positive for AurA (80% ± 17%) while the carcinomas displayed a lower percentage of

cells with positive staining (61% ± 22%)(**TABLE 4**). Patients with neoadjuvant therapy

displayed an intermediate percentage of cells staining positive for AurA (73% ± 15%),

but these differences were not statistically significant with this small patient sample size.

The overall number of cells that stained positive for AurA was higher in the carcinomas

due to increased epithelial content, but the intensity of staining was equivalent with

benign ovarian epithelial cells (**FIGURES 2D-2G**). Average staining intensities were

2.5 ± 0.5 for benign tissues, 2.2 ± 0.6 for carcinomas with adjuvant chemotherapy, and

2.1 ± 0.5 for carcinoma`as without adjuvant chemotherapy. Thus, the higher mRNA

Tumor Type	Stage	Grade	No. of Patients	Age at Surgery	Survival (Months)	TMA Score	% Cells Aurora-A Positive
Benign	-	-	7	65 (10)	-	2.5 (0.5)	80 (17)
Carcinoma No Chemotherapy	I	3	1	47	-	2.9	84
	II	3	1	61	-	2	44
	III	2	1	45	-	2.4	75
		3	3	61 (14)	-	2.4 (0.7)	65 (20)
	IV	3	1	74	-	1.3	30
Carcinoma With Chemotherapy	III	1	1	55	53	1.8	59
		2	9	63 (13)	29 (16)	2 (0.6)	70 (15)
		3	9	61 (8)	33 (6)	2.3 (0.5)	79 (14)
	IV	2	1	51	62	1.8	48
		3	1	72	22	2	78

Table 4: Summary of staining and detailed patient data for the ovarian tumor tissue

microarray stained with anti-Aurora-A antibody. Brackets represent standard

deviations.

signal for AurA in ovarian cancers is likely due to the fact that there is much higher

epithelial than stromal content compared to benign tissues (**COMPARE FIGURES 2E**

AND 2F). Nevertheless, the ovarian cancer cells could be more sensitive to inhibition of

AurA than normal cells, and thus determination of the optimal dose of AurA inhibitors

will be crucial for optimizing treatment regimens.

2.3 Aurora Kinases are expressed in Ovarian Cancer Cell lines

It has previously been demonstrated that over expression of AurA can induce

resistance to paclitaxel in a cell culture model system (9). To assess the effect of Aurora

kinase inhibition on taxol-sensitive and taxol-resistant ovarian cancer cell lines, we

examined taxol-sensitive 1A9 and taxol-resistant PTX10 cells that are derived from the

parental 1A9 cell line (59). Unfortunately, the mechanism of taxol resistance in PTX10

cells is not by AurA over expression. Rather, PTX10 cells harbor a point mutation in the

M40 β-tubulin isotype, resulting in a phenylalanine to valine mutation (59) that is

hypothesized to alter binding of paclitaxel to microtubules. In fact, 1A9 cells express a

roughly two-fold higher level of AurA than PTX10 cells as determined by western blot,

and although 1A9 cells expressed lower levels of AurB than AurA, expression of AurB

was higher in 1A9 cells than PTX10 cells (**FIGURE 3A**). Thus, it was unclear whether

Aurora kinase inhibition would alter the effect of paclitaxel, or induce apoptosis via other

mechanisms. Consequently, we proceeded to test both taxol-sensitive 1A9 and taxol-

resistant PTX10 cells with the Aurora kinase inhibitor VE-465.
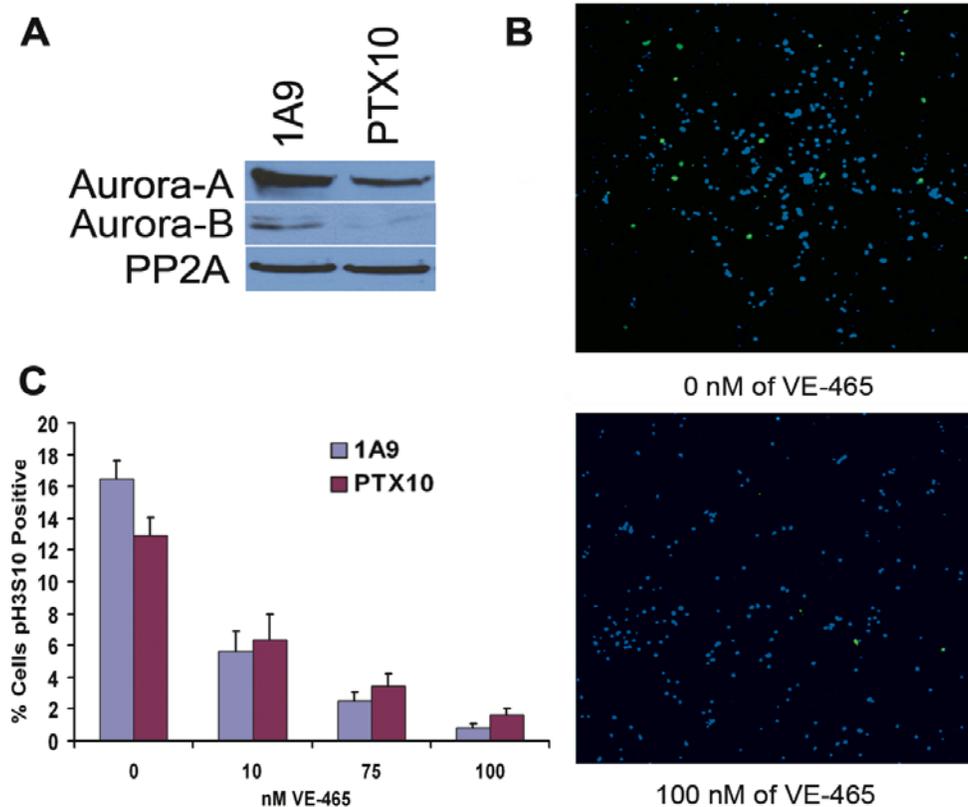
Figure 3: VE-465 inhibits the Aurora kinases. **(A)** Immunoblot analysis of whole cell lysates from 1A9 and PTX10 cell lines probed for AurA, AurB and PP2A as a loading control. **(B)** Paclitaxel-resistant PTX10 and sensitive IA9 cells were treated for 48 hours with VE-465. Following treatment, mitotic cells were assessed by staining for Histone H3 phosphorylated on Ser10 (pH3S10), a marker of mitosis and an AurB substrate (green). Nuclear chromatin was visualized with the To-Pro (blue) counter stain to indicate total number of cells. **(C)** Ten random fields were sampled for each concentration in both the 1A9 and PTX10 cell lines, values averaged and graphed to show percentage of pH3S10 positive cells at each concentration.

2.4 VE-465 Inhibits the Aurora Kinases

The Aurora kinase inhibitor VE-465 (gift of Merck & Co., West Point, PA and Vertex Pharmaceuticals, Oxford, UK), has a slightly higher K_i than the parent compound VX-680, but is still highly specific for the three Aurora kinases (AurA K_i = 1 nM, AurB K_i = 26 nM, AurC K_i = 9 nM, FLT-3 K_i = 29 nM, and ABL K_i = 44 nM) (Data from Merck & Co.). VE-465 has been shown to have activity against mutant BCR-ABL kinase in mice at 75 mg/kg (6) and to induce apoptosis in multiple myeloma cells at concentrations of 100-500 nM (48). Serine 10 on Histone H3 is a highly conserved residue that is phosphorylated by AurB kinase upon entry into mitosis (38, 80). We used immunocytochemistry to determine the percentage of cells positive for histone H3 phosphorylated on serine 10 (pH3S10) after treatment with VE-465. Treatment with 100 nM of VE-465 resulted in significant decrease in pH3S10 positive cells, whereas a DMSO control treatment had no effect (**FIGURE 3B**). Quantification of ten random fields at each concentration indicated a decrease of 7.9-fold in PTX10 and 20.9-fold in 1A9 mitotic cells when treated with 100 nM of VE-465 (**FIGURE 3C**). These results demonstrate that VE-465 effectively inhibits AurB kinase in a dose dependent manner and prevents the phosphorylation of a known mitotic marker in ovarian cancer cells. Thus, inhibitors of Aurora kinases may merit further exploration as tractable targets against ovarian cancer.

2.5 VE-465 Induces Apoptosis in Ovarian Cancer Cells

We hypothesized that treatment with VE-465 would induce apoptosis due to misregulation of the cell cycle and/or because of the instability of polyploidy cells that did manage to complete mitosis. We treated 1A9 and PTX10 cells with DMSO (control) or 10, 25, 50, 75, and 100 nM of VE-465 for 96 hours and examined DNA content by propidium iodide staining followed by flow cytometry (**FIGURE 4A AND 4B**). Fragmented DNA was calculated as a sub G0/G1 peak and was analyzed as a measure of apoptosis. After 96 hours, cell death in the parental 1A9 cell line was increased from 2.15% to 43.6% (**FIGURE 4C**) and from 4.2% to 22.6% (**FIGURE 4D**) in the paclitaxel-resistant PTX10 cell line, a roughly 5-fold increase. It is important to note that as the concentrations of VE-465 increased, both cell lines became increasingly aneuploid (**FIGURE 4B**). After 96 hours we detected cells with an array of DNA content ranging from 4n to 10n, suggesting that many ovarian cancer cells treated with VE-465 were able to bypass the spindle checkpoint, producing errors in chromosome segregation.

Consistent with the higher level of expression of AurA, and especially AurB (**FIGURE 3A**), the 1A9 cells were more sensitive than PTX10 cells to VE-465 treatment at doses of 50, 75, and 100 nM (compare **FIGURES 4C AND 4D**). To further confirm that the sub G0/G1 peak was due to apoptosis and not necrosis, we performed Caspase 3/7 assays using a luminescent detection method. Treatment of 1A9 and PTX10 cells with VE-465 resulted in a dose dependent increase in Caspase 3 and Caspase 7 activity that was inhibited by pretreatment with the general caspase inhibitor Z-VAD (**FIGURES 4E AND 4F**).
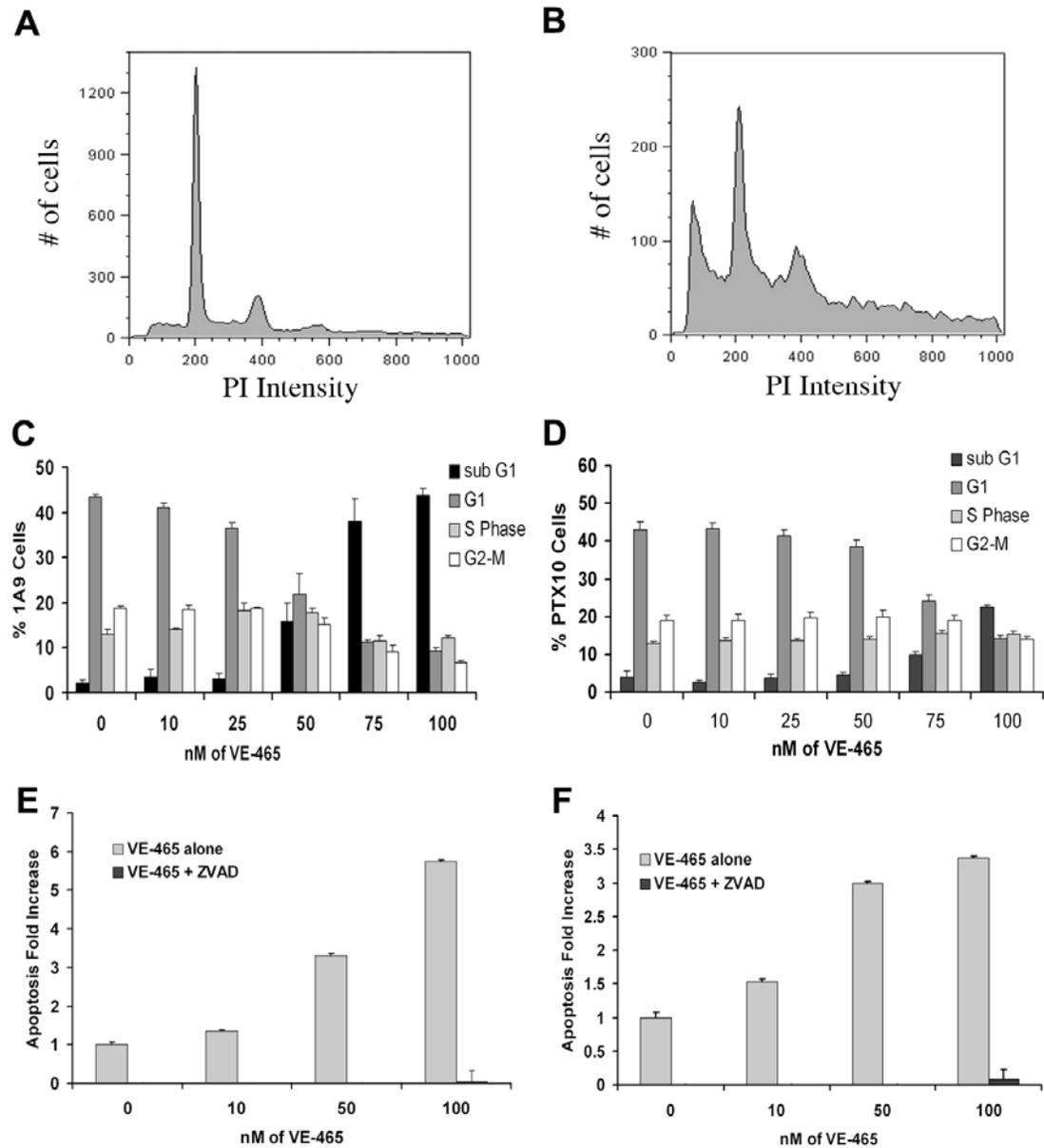
Figure 4: Inhibition of Aurora kinases results in cell death. Cells were treated for 96 hours with differing doses of VE-465. Following treatment, cells were harvested, fixed and stained with propidium iodide before analysis by Flow Cytometry. The sub G0/G1 population represents apoptotic cells. PTX10 cells were treated with DMSO Control (**A**) or 100 nM of VE-465 (**B**) for 96 hours. Significant cell death and anneuploidy was seen following treatment with 100 nM of VE-465 but not a DMSO control.

Figure 4: Graph depicting percentage of 1A9 **(C)** and PTX10 cells **(D)** in each category as determined by propidum iodide staining. Each data point represents at least 3 independent experiments. Caspase 3/7 assays of 1A9 **(E)** and PTX10 **(F)** cells treated with increasing doses of VE-465 demonstrate a dose-dependent increase in apoptosis. The caspase activity was blocked by the pan-caspase inhibitor Z-VAD.

2.6 VE-465 Promotes Apoptosis in a Paclitaxel-Resistant Cell Line at high doses

We treated 1A9 and PTX10 cells with DMSO (control) and 10, 25, 50, 75 and 100 nM of VE-465 in the presence or absence of 15 ng/mL paclitaxel for 96 hours to determine if VE-465 could induce apoptosis in the presence of paclitaxel. In the parental 1A9 cell line, paclitaxel alone caused a slight increase in the number of apoptotic cells, and the addition of VE-465 significantly increased the number of sub G0/G1 cells (**FIGURE 5A**). Consistent with their phenotype (59), PTX10 cells were resistant and proliferated in the presence of 15 ng/mL paclitaxel. The PTX10 cell line exhibited little cell death at low doses of VE-465, but as concentrations approached 100 nM the percentage of apoptotic cells increased 8-fold (**FIGURE 5B**). The presence of both drugs, paclitaxel and VE-465, did not act synergistically in the PTX10 or 1A9 cell lines at high concentrations as the levels of cell death were only slightly increased when cells were treated with VE-465 in the presence of paclitaxel (**FIGURE 5C AND 5D**). Caspase 3/7 assays of PTX10 cells confirmed that there were no statistically significant differences in induction of apoptosis between cells treated with VE-465 alone or in combination with 15 ng/mL paclitaxel (**FIGURE 5E**).
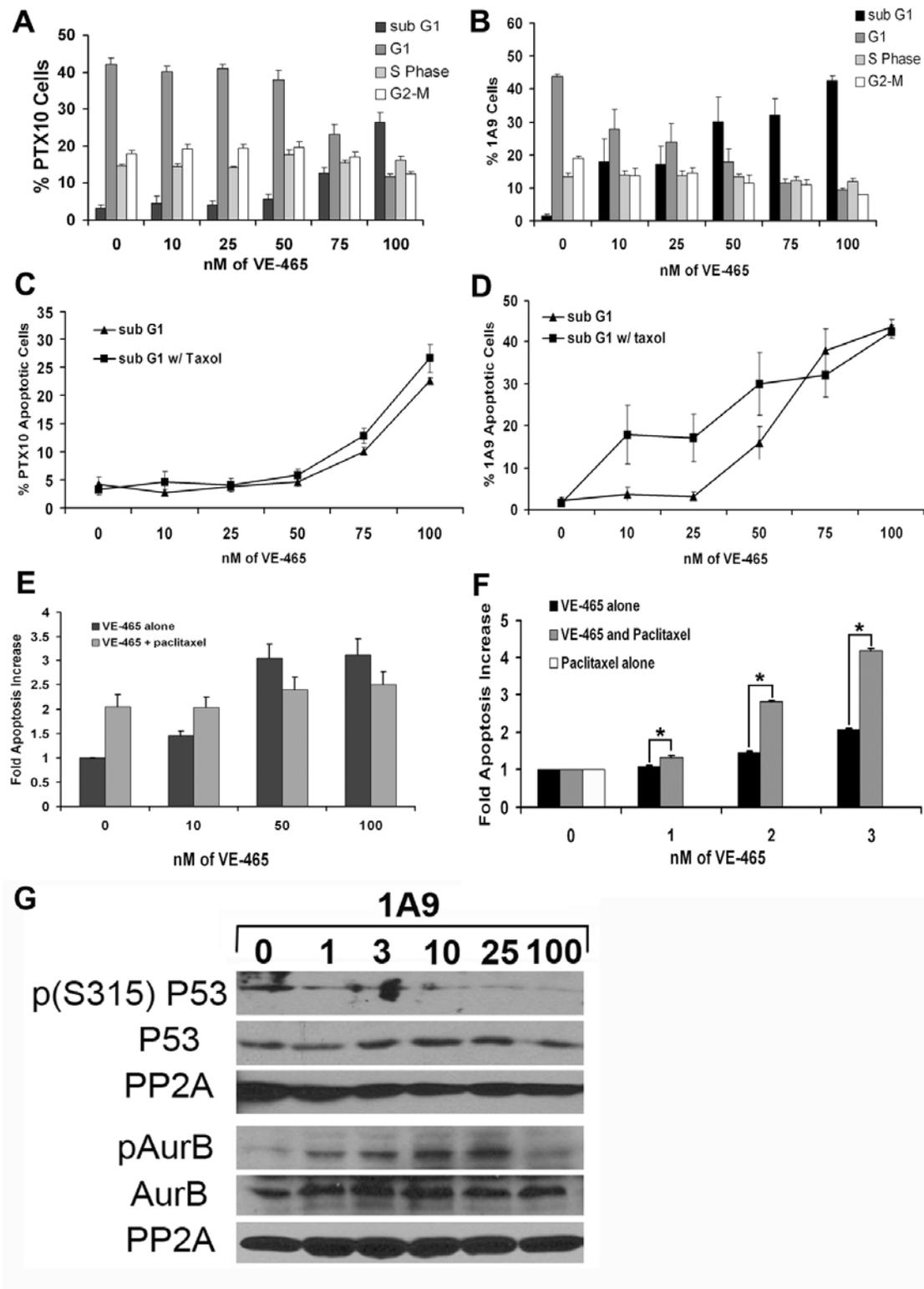
Figure 5: VE-465 induces cell death in the presence of paclitaxel. Cells were treated for 96 hours with differing doses of VE-465 in the presence of 15 ng/mL paclitaxel. **(A)** PTX10 cells **(B)** 1A9 cells. Analysis was performed as described in Figure 3. The sub G0/G1 population represents apoptotic cells. Each time point represents data from at least 3 independent experiments. Paclitaxel and VE-465 did not synergize to cause apoptosis in PTX10 **(C)** or 1A9 **(D)** cells at high doses. Percent of apoptotic cells are plotted for cells treated for 96 hrs with VE-465 alone or VE-465 and 15 ng/mL paclitaxel. Triangles – cells treated with increasing concentrations of VE-465. Squares – cells treated with increasing concentrations of VE-465 in the presence of 15 ng/mL paclitaxel. **(E)** Caspase 3/7 assays of PTX10 cells treated with 10-100 nM of VE-465 alone or in combination with 15 ng/mL paclitaxel. Confirming flow cytometry data, combination treatment with paclitaxel and VE-465 did not synergistically increase apoptosis in the PTX10 cell line. **(F)** Caspase 3/7 assays of 1A9 cells treated with 1-3 nM of VE-465 alone, 15 ng/mL paclitaxel alone, or in combination with 15 ng/mL paclitaxel. A dose of 3 nM VE-465 alone induced 2-fold more apoptosis than 15 ng/mL paclitaxel, whereas combined 3 nM VE-465 and 15 ng/mL paclitaxel synergistically induced 4.5-fold more apoptosis than 15 ng/mL paclitaxel alone. (* = p-value less than 0.0025 by students T-test.) **(G)** Immunoblot of 1A9 cells treated with increasing concentrations of VE-465 for 96 hours. The kinase activity of Aurora-A and Aurora-B is suppressed in a dose-dependent manner consistent with the known K_i values of VE-465. Phosphorylation of the Aurora-A target p53 (S315) is inhibited at doses of 1 nM and higher whereas auto-phosphorylation of Aurora-B (T232) is only inhibited at doses exceeding 25 nM.

2.7

2.7 VE-465 Synergizes with Paclitaxel to induce apoptosis at low doses specific to Aurora-A

We observed increased apoptosis at low doses of VE-465 in combination with 15 ng/mL paclitaxel in the paclitaxel sensitive 1A9 cells (**FIGURE 5C**). We tested if doses of VE-465 that were specific to AurA (3 nM or less) could synergize with paclitaxel to induce apoptosis in the 1A9 cell line. VE-465 alone induced 2-fold more apoptosis than 15 ng/mL paclitaxel alone (**FIGURE 5F**). Compared to 15 ng/mL paclitaxel alone, 3 nM of VE-465 combined with 15 ng/mL paclitaxel synergized to cause a roughly 4.5-fold increase in cell death as measured by Caspase 3/7 activity assay (**FIGURE 5F**). To confirm the effects were due to Aurora-A specific inhibition, we treated 1A9 cells with both low and high doses of VE-465 for 96 hours and probed immunoblots for phospho-AurB (T232) and phospho-p53 (S315) (**FIGURE 5G**). p53 (S315) is phosphorylated by AurA but not AurB (88) and AurB autophosphorylates threonine 232 (T232) upon activation (134). Following VE-465 treatment, phospho-p53 levels were reduced at doses of 1 nM and higher, indicating an inhibition of AurA activity. As expected, AurB kinase activity was inhibited only at doses of VE-465 that exceeded 25 nM. The level of inhibition is in agreement with K_i values for AurA (1 nM) and AurB (26 nM), respectively. These results show that VE-465 by itself can induce apoptosis, and can synergize with paclitaxel at AurA specific concentrations (< 10 nM) to enhance cell killing.

Chapter

III

Aurora Kinases:

Discussion

We identified AurA kinase to be significantly over expressed in ovarian carcinoma patient tissues compared to adenomas (128). Our data suggested that reduced p53 activity can lead to improved clinical outcome for ovarian cancer patients undergoing chemotherapy (128). One mechanism that might contribute to this phenomenon is that AurA renders cells resistant to paclitaxel-induced apoptosis and stimulates AKT1 and AKT2 activity in wild-type p53, but not p53-compromised ovarian cancer cells (202). Thus, in tumors with high AurA levels, those tumors that are p53 null would be more responsive to chemotherapy regimens. We established that the mitotic kinase AurA is over expressed in ovarian carcinomas compared to adenomas at the mRNA and protein level. Furthermore, we demonstrated that the pan-Aurora inhibitor, VE-465, can synergize with paclitaxel to induce apoptosis and is a potent killer of taxane-sensitive and resistant ovarian cancer cells.

Although other Aurora family members were not over expressed, additional genes known to interact with AurA kinase were significantly increased (**FIGURE 2C and TABLE 3**). General misregulation of a host of genes involved in regulating the cell cycle can contribute to genomic instability, and activation of specific 'driver' genes can increase this process. One of the most significantly over expressed AurA network genes was TPX2, an activator and substrate of AurA (96, 164). Recently, a link between another AurA substrate, BRCA1, and TPX2 has been demonstrated (85). Juokov *et. al.* showed that loss of BRCA1 expression leads to mislocalization of TPX2 along microtubules instead of at the aster poles, suggesting a mechanism by which BRCA1 mutation could lead to chromosomal instability (85). This finding also suggests that BRCA1 could lie upstream of TPX2, regulating spatial localization of TPX2 and in turn,

controlling temporal activation of AurA. Correct spatial and temporal regulation of proteins that regulate the cell cycle is critical to maintain a healthy cell. We did not test AurA kinase activity directly in ovarian tumors, however TPX2 was over expressed 15-fold in carcinomas. TPX2 binds AurA and induces conformational changes that result in catalytic activation of AurA and subsequent auto-phosphorylation. This may provide a possible mechanism for increased AurA kinase activity in ovarian tumors. Future experiments determining AurA kinase activity in tumors that express high and low levels of TPX2 would confirm this hypothesis. It may be possible to develop drugs that inhibit the AurA-TPX2 interaction, and thus limit the amount of active AurA kinase in tumors.

These observations have implications for ovarian cancer because over expression of AurA can induce resistance to the chemotherapeutic paclitaxel (9). We predict that ovarian cancer patients who over express the kinase would have a higher chance of becoming resistant to the taxanes and possibly benefit from a different treatment strategy targeted at AurA and other Aurora family members. To test this prediction, we evaluated the compound VE-465 as a pan-Aurora inhibitor and inducer of apoptosis in ovarian cancer cell lines. Although VE-465 is not specific to AurA, it is highly selective and effective at inhibiting the Aurora family of kinases and offered a unique opportunity to evaluate the entire family of kinases as a therapeutic target. Our results indicate that VE-465 is able to induce apoptosis in the paclitaxel-resistant, ovarian cancer cell line PTX10 in a dose dependent manner and synergize with paclitaxel in the 1A9 paclitaxel-sensitive cell line.

VE-465 and paclitaxel are both drugs that function by targeting mitotic cells, but induce apoptosis by different mechanisms. Paclitaxel alters microtubule dynamics and

induces the spindle checkpoint resulting in mitotic arrest and eventual apoptosis. VE-465, on the other hand, inhibits the activity of the Aurora kinase family and subsequent mitotic regulation. We found that many PTX10 cells treated with VE-465 bypass the spindle checkpoint resulting in missegregation of chromosomes and aneuploidy, possibly due to the inhibition of other Aurora family members such as AurB. Thus, VE-465 appears to induce apoptosis by causing catastrophic chromosomal abnormalities in cells that do manage to crash through mitosis due to the absence of an intact spindle assembly checkpoint.

Intriguingly, 1A9 cells were more sensitive to VE-465 than PTX10 cells and this correlates with the roughly 2-fold higher expression of AurA in the 1A9 cell line. Significant cell death was observed at low concentrations in 1A9 cells (1-25 nM) relative to the widespread apoptosis seen at higher concentrations (50-100 nM) for both cell lines. This indicates that at low doses VE-465 synergizes with paclitaxel in taxol-sensitive ovarian cancer cells. Interestingly, at low concentrations VE-465 has a K_i more specific to AurA (1 nM) than AurB (26 nM) or C (9 nM). This suggests the synergistic effects are due to the specific inhibition of AurA and not the other family members. However, at higher concentrations, we found no evidence that paclitaxel and VE-465 synergized to induce apoptosis in PTX10 cells. This could be because a very high percentage of cells are undergoing apoptosis at high doses, or possibly due to the inherent nature of resistance of PTX10 cells. PTX10 cells harbor a point mutation in the M40 β-tubulin isotype resulting in a phenylalanine to valine mutation (59), which may alter the binding of paclitaxel to microtubules. It is possible that this particular form of resistance does not coincide with the function of Aurora kinases and therefore no synergism is seen when

treating with a combination of both drugs. Tumors that exhibit other forms of taxane-resistance such as AurA over expression, alternate point mutations, modulations in tubulin isotypes, decreased tubulin expression and changes in post-translational modifications may respond synergistically when treated with VE-465 and paclitaxel. Alternatively, a synergistic effect may be observed prior to the acquisition of taxol-resistance, or in combination with other drugs that target different cellular pathways such as tyrosine kinase receptor signals or apoptosis resistance pathways. In line with our tumor model, the 1A9 cells did express a roughly 2-fold higher level of AurA than the PTX10 cell line. One can infer that tumors with higher AurA levels may be more 'addicted' to the AurA pathway and therefore would be more responsive to Aurora inhibition. Nevertheless, Aurora kinase inhibitors represent a promising alternative combination to taxane therapy, especially for patients who over express the mitotic kinase AurA, or other family members, or whose disease continues to progress in spite of taxane therapy (57).

Treatment of patients with different drugs in a serial fashion allows for clones that are resistant to one or more therapies to arise by drug-resistance selection. However, combinatorial therapies may be more effective, as has been shown using cocktail therapies for the treatment of the rapidly evolving human immunodeficiency virus (15). Thus, initial combinatorial chemotherapy using Aurora inhibitors, paclitaxel, and other chemotherapeutic agents could be an effective approach to prevent the development of chemoresistant cancers. Alternatively, molecular tumor typing to determine the genes and subsequent pathways that are misregulated may provide insight and guidance for possible treatment options. Whereas the required biopsy and molecular diagnostics are

not possible in all tumor types and even more difficult in many solid tumors, first-line therapy for ovarian cancer is primarily surgery followed by chemotherapy. Therefore, it may soon be possible to screen a molecular diagnostic array, assessing the status of genes such as the Aurora kinases, to determine the therapeutic treatment options that will provide the best possible outcome with the lowest chance of developing resistant disease.

In summary, we have shown the mitotic kinase AurA to be over expressed in ovarian carcinomas compared to adenomas. Furthermore, we demonstrated the pan-Aurora inhibitor VE-465 can synergize with paclitaxel to induce apoptosis and is a potent killer of taxane-sensitive and resistant ovarian cancer cells. Our *in vitro* results suggest that Aurora kinase inhibitors could be useful for treatment of taxane resistant ovarian tumors and possibly as a first-line therapy against those tumors with high expression levels of the Aurora kinases. Future experimental directions include testing the efficacy of Aurora inhibitors using *in vivo* mouse xenografts models. 1A9, PTX10, and ovarian cancer cell lines with elevated AurA expression can be grafted onto immunodeficient mice that are subsequently treated with AurA and Aurora family specific doses of VE-465. This will determine the *in vivo* efficacy of Aurora inhibition and other possible gross side effects. VX-680 has already demonstrated little toxicity in previous mouse xenograft experiments (71), nevertheless these studies are critical as proof-of-principle before progressing into humans. The development of a variety of targeted, small molecule inhibitors towards the Aurora kinases is a promising step in the search for new and improved cancer treatment options.

Chapter

IV

SOX4:

The Diversity of Mammalian SOXs

Complex eukaryotes, such as mammals, begin life as a single cell and undergo rounds of cell division and specialization to expand beyond a trillion cells. Roughly two meters of DNA is exquisitely packaged into a nucleus roughly 6 μm in diameter (7). DNA must be condensed and packaged in a manner that is dynamic enough to allow different sets of genes to be expressed as a cell differentiates and the organism develops, yet maintain the compactness to stay within the nuclear boundaries. Eukaryotes compact and limit access to their DNA through varying levels of organization and regulation, including chromatin compactness and spatial location within the varied nuclear compartments. The process of transcriptional regulation involves a host of DNA binding proteins that require access to specific DNA sequences throughout the genome to exert either a positive or negative affect. DNA binding proteins can function in the direct regulation of DNA compaction and accessibility such as histones, directly in the transcriptional process like the transcription initiation factors, or as in transcription factors, serve dual roles and regulate both processes. This chapter will concentrate on the SOX family of transcription factors, focusing on SOX4, and their role as vertebrate regulators of development and differentiation, as well as carcinogenesis.

4.1 A Collection of SOXs

Multicellular organisms require precise cellular control to ensure the correct differentiation of the required tissue types and accurate placement throughout the body of each organ. Evolution has supplied a myriad of transcription factors that function in development to ensure differentiation functions normally. Examples of developmental transcription factor families include, but are not limited to: the POU domain, RUNT domain, Forkhead, helix-loop-helix, homeodomain, and the SOX family. The SOX family of transcription factors is defined, as are the other transcription factor families listed above, by their DNA binding domain. SOX proteins contain a high-mobility group (HMG) box DNA binding domain that was originally classified according to their ability to bind AT rich DNA (63). HMG DNA binding proteins fall into 4 distinct categories comprising the HMG domain superfamily: HMG1/2, HMG domain family, HMG-I(Y) or AT-hook family, and the non-canonical HMG box domain proteins that include the SOX and TCF family of transcription factors. The founding member of the SOX family, SRY, was first isolated from the Y chromosome as the gene responsible for sex reversal in males (67, 166). It was found to contain a unique HMG domain and subsequent PCR based searches looking for similar DNA sequences identified SOX family members and homologs in all vertebrates (41, 52, 199). The advent of high throughput DNA sequencing technology allowed for bioinformatic characterization of the SOX family in sequenced genomes. The SOX family was then refined in mouse and humans to 20 members comprising eight different subgroups (162). The various SOX family subgroups, including protein domains are summarized in **FIGURE 6.**

Group	Gene	Schematic
A	*Sry*	1 3 81 263 345 395
B1	*Sox1*	1 48 127 391
	Sox2	1 41 119 319
	Sox3	1 67 145 375
B2	*Sox14*	1 6 84 240
	Sox21	1 6 84 125 276
C	*Sox4*	1 58 136 298 440
	Sox11	1 47 125 273 395
	Sox12	1 39 109 314
D	*Sox5*	1 77 116 182 261 391
	L-Sox5	1 158 240 364 403 469 547 679
	Sox6	1 181 262 487 515 617 696 827
	Sox13	1 120 196 396 475 595
E	*Sox8*	1 57 97 176 349 441 646
	Sox9	1 63 103 182 406 507
	Sox10	1 61 101 180 234 466
F	*Sox7*	1 43 122 380
	Sox17	1 68 147 343 419
	Sox18	1 166 245 278 345 468
G	*Sox15*	1 44 123 125 213
H	*Sox30*	1 365 435 782
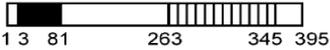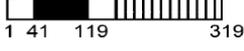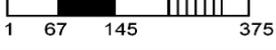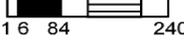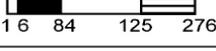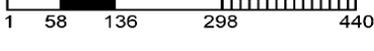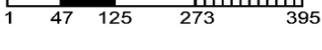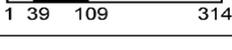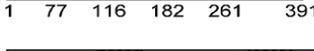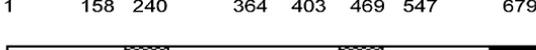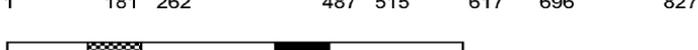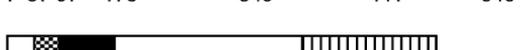
Figure 6: SOX Family Subgroups and Protein Domain Architecture. SOX proteins are depicted as boxes and numbers represent the first and last amino acids as well as internal boundaries for the functional domains. Black boxes = HMG domain, checkered boxes = dimerization domain, horizontal striped boxes = transrepession domain, and vertical boxes = transactivation domain. Adapted from Levefbre *et. al. (100)*

SOX proteins within the same group share a high degree of sequence identity within and outside the HMG domain whereas different groups share a limited homology only within the HMG domain. The HMG domain is responsible for DNA binding, and has been shown to harbor two nuclear localization signals (NLS) (175) and facilitates protein-protein interactions (196). By definition, transcription factors must exert some of their functions in the nucleus where the cellular DNA is housed. Nuclear import is mediated in part by the importin family of proteins that must recognize an NLS in their cargo prior to nuclear translocation. Non-nuclear functions have been reported for SOX4 (58), but, evidence suggests that nuclear localization is required for full functionality. This is highlighted by studies of SRY in which mutation of only the NLS was sufficient to cause sex reversal in mice (58). Although only two SOX proteins, SRY and SOX9, have had extensive biochemical NLS analysis, the two NLS motifs are highly conserved throughout the family.

Protein-protein interactions are crucial for SOX activity and can be classified into three distinct functional categories: those that bind nuclear import proteins, adaptor proteins, and DNA binding cofactors (196). While the importance of the first category has already been discussed, the ability of SOX proteins to bind cellular adaptors, such as

PDZ domain proteins, is still poorly characterized. The PDZ acronym is derived from the three proteins the domain was initially identified in: PSD95, DIgA and ZO-1. PDZ domain proteins exist in all metazoans and recognize short peptide motifs that act to bring proteins together into signaling complexes (72). The best characterized example is the interaction of SOX4 with the PDZ protein Syntenin (58). Here SOX4 transcriptional activation from the IL-5α receptor requires the PDZ adaptor protein Syntenin. SRY has also been shown to interact with the PDZ domain protein SIP-1 (Entrez ID: 8487) (147). These interactions are DNA independent and could provide the framework to assemble multiprotein complexes that function as transcriptional or non-nuclear signaling nodes.

The final category involves protein-protein interactions with other DNA binding proteins. The SOX family HMG domain recognizes the core motif WWCAAWG, where W = A or T. Therefore, interaction with specific partner proteins can provide sequence specificity for a SOX protein and partner to affect transcription of a specific set of genes. Multiple SOX family members can be expressed in the same cellular context and specific protein partners can provide target specificity to activate or repress transcription of a distinct gene set. There is a growing list of DNA binding proteins with which SOX family members can interact (196). The most thoroughly characterized interaction is that of SOX2 and the OCT3/4 complex binding to the FGF-4 enhancer in embryonic stem cells (152). Crystal structure analysis of SOX2 and OCT3/4 pair bound to DNA reveals how DNA binding places the C-terminus of the HMG domain in perfect position to interact with OCT3/4 (152). Unlike other transcription factors that bind the DNA major groove, HMG domain proteins, like SOX2, bind the minor grove. It is this placement, roughly one-half helix turn away from OCT3/4's binding site that positions SOX2 for the

interaction. In this example, the transcriptional complex functions to activate the FGF-4 enhancer; however other DNA dependent protein-protein interactions can function to stabilize SOX protein binding to a low affinity binding site and act to either activate or repress transcription of target genes. The ability to bind different partners that are also expressed in specific cell types allows for factor specific transcriptional programs in defined cell types and adds functional diversity to a family with similar DNA binding domains.

Molecular characterization of the SOX family has been ongoing from the moment of discovery; however, the advent of genetic manipulation has allowed for the tissue and animal specific functions of SOX proteins to be elucidated. Current research has identified *in vivo* roles for SOX factors in sex differentiation in mammals, promoting various levels of potency for stem cells, neurogenesis, neural crest formation, skeletogenesis, hematopoiesis and endoderm development (100). Most of the information is derived from phenotypes of genetic knock-out or knock-in mice. The detailed molecular identification of transcriptional targets and signaling pathways is still in its initial stages. Nevertheless, it is clear that the SOX family of transcription factors is critical and required for vertebrate differentiation and development.

As described earlier, the founding member of the SOX family, SRY, was originally cloned as the gene responsible for 'maleness' (67, 166) in mammals. Female XX mice that have a genetic knock-in for SRY develop as fertile males (166). Like SRY, SOX9 mutations can affect sex determination in mammals. It is estimated that two thirds of patients with gonad defects and XY females contain mutations in SOX9 (54, 191). Other SOX proteins that are not exclusively required for sex differentiation but are

essential for correct gonad development are the SRY homolog SOX3 (194), and SOX8 (32) which can partially substitute for SOX9 function. The role of SRY in sex determination and function is specific to the mammalian lineages of the animal kingdom but SOX9 has roles in sex determination outside of mammals (130).

Of all the SOX genes, SOX2 has the best studied and most famous due to its role as one of four transcription factors that regulate pluripotency in embryonic stem cells (12). One critical function is to dimerize with the POU domain transcription factor OCT3/4 and activate transcription of the FGF-4 enhancer, a critical signaling molecule for early blastocyst development (204). A landmark paper demonstrated the reversal of differentiated adult fibroblasts to stem cell 'like' cells by the addition of four transcription factors: SOX2, C-MYC, KLF4 and OCT3/4 (177). This finding suggests that SOX2, in combination with the other three factors, are inherently programmed to transcribe gene sets involved in maintaining pluripotency and can overcome and reset the current transcriptional program of a fully differentiated cell type.

While SOX2 is critical for early developmental stages, other SOX family members have been shown to be involved in regulating differentiation of the neural crest and the central nervous system. Examining neural development highlights the complicated and essential role of SOX proteins in development. SOX genes regulate each other and only fine-tuning of specific SOX family member expression at distinct developmental time points allows for correct neural development. Some SOX factors, such as SOX4 and SOX11, drive pan-neural gene expression and commit cells to the neural lineage (21). Activation of other SOX family members, such as SOX9, further commit neuronal cells to specific lineages which specify the glial cell population (170-

172). Interestingly, the same SOX family genes are involved in the specification of neural crest cells. Mutations in the group E protein SOX10 cause Waardenburg-Hirschsprung disease characterized by defects in pigmentation and the peripheral nervous system (144) whereas group C members SOX4 and SOX11 are required for development of heart septums (163, 168). This suggests that these proteins control similar processes or stages of development in both the central and peripheral nervous systems.

Although the majority of the work deciphering the various processes SOX proteins control has been performed in neuronal lineages, it is also understood that SOX genes control skeletal, hematopoietic (163), and endodermal derived tissue development (180). Similar to glial development, SOX9 activates transcription of the group D members SOX5 and SOX6. These two factors, in contrast to repressing SOX9 activity in glial development, cooperate to activate transcription of genes critical for all stages of chondrocyte differentiation and maturation, highlighting the diverse and complex roles of SOX proteins (101). The same sets of transcription factors can drive the development of multiple cell types through lineage specific collaboration with other transcription factors and signaling molecules. In order to act in a cell type-specific manner, the influence of either post-translational modifications or direct protein binding partners must play a major role. In some cases the tissue specific factors are known, but it is less well understood which downstream transcriptional programs are activated in specific cell types.

Understanding the diverse cellular roles of SOX proteins in controlling all stages of development from the earliest cell divisions in the blastocyst to organ specific cell differentiation is in its infancy. From the first molecular cloning of SRY and subsequent

identification of SOX family genes two decades ago, the physiological and molecular roles of SOX proteins continue to expand and diversify, further complicating a complete understanding of this distinct and ancient gene family.

4.2 SOX4

Sex-Determining Region Y Box 4 (SOX4) is a 47 kDa, single exon gene located on human chromosome 6p22. Homologs of SOX4 have been identified in all sequenced vertebrate genomes from Zebra fish and Chicken to Mouse and Humans. Sequences outside the HMG DNA binding domain show little evolutionary conservation while the HMG domain demonstrates little divergence with greater than 95% identity across sequenced species. SOX4 resides in subgroup C along with SOX11 and SOX12 who all share sequence homology in their DNA binding domain and surrounding residues (**FIGURE 6**). Since the initial cloning and characterization of SOX4 by Farr *et. al.* (52), an explosion of data has implicated a role for SOX4 in controlling diverse developmental process in multiple tissues, defined molecular interactions with cellular regulatory proteins, identified bona fide SOX4 transcriptional targets, and finally proposed roles in the regulation of cell death and carcinogenesis. In the following three sections I will explore the current knowledge of SOX4 as it relates to these general areas.

4.2.1 Developmental Regulator

Current knowledge indicates SOX4 is involved in early developmental processes, but it is not known when SOX4 expression is first initiated or in which specific cell types. SOX4 is known to be expressed in the developing mouse central nervous system (CNS) and neural tube as early as day E8.5, with expression being restricted to specific differentiating neural lineages in the developing brain by day E11.5 (35). In humans, SOX4 has a higher expression in the developing fetus than in adult tissues. Screening an RNA Master Blot which contains poly(A+) RNA from various adult and fetal human tissues, Reppe *et. al.* found that SOX4 is most highly expressed in the fetal brain, kidney and lung. SOX4 expression is detectable at low levels in all adult tissues analyzed with dramatically higher expression in the ovary and thymus (154). These data are in line with known roles for SOX4 in T-cell development (189, 198) and its upregulation during the female reproductive hormonal cycle (82).

The power of murine transgenics has enabled experiments in the whole mouse or in specific tissues, either ablate SOX4 entirely or ectopically express SOX4. Two groups have independently engineered whole mouse homozygous deletions for SOX4 roughly 11 years apart (142, 163). Deletion of SOX4 results in embryonic lethality around day E14 due to the failure of the endocardial ridges to fuse properly and develop atrioventricular valves, resulting in cardiac failure. Interestingly, the cells that form the endocardial ridge derive from the neural crest where SOX4 is known to be expressed early in development (35). All other tissues examined were able to develop normally by histological analysis up to day E14 except there was a disruption during B-cell development with a lack of the pre-B-cell population (characterized by the cell surface markers CD43$^-$ B220^{++}). This

suggests SOX4 is also involved at specific stages of hematopoietic differentiation and development.

Further studies in the neuronal lineages have elucidated roles for SOX4 in controlling the differentiation of specific cell populations. Prolonging SOX4 expression in glial cells throughout development resulted in severe anatomical defects of the brain due to migratory defects of specific cells (78). When SOX4 expression is maintained in oligodendrocytes, they fail to promote myelination, suggesting that SOX4 holds this cell population in a pre-differentiation phase. Only when SOX4 is down regulated does myelination proceed normally (146). Interestingly, CNS specific knockout of other SOX subgroup C members SOX11 and SOX12 suggests functional redundancy within the subfamily. They all exhibit similar molecular properties (46) and the SOX12 deletion revealed little phenotypic defects (79), suggesting other family members can compensate for loss of SOX12. Outside of the CNS, SOX4 has a role in pancreatic development in which SOX4 null mice fail to undergo a second round of differentiation and expansion, resulting in a lack of β-cells that are essential for insulin secretion (197). Furthermore, SOX4 has been shown to be involved in pancreatic differentiation in Zebra fish (*121*) and insulin secretion from β-cells in mice (62). Although inconclusive, other studies have suggested that SOX4 expression must be efficiently down regulated for secretory activation in mammary gland (132), as well as bone development (22). These studies clearly show that SOX4 is not required for complete organ development, but SOX4 must be tightly regulated during development in specific cell types. Differentiation is determined by tissue specific temporal regulation of SOX4 and as these studies show, having too much or too little can result in severe developmental consequences.

4.2.2 Molecular Functions

Although tissue specific expression and roles in differentiation of specific cell types are being deduced genetically, additional experimentation has addressed the molecular functions of SOX4. It is clear that SOX4 itself can be subdivided into distinct domains that function in DNA binding and transcriptional activation, but there are also reports that discrete sequences within SOX4 can contribute to protein-protein interactions; including promoting apoptosis. Interestingly, like the transcription factor p53, there appears to be functions for SOX4 that are independent of DNA binding.

Initial characterizations of SOX4 began when it was independently cloned as a gene responsible for activating the sequence AACAAAG (189), a sequence also identified in colon carcinoma cells to be bound by SOX4 (124). Independently, SOX4 was shown to bind and activate the sequence AACAAT present in the CD2 enhancer in T-cells (198). Both of these sequences match the consensus SOX family binding site WWCAAWG but most likely are not specific for SOX4. It is not known if there is a binding site specific for SOX4 or if the influence of protein binding partners or post-translational modifications determines the precise sites SOX4 can bind to. For that matter, there have been only two reports of other proteins that can bind SOX4 and influence transcription both negatively or positively. The first study identified upstream activators of SOX4 through a yeast 2-hybrid approach (58). The authors mapped the binding of Syntenin and SOX4 to show that binding of Syntenin to the cytoplasmic tail of the IL-5 receptor caused recruitment of SOX4 to Syntenin and subsequent transcriptional activation of a SOX4 reporter plasmid. The second study reported the interaction of SOX4 and ubiquitin conjugating enzyme UBC9 (139). While little molecular

information about the interaction is known, the authors reported that over expression of

UBC9 suppressed SOX4's ability to activate a transcriptional reporter following

treatment of breast cancer cells with progesterone, a known inducer of SOX4 (64, 125).

It is not known if UBC9 can polyubiquitinate SOX4 and induce proteasomal degradation.

Both studies identified protein binding partners that either positively or negatively

influenced transcription. Although their conclusions were generalized, they indicated

that specific binding partners have the ability to influence the activity of SOX4 in distinct

circumstances, further alluding to the importance of protein binding regulators to SOX4

function.

Functionally, SOX4 is a transcription factor and is defined by the HMG DNA

binding domain where it is most similar to other family members. However, the

sequences outside of the DNA binding domain definitively place SOX4 as a SOX group

C protein and allow for SOX4 specific functions. The extreme C terminus is highly

conserved between SOX group C proteins, and has been shown by multiple groups to

harbor a transcriptional activation domain (TAD) (46, 189). Recently, the HMG and

TAD were demonstrated to interact with p53 by co-immunoprecipitation and deletion

analysis, demonstrating the first transcriptional binding partner for SOX4 (140). Outside

of these two regions there are suggestions that SOX4 has a pro-apoptotic domain.

Specific over expression of the glycine-rich region (amino acids 152-227) caused

apoptosis and DNA fragmentation in HEK293 cells and SOX4 is upregulated during

prostaglandin induced cell death (4, 5, 83, 92). Interestingly, there is equal evidence

from our lab and others that complete loss of SOX4 expression can induce apoptosis

(112, 148), however, this is not the case in every tissue (133). In prostate cells, SOX4,

possibly in cooperation with p53, regulates PUMA and Survivin to induce cell death

(112). Either way, there is no clear evidence outlining how SOX4 functions in apoptosis

and if this is dependent on SOX4's pro-apoptotic domain, p53 binding, or downstream

transcriptional targets of SOX4. It is possible that cells must maintain a fine level of

control over SOX4 levels and cells can not tolerate having too little or too much SOX4 is

toxic and induces cell death.

Functional characterization of SOX4 will determine how SOX4 operates as a

transcription factor, and identification of transcriptional target genes and biologically

relevant binding sites is necessary to develop models of the gene regulatory networks

SOX4 influences. DNA microarrays are powerful technologies that allow the

identification of all mRNAs present in a sample at the time of RNA isolation.

Comparison of two different experimental conditions can identify gene expression

changes between the two conditions. To identify putative transcriptional targets of a

transcription factor, such as SOX4, the two conditions can be loss of SOX4 versus

normal or over expression of SOX4 versus normal. Either condition will determine the

genes whose expression changes when SOX4 levels are perturbed. However, this does

not definitively classify genes as direct SOX4 targets because intermediary factors may

be involved. Nevertheless, there have been attempts by three separate groups in addition

to our laboratory to use this technology to identify SOX4 direct target genes. Two groups

eliminated SOX4 using siRNA (148) or CRE recombinase embryonically (197), one

group performed a time course of SOX4 over expression (1) and our laboratory used both

siRNA knockdown and over expression (112) to identify SOX4 regulated genes. All

experiments identified hundreds of genes as possible SOX4 targets. Unfortunately, each

study was performed in different tissues or cell lines, and as a result, there is little overlap between the genes identified. Whereas, this result could be due to the difference in experimental design, it also suggests that SOX4 has distinct transcriptional targets that are cell type specific.

4.2.3 Carcinogenesis

Carcinogenesis is the process by which normal cells undergo transformation and become cancerous. In their landmark paper (69), Hanahan and Weinberg outlined six processes or traits that all cancers have in common: (1) insensitivity to anti-growth signals, (2) insensitivity to pro-apoptotic signals, (3) sustained angiogenesis, (4) self-sufficiency in growth signals, (5) tissue invasion and metastasis, and (6) limitless replicative potential. SOX4 has been implicated in regulation of apoptosis, cell growth, and metastasis; all processes that are directly related to cancer progression. In fact, four independent studies, using replication-deficient retroviruses to search for oncogenes, identified integration sites within the SOX4 locus that affect SOX4 expression (45, 106, 108, 115) with one study identifying SOX4 as the most commonly disrupted gene (106). Microarray analysis of primary tumor samples has identified SOX4 upregulation in leukemias (10), melanomas (179), glioblastomas (176), medulloblastomas (99), pineal tumors (53), ovarian tumors (128), colorectal tumors (151, 167), tumors of the lung (55), cancers of the bladder (1), and in breast cancer cell lines (64). A recent meta-analysis identified SOX4 as one of 64 genes upregulated as a general cancer signature (156), further suggesting SOX4 has a role in the formation of multiple tumor types.

Prostate cancer is the most common non-cutaneous cancer among American men with an estimated 186,000 new cases in 2008, resulting in almost 29,000 deaths (3).

Serum PSA tests can predict prostate size; however, this is not an indication of malignancy and molecular diagnostic tests and markers are sorely needed (39). Transcription factors, such as the androgen receptor (AR)(186), NKX3-1 (47) and HOXC6 (149), are known molecular drivers of prostate cancer. Our lab has previously shown SOX4 to be over expressed in primary prostate tumors at the mRNA and protein level, and these levels correlated with Gleason score (112). In fact, no less than seven independent studies have also observed SOX4 mRNA up regulation in prostate tumors (42, 47, 98, 116, 118, 155, 195). It is not previously known how SOX4 is upregulated in prostate, or other tumors. Over expression of SOX4 in the non-transformed prostate cell line RWPE-1 causes anchorage-independent growth in soft agar assays, suggesting SOX4 has transformation potential in prostate cells (112). However the molecular role of SOX4 in tumorigenesis of the prostate, or other tissues, is not known.

In the next chapter I will describe experiments focused on identifying the transcriptional targets of SOX4 using a combination of a genome-wide localization ChIP assay and transient over expression followed by expression profiling in a prostate cancer model cell line. We have also used protein-binding microarrays to derive a novel SOX4-specific position-weight matrix and determined that SOX4 binding sites are enriched in SOX4-bound promoter regions. Direct transcriptional targets of SOX4 include several key cellular regulators, such as Epidermal growth factor receptor (EGFR), Heat shock protein 70 (HSP70), Tenascin C, Frizzled-5 (FZD5), Patched-1 (PTCH1), and Delta-like 1 (DLL1). We also show that SOX4 targets 23 transcription factors, such as Mixed-lineage leukemia (MLL), Forkhead box A1 (FOXA1), Zinc finger protein 281 (ZNF281), and NK3 homeobox 1 (NKX3.1). In addition, SOX4 directly regulates expression of

three components of the RNA-induced silencing complex, namely Dicer, Argonaute 1, and RNA Helicase A. These data provide new insights into how SOX4 affects developmental signaling pathways and how these changes may influence cancer progression via regulation of gene networks involved in microRNA processing, transcriptional regulation, the TGFβ, WNT, HEDGEHOG, and NOTCH pathways, growth factor signaling, and tumor metastasis.

Chapter

V

Results:

GENOME-WIDE PROMOTER

ANALYSIS OF THE SOX4

TRANSCRIPTIONAL NETWORK IN

PROSTATE CANCER CELLS

5.1 SOX4 Transcriptionally Regulates EGFR, ERBB2, TLE1, and PUMA

Expression profiling is a powerful technique that is particularly applicable to understanding the function of transcription factors. Using expression profiling to determine the genes whose mRNA levels change when SOX4 is either over expressed or eliminated using siRNA, we identified 466 transcripts which were differentially expressed according to SAM analysis (112). Among the most significantly changed, and excellent candidate SOX4 target genes, were EGFR, TLE1, CSF1 and PUMA (**FIGURE 7A**). Bioinformatic analysis of the promoter and first intron of TLE1, CSF1, PUMA and the EGFR family of receptors using CONFAC software (87) revealed the presence of potential SOX4 binding sites within the upstream promoter and first intron of each gene. We also identified possible SOX4 binding sites for the EGFR family member ERRB2. CONFAC functions by identifying the conserved sequences in the 3 kb proximal promoter region and first intron of human-mouse ortholog gene pairs and then detects transcription factor binding sites (TFBS), defined by position weight matrices from the MATCH software (91), which are conserved between the two species (87). Despite the requirement of conservation, we found anywhere from three to seven potential SOX4 binding sites per gene.

Limited commercial antibodies exist for SOX4 and show varied activity in immunoblots. However, in our hands, none of them have been useful in a ChIP assay. Therefore, we used epitope-tagged SOX4 to perform the immunoprecipitation, as described in other SOX4 ChIP studies (109, 112). Although the FLAG-SOX4 protein was not tested directly for activity, a similarly epitope tagged, glutathione S-transferase (GST)-SOX4 construct showed binding to a known SOX4 motif and not a control motif
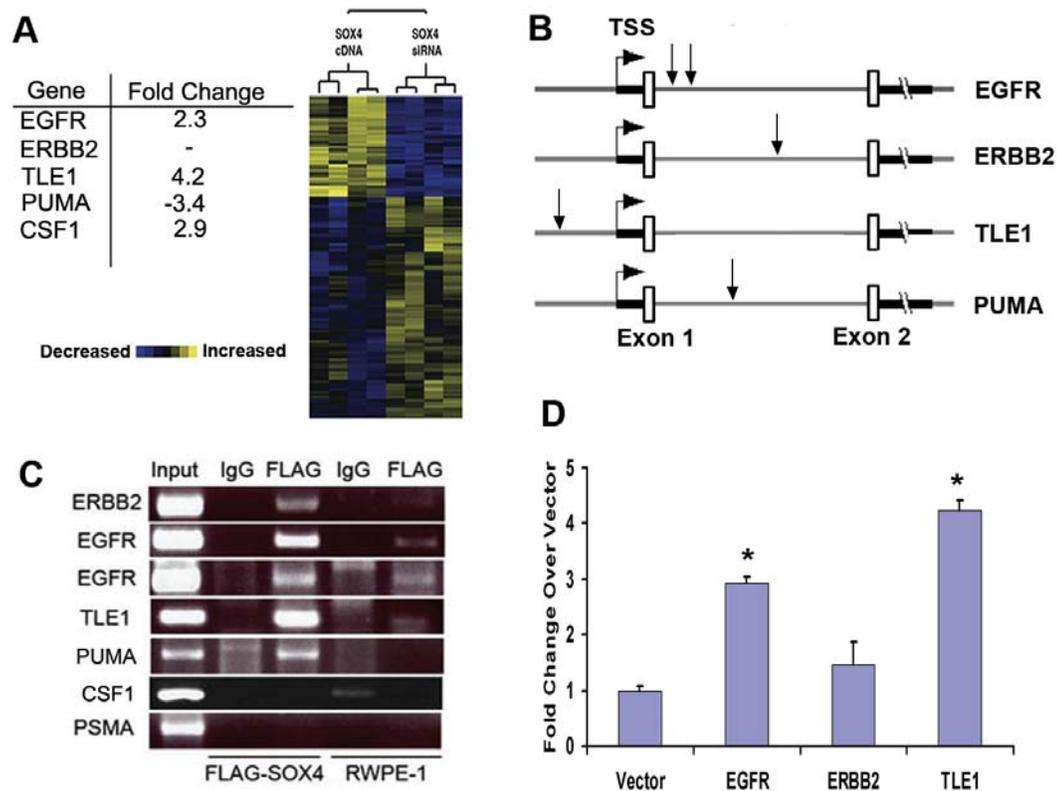
Figure 7: **(A)** Affymetrix U133A GeneChip microarray analysis of *SOX4* over expression and knockdown in LNCaP prostate cancer cells. Over expression of *SOX4* leads to increased *EGFR, TLE1,*and *CSF1* and decreased *PUMA* expression while siRNA knockdown of *SOX4* results in decreased expression. Table denotes expression changes. **(B)** Schematic showing the location of the *SOX4* binding sites for *EGFR, ERBB2*, *TLE1* and *PUMA* genes. Arrows denote location of the *SOX4* binding site. **(C)** ChIP assay of FLAG-SOX4 bound to the sequences of the *EGFR, ERBB2, PUMA* and *TLE1* proximal promoters. The SOX4 was not bound to the *CSF1* promoter and *PSMA* is shown as a negative control. *SOX4* bound DNA is specifically amplified in the FLAG IP lane from FLAG-SOX4 expressing cells (lane 3) and not control cells (lane 5) or with a non-specific antibody (lanes 2 and 4).

Figure 7: **(D)** Luciferase reporter assays with *SOX4* binding sites showing activation in the presence of *SOX4* compared to empty vector. * indicates *P*-value less than 0.01 by students T-test and error bars indicate 1 SD (n = 3 biological replicates).

(**FIGURE 11B**), validating that the epitope tag does not interfere with SOX4 DNA binding activity. To determine if SOX4 bound directly to the predicted binding sites for our candidate genes, we performed ChIP analysis on RWPE-1 prostate cancer cells stably infected with FLAG-SOX4 or a control lentiviral vector. DNA representing the predicted SOX4 sites was specifically amplified from the FLAG-SOX4 cell line and not from the control cell line, indicating that SOX4 binds to proximal promoter sequence within the first intron of PUMA, EGFR and ERBB2 and upstream of the TLE1 transcription start site (TSS) (**FIGURE 7B AND 7C**). Despite the frequency of predicted SOX4 binding sites, we were only able to verify one site in each gene, with the exception of EGFR in which two binding sites were identified. This result shows that bioinformatics can predict the presence of a potential binding sequence, yet despite employing stringent criteria, do not necessarily predict the possibility of *in vivo* occupancy.

To characterize the transcriptional effect of SOX4 levels on the regions bound in ChIP assays, the amplified ChIP fragments were cloned in front of a minimal promoter luciferase reporter plasmid and tested in transient transfections in LNCaP cells. Compared with a vector control, SOX4 significantly increased transcription of the EGFR fragment 3-fold and the TLE1 fragment roughly 4-fold. Although not found significant, ERBB2 was activated 1.5-fold compared with the vector control (**FIGURE 7D**).

Consistent with microarray data, SOX4 transcriptionally activates the EGFR and TLE1

enhancers in prostate cancer cells.

5.2 Genome-wide Localization Analysis

We were able to identify five genomic SOX4 binding sites that regulate four

genes, but our success rate of validating bioinformatically predicted sites was low.

Recent advances in microarray technology allow the coupling of ChIP assays to tiling

DNA microarrays (ChIP-chip) for high-throughput identification of all the SOX4 bound

regions. Again, we employed an epitope tagged SOX4 to facilitate immunoprecipitation

but instead used a hemaglutin (HA) tag, which shows less background in our hands than

the FLAG epitope tag. We cloned HA-SOX4 into the pHR-UBQ-IRES-eYFP-ΔU3

lentiviral vector and stably infected LNCaP and RWPE-1 prostate cancer cells (**FIGURE**

8A). The vector contained an ubiquitin promoter (UBQ) to express SOX4 at close to

physiological levels and an internal ribosomal entry site (IRES) followed by the Yellow

fluorescent protein (YFP) to allow isolation of a near pure population of infected cells by

fluorescence activated cell sorting (FACS) (**FIGURE 8B**). For each cell line, we created

a YFP only control line and a HA-SOX4-IRES-YFP experimental cell line. HA-SOX4

expression and activity was demonstrated by immunoprecipitations (IPs) with the

12CA5, monoclonal HA-antibody. Our results show we could specifically IP SOX4

from the HA-SOX4 cell line and not the YFP only control line (**FIGURE 8C**). Finally,

the two cell lines were tested in a ChIP assay using the known TLE1 binding site as a

positive control. The TLE1 binding site was specifically amplified in the LNCaP-HA-

SOX4 specific IP and not from the control LNCaP-YFP cell line or from control IPs

(**FIGURE 8D).**

A

5'LTR —Ψ— P-UBQ —— IRES-eYFP —— ΔU3-LTR

5'LTR —Ψ— P-UBQ — HA-SOX4 —— IRES-eYFP —— ΔU3-LTR

B

C

D
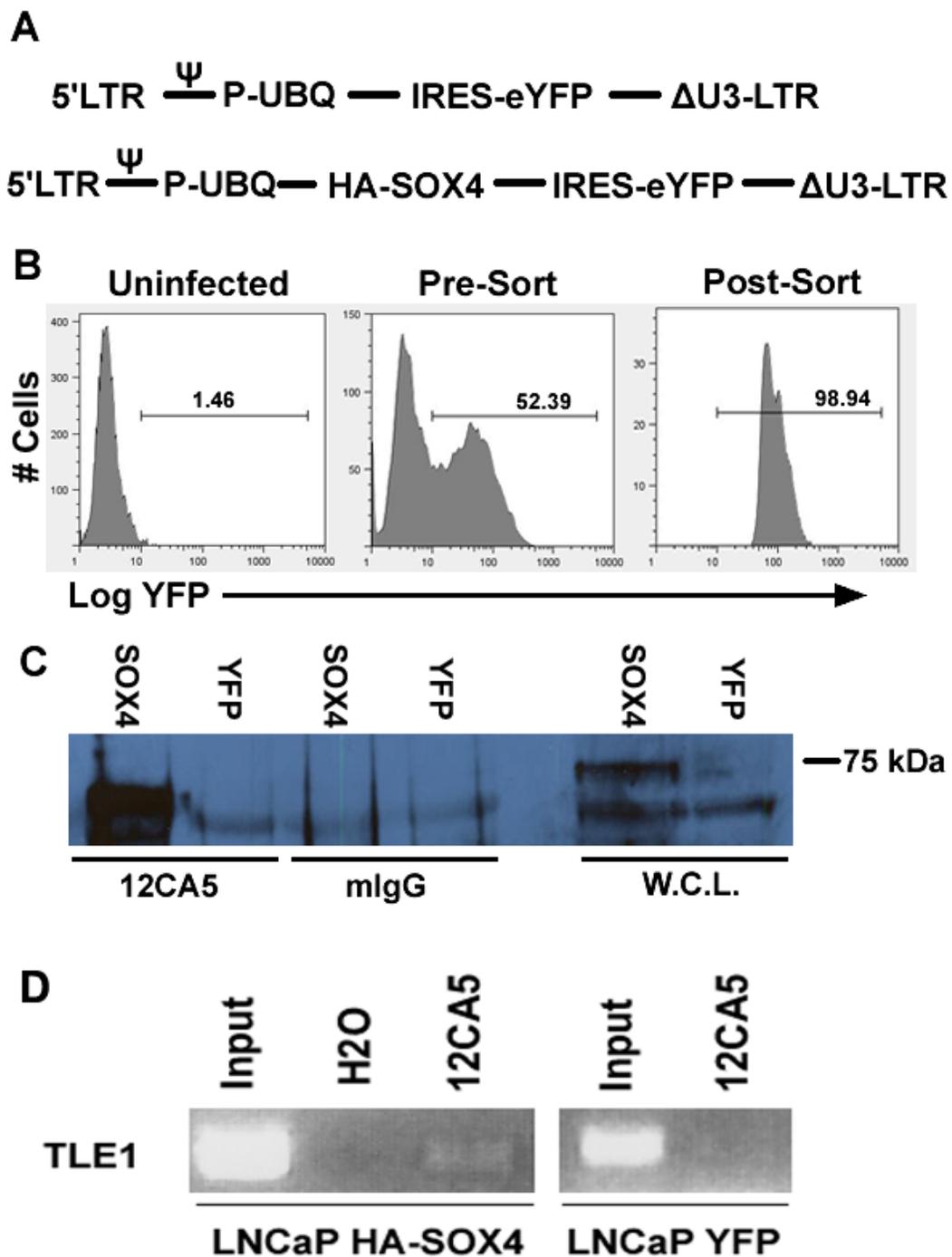
Figure 8: **(A)** Schematic diagram of the lentiviral constructs used to stably infect

LNCaP and RWPE-1 prostate cancer cells showing the locations of LTRs and

promoters. The top figure represents the control, eYFP only construct, and the

lower figure represents the HA-SOX4 construct. **(B)** Histogram showing the

control uninfected, pre-sorted and post-sorted cell populations. X-axis displays

YFP signal intensity. **(C)** Immunoblot showing that HA-SOX4 is expressed and

specifically immunoprecipitated from the LNCaP-HA-SOX4 cell line and not the

control LNCaP-YFP cell line (upper band). **(D)** ChIP assay of SOX4 binding to a

known locus in the TLE1 promoter. TLE1 promoter sequence is specifically

amplified from the LNCaP-HA-SOX4 cell line and not the control LNCaP-YFP

only cell line.

To determine the direct SOX4 target genes on a global scale, we performed ChIP

assays in triplicate from the LNCaP-HA-SOX4 stable cell line and in duplicate from the

control LNCaP-YFP cell line. Input and immunoprecipitated DNA were sent to

NimbleGen and hybridized to their dual-chip, 25K promoter array. This array queries

roughly 3 kb up-stream and 1 kb downstream of the TSS, focusing on proximal promote

elements of 25,000 known transcripts. Raw signal values were Z-score normalized,

ratios of immunoprecipitated over input DNA were determined and ChIPOTle software

(28) was applied to each data set to identify enriched peaks. ChIPOTle software uses a

sliding window to search neighboring probe sets for enrichment above the overall

background. Peaks ($P < 0.001$) that overlapped in at least two of the three data sets and

were not present in the LNCaP-YFP cell line were called significant (**FIGURE 9A**).

Based on these variables, we classified 3,600 significant, overlapping peaks as SOX4

target sequences (**APPENDIX A**). We mapped the 3,600 peaks to 3,470 different genes.

It was not possible to assign a unique gene to every peak because some TSS are quite

close to each other (< 3 kb), and many genes had multiple peaks in their promoters

(**APPENDIX A**).

To verify the set of 3,600 SOX4 peaks, 28 candidate SOX4 target sites

representing a range of *P* values in promoters of genes of biological interest were chosen,

primers were designed around the peaks and enrichment verified by conventional ChIP.

Ten of the 28 candidates were analyzed by ChIP-qPCR and 18 by ChIP-PCR. Overall,

24 of 28 (86%) of the candidate targets were confirmed, validating our data set. All 10 of

the peaks chosen to validate by qPCR were reproducibly enriched over the YFP control

in both the LNCaP-HA-SOX4 cell line and the RWPE-1-HA-SOX4 cell line

(**FIGURE 9B**). Of the target sites validated by conventional PCR, 14 of 18 genes were

Figure 9: **(A)** Graph showing enrichment in the three HA-SOX4 lanes over the

average of the two YFP replicates for the *SOX4* target gene *FMO4*. Y-axis is the

signal intensity across the genomic coordinates on the X-axis. **(B)** qPCR analysis of

10 randomly selected genes verified in both the RWPE-1 and LNCaP cell lines.

Graph shows fold enrichment of the HA-SOX4 IP over the YFP negative control IP.

Error bars indicate 1 SD (n = 3 biological replicates. **(C) and (D)** Genes that were

verified by conventional ChIP assay. HA-SOX4 and YFP cells were subjected to

conventional ChIP followed by PCR in both the LNCaP **(C)** and RWPE-1 **(D)** prostate

cell lines. Six genes verified in the LNCaP cell lines and five in the RWPE-1 cell

lines.
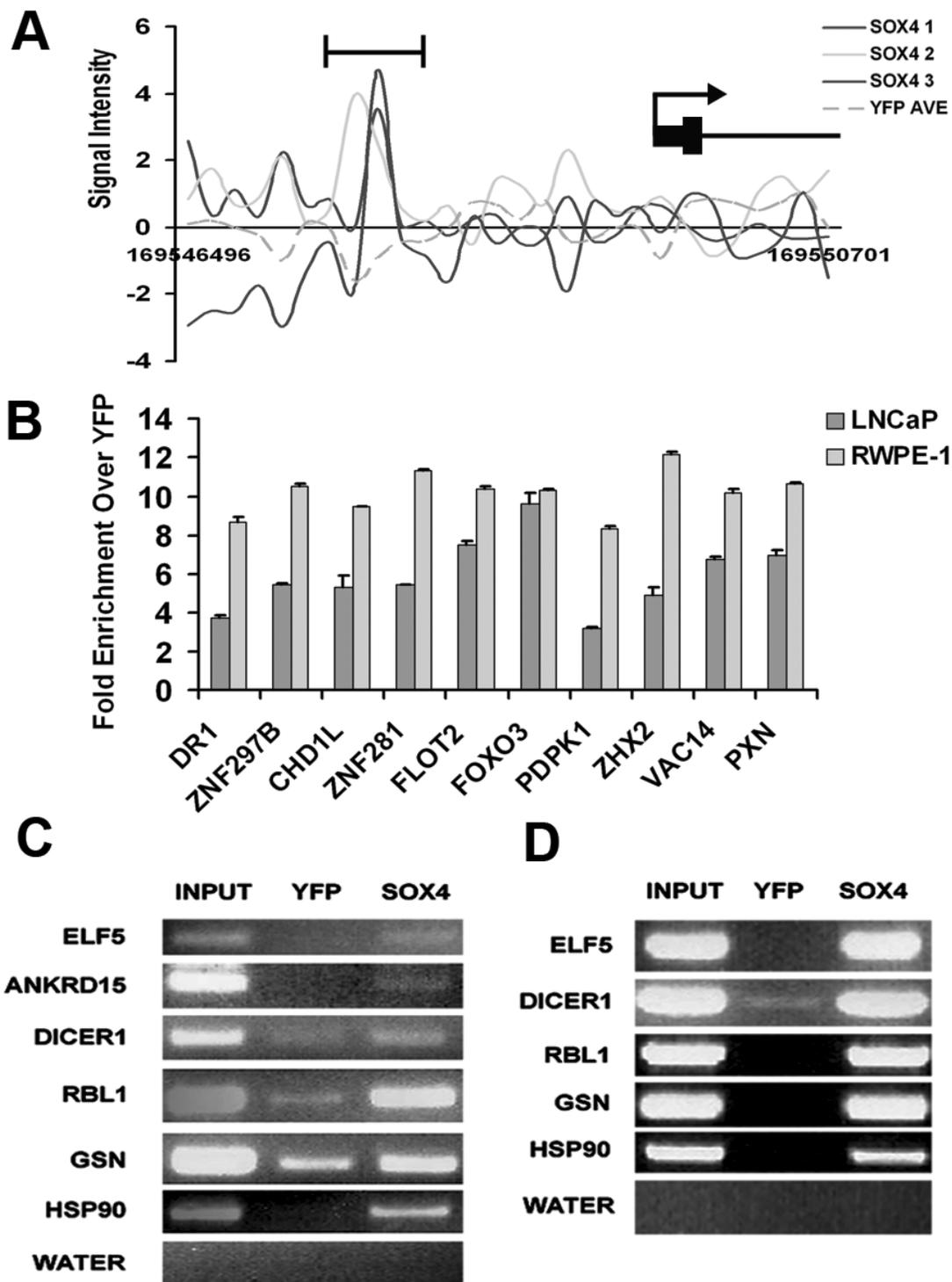
confirmed in both the LNCaP and RWPE-1 cell lines, whereas a mock, control PCR was negative (**FIGURE 9C AND D**). One peak in the promoter of ANKRD15 was enriched and validated only in the LNCaP cell line and not in the RWPE-1 line.

5.3 Target Gene Expression Analysis

To determine if SOX4 binding affects transcription of the 3,470 genes that have SOX4 bound at their promoters, we performed whole genome expression analysis on LNCaP cells after transfection with SOX4 or a control vector. To increase likelihood of identifying direct SOX4 targets, total RNA was isolated at a relatively early time point (24 hours post-transfection) and hybridized to Illumina Human 6-v2 whole genome arrays. Following normalization, SAM analysis identified a total of 1,766 genes that were significantly changed at least 1.5-fold with a false discovery rate of 0.749% (**FIGURE 10A AND APPENDIX B**). Of the 1,766 genes, 244 were also identified as direct SOX4 targets by ChIP-chip analysis (**FIGURE 10B AND APPENDIX C**). We confirmed the expression changes of seven of these genes following SOX4 transfection by qPCR (**FIGURE 10C**).

Previous expression profiling of LNCaP cells after SOX4 siRNA knockdown (112) identified 466 downstream targets, and we confirmed that SOX4 regulates the expression of DICER, DLL1 and HES2 in LNCaP cells by qPCR (**FIGURE 10C**). In addition, we confirmed SOX4 regulation of DICER at the protein level (**FIGURE 10D**). Of the 466 candidate targets, 47 genes overlapped with the 3,470 ChIP-chip targets (**APPENDIX C**), increasing the number of direct SOX4 targets to 282 genes (**FIGURE 10B AND APPENDIX C**). These 282 genes bound by SOX4 in ChIP-chip and significantly changed by expression profiling were classified as high-confidence direct
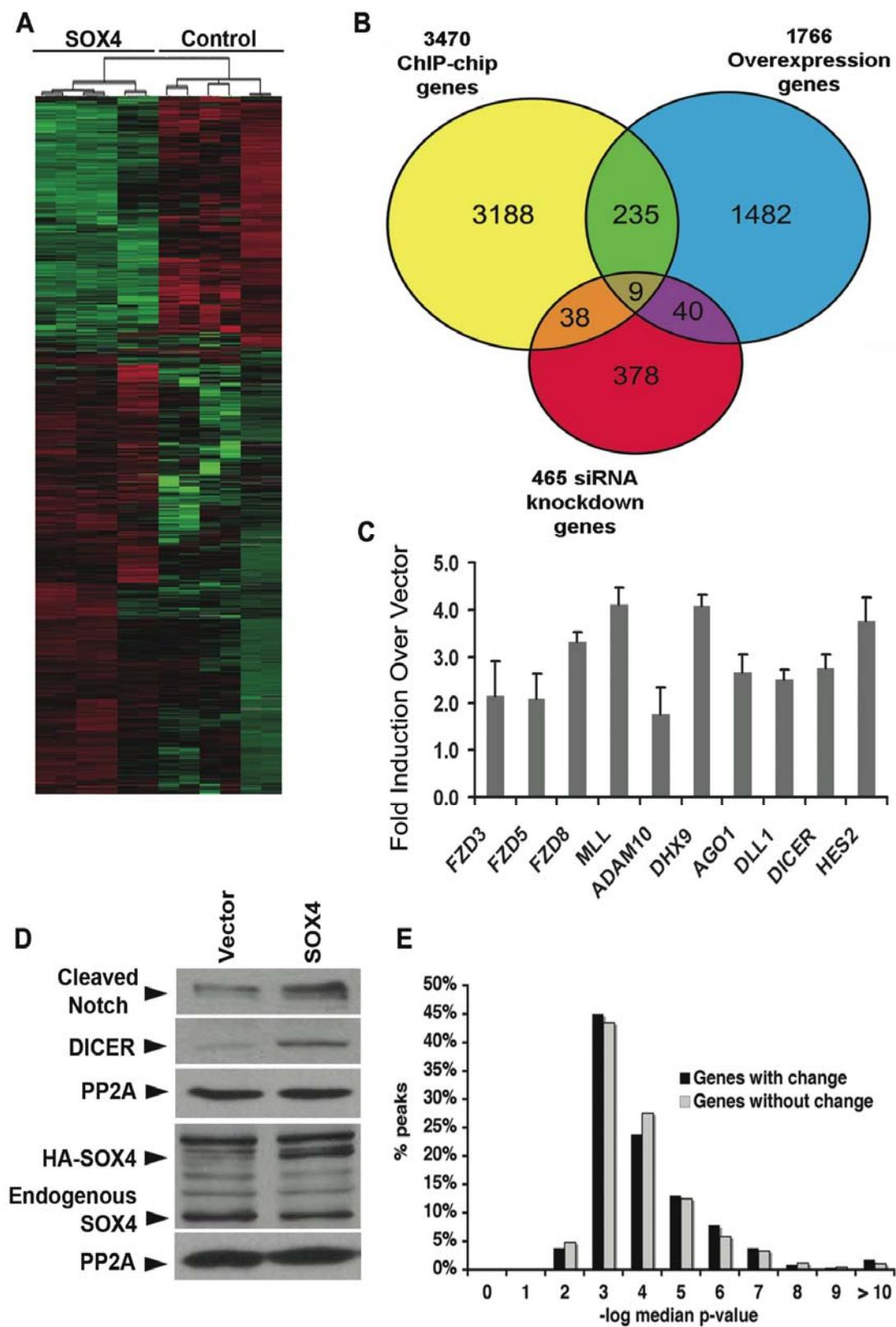
Figure 10: (A) Heat map illustrating Illumina expression data of the 1,766 significant

genes as determined by SAM analysis. Red indicates over expressed and green

denotes under expressed genes. **(B)** Venn diagram showing the overlap between 3,470

ChIP-chip *SOX4* direct target genes, the Illumina expression data set of 1,766 genes,

and the Affymetrix expression dataset of 465 genes. **(C)** qPCR data of *SOX4* direct

target genes after *SOX4* over expression in LNCaP cells. All seven genes were

upregulated over a vector control transfection, similar to values determined by the

Illumina array with a *P*-value less than 0.005 by students T-test. Error bars indicate 1

SD (n = 3 biological replicates). **(D)** *DICER* protein expression and cleaved, activated

NOTCH1 is upregulated by *SOX4*. HA-SOX4 or vector control was transfected into

LNCaP cells and immunoblots were probed for *DICER*, *SOX4*, cleaved *NOTCH1*, and

PP2A as a loading control. **(E)** Graph depicting the genes identified by ChIP-chip,

split into ten bins based on the median –log10 of the *P*-value for the three replicates.

Shown are the distributions as a percentage of the total number of genes for those that

were significantly changed in expression profiling and those that showed no

significant change. No clear *P*-value is evident to separate the two gene sets.

SOX4 target genes. Nine genes (PIK4CA, DHX9, BTN3A3, CDK2, MVK, ADAM10,

RYK, ISG20 and DBI) overlapped in all three data sets. DLL1 and ADAM10 can

directly activate NOTCH signaling through ligand binding and proteolytic cleavage

respectively (24). We transfected LNCaP cells with either SOX4 or a control expression

vector and probed immunoblots for activated, cleaved NOTCH1. Only in the SOX4

transfected LNCaPs was cleaved NOTCH1 detected, indicating SOX4 stimulates

activation of this pathway **(FIGURE 10D)**. Interestingly, the transcription factor SON

and purine biosynthetic enzyme GART, two genes on chromosome 21 that are transcribed in opposite directions and regulated by a bidirectional promoter, were affected in opposite ways. SON was activated by SOX4 1.8-fold, as detected by SOX4 over expression, whereas GART was increased almost 3-fold as determined by SOX4 siRNA knockdown, suggesting that SOX4 contributes to the regulation of this bidirectional promoter.

We next analyzed the peaks in our ChIP-chip data set, comparing the *P*-values of the genes that were altered by transient over expression of SOX4 with those that were not and found no difference in the distributions (**FIGURE 10E**). Based on our ChIP-chip validation experiments and the similar *P*-value distributions, we concluded that SOX4 is genuinely bound at the promoters of the 3,188 genes that did not change but that SOX4 by itself is not limiting or sufficient to generate changes in transcription without corresponding changes in the cellular context, such as activation of other transcription factors or signaling pathways.

5.4 Novel SOX4 Position-Weight Matrix

We determined the DNA binding preferences of SOX4 using universal Protein-binding microarrays (PBMs) (20) to facilitate computational analyses of SOX4 DNA binding sites. This universal PBM array allows recombinant SOX4 protein to interact with and bind to every possible 8-mer sequence, allowing *in vitro* binding site specificities to be calculated. This powerful technology has already been applied to each mouse homeodomain transcription factor (19). We generated an NH$_2$-terminal, GST-SOX4-DBD fusion protein consisting of amino acids 1-340, expressed, and purified it from *E. coli* (**FIGURE 11A**). To test our GST-SOX4-DBD for activity, we used a

63
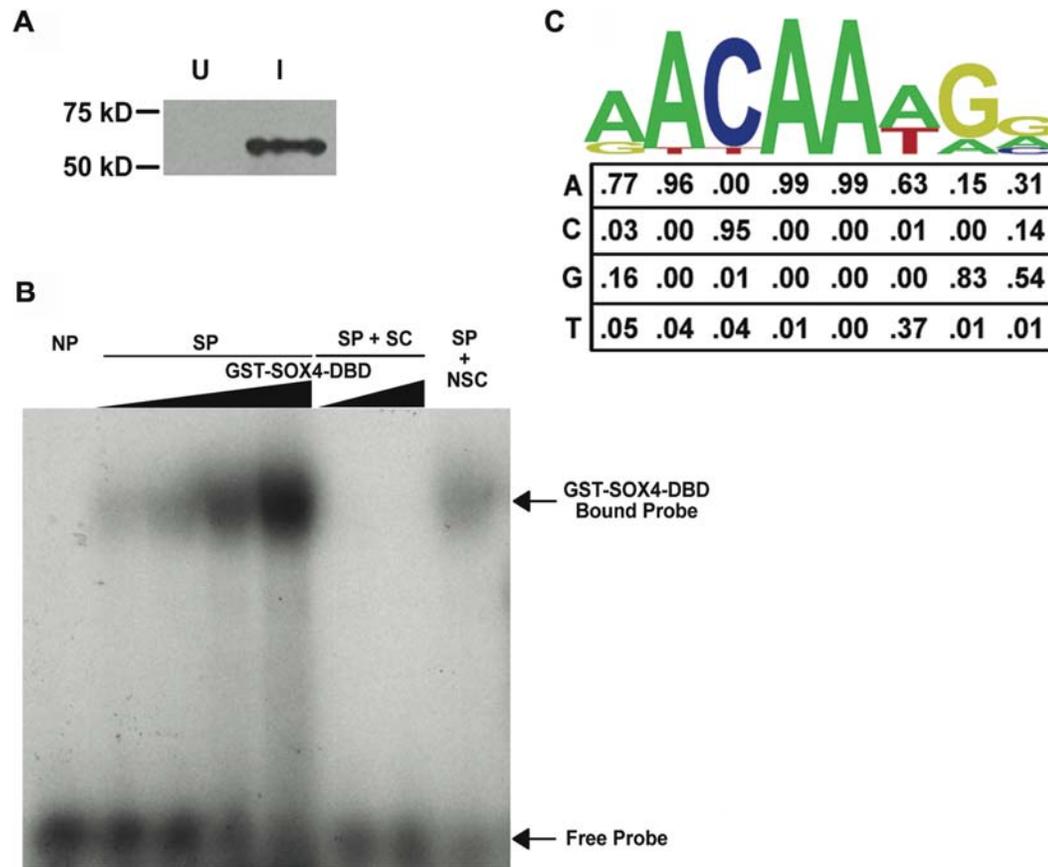
Figure 11: **(A)** SDS-PAGE gel of GST-SOX4-DBD from an IPTG uninduced (U) or induced (I) bacterial cells. **(B)** EMSA assay of recombinant GST-SOX4-DBD binding to a known SOX4 binding motif centered on a 35-mer oligo. NP – no protein, SP – specific probe, SC – specific cold competitor probe, NSC – non-specific cold competitor probe. **(C)** PBM-derived 8-mer PWM for *SOX4* displayed both graphically and numerically for each base position derived from incubation of recombinant GST-SOX4-DBD with a universal 'all 8-mer' double-stranded DNA PBM.

known SOX4 binding motif found inside the EGFR enhancer in an electromobility shit assay. GST-SOX4-DBD caused a shift when incubated with the specific radiolabled

oligo but not with a mutated oligo, and the shift was abolished when incubated with unlabeled-specific oligo (**FIGURE 11B**). We concluded that the GST-SOX4-DBD is functional despite the NH$_2$-terminal GST epitope tag and truncated protein. The GST-SOX4-DBD was incubated with the protein binding microarray and a novel PWM (RWYAAWRV) was calculated from the PBM data (**APPENDIX D**) using the Seed-and-Wobble algorithm (**FIGURE 11C**)(20) where R is A or G, W is A or T, Y is C or T and V represents A, G or C. Three groups have previously reported similar binding site sequences for SOX4: AACAAAG (189), AACAAT (198), and WWCAAWG (109). Our PWM confirms the SOX4 core binding sequence of the previously known binding sites. Our study queried every possible 8-mer whereas earlier groups used no more than 31 sequences to develop their binding motif. This may account for the differences in the specificity at the 1st and 7th positions where we found a bias towards A or C, and at the 8th position where we found a bias towards G.

5.5 SOX4 Peaks Contain SOX4 Binding Sites

We employed two different methods to analyze our ChIP-chip peaks for the presence of SOX4 binding sites. Firstly, using our newly derived PWM, we applied CONFAC software (87) to analyze the enriched sequences for the presence of SOX4 binding sites. We analyzed the sequences of the peaks in the promoters of our 282 high-confidence genes against 10 sets of control promoter sequences to see if SOX4 sites were enriched in our target gene set. Control promoter peaks of equal size to SOX4 peaks were chosen randomly from sequences covered by the NimbleGen array, and each control set contained equal total sequence coverage as our 282 high-confidence peaks. Using stringent criteria (core similarity, > 0.85; matrix similarity, > 0.75), we found that

60% of the SOX4 peaks contained SOX4 binding sites. We also found that SOX4 sites were significantly enriched relative to 10 sets of random promoter sequence by Mann-Whitney U test using Benjamini correction for multiple hypothesis testing (q < 0.0019).

To further characterize the SOX4 binding sites, we secondly searched the entire set of 3,600 SOX4 peaks and 10 equal sets of random promoter sequence for the presence of PBM-bound k-mers. A k-mer is an ungapped 8-mer sequence that was bound by SOX4 in the PBM assay. The specificity of PBM k-mers can be quantified by the enrichment score (ES), which ranges from -0.5 to 0.5 (19). We analyzed the enrichment of PBM k-mers with 0.45 > ES > 0.40 (moderate) and ES > 0.45 (stringent). Using two-tailed Mann-Whitney test, SOX4 peaks contained significantly more stringent ($P =$ 0.0002) and moderate ($P = 1.08$ x 10^{-5}) k-mers than random promoter sequence (**FIGURE 12A**).

To investigate the possible interaction with protein partners that may increase SOX4 affinity for poor matching sites *in vivo*, we searched for enrichment of co-

Transcription Factor	Family	Benjamini Corrected q-value
E2F4	E2F	1.78E-11
E2F1	E2F	3.06E-11
PAX5	Paired Box	2.07E-10
WHN	Forkhead	2.94E-10
SMAD3	SMAD	1.82E-09
SMAD4	SMAD	3.33E-09
MYC	MYC	6.25E-09
NFKAPPAB	NF-κB	2.95E-08
LEF1/TCF1	TCF/LEF	1.12E-06

Table 5: Benjamini corrected q-values for co-occurring transcription factor binding sites.

occurring TFBS in the SOX4 peaks. We applied CONFAC software to search our data

for the presence of co-occurring TFBS within the same peak (**TABLE 5**). Using the

same criteria as above (core similarity, > 0.85; matrix similarity, > 0.75), we determined

that the E2F family had the most frequently co-occurring motif (similar to TTTCGCGC,

$q = 1.78 \times 10^{-11}$). Interestingly, Ingenuity Pathway Analysis (IPA) identified cell cycle as

a functionally enriched process in the 3,470 SOX4 target genes ($P = 0.00916$), suggesting

Entrez ID	Symbol	Microarray Fold Change
196528	ARID2	1.99
2001	ELF5	-2.65
3169	FOXA1	-2.47
2976	GTF3C2	-3.12
64412	GZF1	2.42
84458	LCOR	2.41
4173	MCM4	1.55
58508	MLL3	2.06
10933	MORF4L1	2.07
8031	NCOA4	2.64
4784	NFIX	-2.83
4824	NKX3-1	-4.53
7799	PRDM2	2.48
5933	RBL1	1.80
55509	SNFT	-2.32
6722	SRF	-2.03
54816	SUHW4	-1.93
9412	SURB7	-2.24
9338	TCEAL1	-1.57
7718	ZNF165	1.53
7738	ZNF184	1.66
23528	ZNF281	1.71
30834	ZNRD1	-1.63

Table 6: DAVID analysis identified 23 transcription factors present in our high-confidence SOX4 target genes. GO Term: transcription, DNA dependent ($P = 3.7 \times 10^{-18}$).

that part of SOX4's function is to control the expression of genes involved in cell cycle progression. CONFAC analysis identified other significant TFBS motifs enriched in the SOX4 peaks (**TABLE 5**), including those for transcription factors in the TGFβ, WNT, and NF-κB pathways. SOX4 modulates WNT signals via interaction with β-catenin to activate TOP-Flash luciferase reporters (167). We investigated this finding in LNCaP cells and found that SOX4 synergizes with constitutively active β-catenin to activate a TOP-Flash reporter
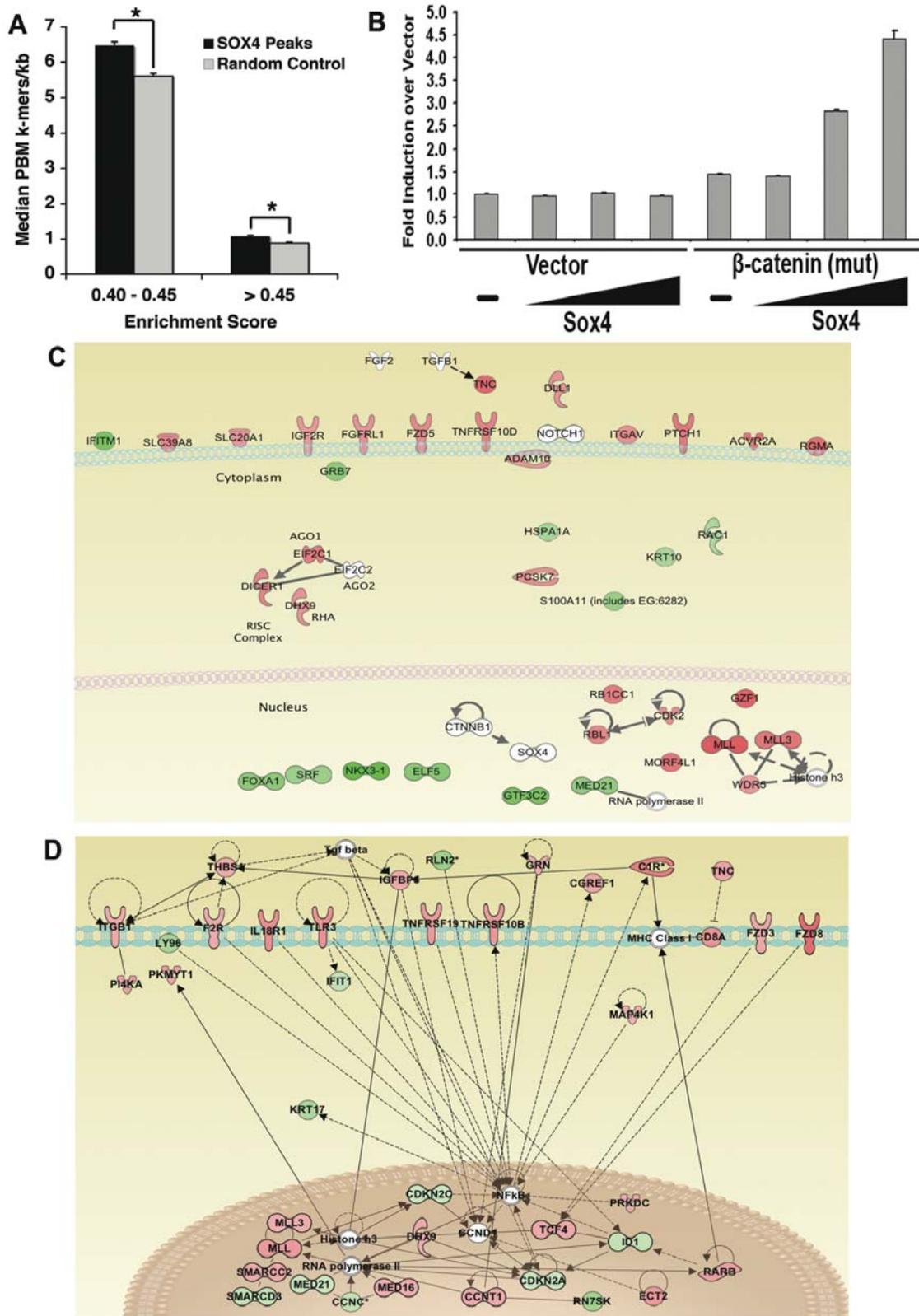
Figure 12: **(A)** Graph of median frequency of PBM k-mer hits in SOX4 ChIP-chip peaks (n = 3600) and random control sequences (n = 3600) for moderate (0.45 > ES > 0.40) and stringent (ES > 0.45) SOX4 k-mers. Significant differences are indicated with an asterisk (*) for stringent k-mers (*P* = 0.0002) and moderate k-mers (*P* = 1.08 x 10^{-5}). **(B)** Luciferase assay of LNCaP cells transfected with either a vector control or 100, 200, or 300 ng of a SOX4 expression vector. LNCaP cells were also co-transfected with either a vector control or the β-catenin S33Y constitutively active mutant. All cells were transected with the TOP-Flash luciferase reporter and luciferase activity was measured 24 hrs post-transfection. SOX4 does not function alone but instead cooperates with β-catenin to activate the TOP-Flash reporter in a dose dependent manner. **(C)** IPA analysis of direct target genes illustrating the *SOX4* transcriptional network. *SOX4* regulates a host of nuclear and membrane localized proteins as well as multiple components of the RISC complex. Red indicates target genes upregulated by *SOX4*, green denotes downregulated genes and white represents genes for which no expression change was detected. **(D)** IPA analysis of Illumina expression genes changed at least 2-fold by SAM analysis. SOX4 regulatory targets include a host of membrane and nuclear proteins. Red indicates genes upregulated by SOX4 over expression and green denotes downregulated genes.

plasmid, further highlighting a role for SOX4 in the WNT pathway in prostate cancer (**FIGURE 12B**).

5.6 Ontology and IPA Analysis

We performed a gene ontology analysis using DAVID software (40) on the 282 high-confidence SOX4 targets to determine the biological processes and functions of the SOX4 targets. Among the SOX4 targets were 23 transcription factors (**TABLE 6**), and DAVID analysis determined that the top functional annotations were transcription ($P =$ 3.7×10^{-18}), transmembrane ($P = 5.59 \times 10^{-10}$), and protein phosphorylation/dephosphorylation ($P = 3.5 \times 10^{-18}/6.6 \times 10^{-7}$). These findings are paralleled by expression profiling of SOX4 over expression in HU609 bladder carcinoma cells where top annotated functions were signal transduction and protein phosphorylation (1).

Using commercial IPA software, we identified biological pathways and functions that are enriched in our 282 high-confidence targets, 1,766 significant genes identified by SAM analysis, and the 3,470 unique genes that had SOX4 bound at their promoters in ChIP-chip. As anticipated, among the most significant annotations were cell cycle, cancer, and tissue development. In the significant expression data set of 1,766 genes, we observed an up-regulation of three Frizzled receptors, FZD3, FZD5 and FZD8, as well as the downstream transcription factor TCF3. Overall, IPA analyses discovered key components of the EGFR, Notch, AKT-PI3K, miRNA, and WNT-β-catenin pathways as SOX4 regulatory targets. Based on these findings, we built SOX4 regulatory networks found in prostate cancer cells (**FIGURE 12C AND 12D**). SOX4 target genes comprise key pathway components, such as ligands (DLL1 and NGR1), receptors (FZD5 and

PTCH1), an AKT regulatory kinase (PDPK1), and downstream transcription factors

(FOXO3 and HES2). We found that SOX4 activates expression of Tenascin C, an

extracellular matrix protein that is a target of TGFβ signaling (141) and β-catenin (17).

In addition, SOX4 regulates three components of the RNA-induced silencing complex

(RISC), DICER1, Argonaute 1 (AGO1) and RHA/DHX9 (**APPENDIX C**). We

confirmed these data by qPCR (**FIGURE 10C**) and western blot for DICER (**FIGURE

10D**).

Gene set enrichment analysis (GSEA)(174) and GSEA Leading Edge analysis

(173) of these gene sets identified TGFβ induced SMAD3 direct target genes (**TABLE 7**)

as enriched in the SOX4 target genes. SOX4 is up-regulated by TGFβ-1 treatment (158,

197), and we found SMAD4 sites are significantly enriched in the SOX4 ChIP-chip peaks

(**TABLE 5**). This suggests that SOX4 may affect key developmental and growth factor

signaling pathways in prostate cancer cells at both the transmembrane signaling and

transcriptional levels.

TGFβ induced target Genes
IGFBP5, THBS1, TNFRSF1A, BSG, EPHB4, JUP, ITGB5, TNC, SEMA3F, COL6A1, LAMB1, EFNA5, MMP11, DSP, DVL1, EPHB3, SMO, IGF2R, ITGA5, CTNNB1, IGFBP2, NEO1

Table 7: GSEA Leading Edge analysis of genes induced by SOX4 over expression.

Significantly expressed genes include 22 TGFβ target genes.

Chapter

VI

SOX4:

Discussion

Many studies have identified SOX4 as a crucial developmental transcription factor that is often over expressed in many types of malignancies, but little is known of the genes SOX4 regulates in cancer and normal cells. We have used a ChIP-chip approach to perform the first genome-wide localization analysis of SOX4. We mapped 3,600 binding peaks that represent a possible 3,470 unique genes possibly under the transcriptional regulation of SOX4. We have also identified 1,766 genes that respond to increased SOX4 levels by whole genome expression profiling on an Illumina platform. Integration of these two data sets with previous expression profiling of SOX4 mapped 282 high-confidence direct targets in the SOX4 transcriptional network. In addition, we have used unique protein-binding microarrays to determine a novel SOX4 specific PWM and have shown that our ChIP-chip peaks are significantly enriched for SOX4 binding sites. This chapter will discuss the insights these data provide into the roles that SOX4 plays in the normal cell and in cancer.

6.1 SOX4 Direct Target Genes

Our ChIP-chip experiment identified 3,600 genomic binding sites for SOX4, but only 10% of the significant, differentially expressed genes identified by expression microarray overlapped with the ChIP-chip data. Although a portion of the genes present in our expression data set could be regulated by factors under the control of SOX4 and are indirect targets, the 282 high-confidence gene list is likely a conservative estimate. The NimbleGen 25K promoter array only queries proximal promoter sequences with an average of 3 kb upstream and 1 kb downstream of the TSS. The total coverage of the array is 110 Mb and represents only a fraction of the entire genome. In fact, the first genomic binding site reported for SOX4 was to the CD2 enhancer, located 3' to the polyadenylation signal (198) and is outside the coverage of the NimbleGen arrays. We also found that SOX4 binds EGFR, ERBB2 and PUMA in the first intron over 20 kb downstream of the TSS and the TLE1 binding site was located further upstream of the TSS than the array tiled. Unsurprisingly, we did not detect these genes in our ChIP-chip experiment. Despite the limitations of the NimbleGen arrays, one advantage to tiling proximal promoter sequences is that genes regulated by a particular binding site are easily identified. By their nature, enhancers can be tissue specific and function in a position and orientation independent manner. Intensive functional studies in a variety of tissues would be required to identify the gene or set of genes regulated by SOX4. Previous analyses have shown that the majority of TFBS in the TRANSFAC database cluster in the 2-3 kb around the TSS (27). Therefore we are confident we have identified the majority of SOX4 binding sites in LNCaP cells.

Whereas 3,600 is a large number of SOX4 bound regions, there is the possibility that some peaks are false positives. Nevertheless, we were able to validate 24 of 28 (86%) candidate binding sites chosen, adding confidence to our data set. In fact, an even higher number of over 4,200 genomic binding sites have been reported for c-MYC in ChIP-positron emission tomography (ChIP-PET) whole genome studies (206). A similar promoter focused study identifying SOX2 binding sites in embryonic stem cells found 1,228 genes with SOX2 bound promoter regions. Thus, our list of 282 high-confidence targets may be a low estimate of the true SOX4 regulatory network and more of the 1,900 genes that responded to changes in SOX4 mRNA levels (but were not detected by ChIP-chip) may be direct targets. Excellent candidates would be the 40 genes that responded to SOX4 on both microarray platforms, such as the IL6 receptor, SOX12, and NME1 (**APPENDIX C**). Whole genome tiling arrays or ChIP-seq studies could provide additional binding sites that may show more overlap with the Illumina and Affymetrix expression data sets.

Many of the bound regions did not respond to changes in SOX4 mRNA levels. This suggests they may not be regulated by SOX4 alone but by multi-protein activator complexes, of which SOX4 is only one component. Furthermore, the stability of SOX4 bound to a promoter could be greater than unbound SOX4, limiting the effects observed by siRNA knockdown. Experiments to date have not investigated the regulation of SOX4 or determined the half-life of the protein. In different cell types or cellular contexts, SOX4 may activate a different subset of these genes. Of the 31 SOX4 target genes previously confirmed by ChIP assays (109), only six are represented in our NimbleGen data set and three found to be changed in our expression profiling data sets.

Three previous studies have used expression profiling to discover SOX4 regulated genes and in each study there is little overlap between the hundreds of genes identified (1, 148, 197). None of these studies used ChIP assays to identify bona fide genomic binding sites, thus it is not possible to determine if the genes are direct or indirect SOX4 targets.

There may be a number of reasons for the small overlap seen between data sets in the studies. Firstly, three of the four expression profiling experiments were performed in different human cancer cell lines and one was performed using mouse pancreatic tissue. The target genes previously identified by ChIP assay (109) were in a hepatocellular carcinoma cell line, and we have examined prostate cancer cells. This suggests that SOX4 regulates distinct sets of genes in a cell line and tissue specific manner. We have seen similar results comparing SOX4's regulation of EGFR in the non-transformed RWPE-1 prostate cell line and the fully transformed LNCaP prostate cancer line. ChIP assays confirmed SOX4 binding to intronic sequence of EGFR in RWPE-1 cells but not in LNCaP cells. It is important to note that RWPE-1 cells require EGF stimulation for growth in cell culture and may be more dependent on this pathway than LNCaP cells. Thus, SOX4 target may change within the same tissues as they undergo processes such as differentiation or cancer development. It would be interesting to perform a comparative ChIP-chip experiment in the RWPE-1 cell line to see the extent of overlap with the LNCaP data. This would provide new insights into how SOX4 target genes change during transformation and other processes. Another reason for the small overlap in data sets may be due to technical aspects of the studies including experimental design, microarray platforms used as well as data analysis and normalization techniques.

Despite the differences, ontology analysis of each data sets does reveal overall functional similarities such as SOX4's role in signal transduction (1). The overall role SOX4 plays in cellular physiology may remain the same, even though specific target genes are not conserved across tissues. Interestingly, DKK3 was one of the six genes that overlapped in both data sets of ChIP verified genes, further implicating SOX4 in the WNT pathway. SOX4 is known to interact with β-catenin and other coactivators and it may be poised at many of these promoters to enable responses to developmental signals from the WNT or TGFβ pathway.

6.2 Modulation of Input/Output Network Signals

Genetic mouse models of SOX4 over expression or deletion have revealed roles for SOX4 in differentiation and development. In the neuronal lineage it has been reported that SOX4 can up regulate the neuronal maker genes TUJ1 and MAP2, driving cells down the path of neuronal maturation (21). It is not known if these genes are direct or indirect SOX4 transcriptional targets, and to date no molecular mechanisms explaining SOX4's role in differentiation have been reported. Our data indicates that SOX4 regulates cellular differentiation through a variety of transmembrane receptors, transcription factors and developmental signaling pathways. We have shown SOX4 directly regulates EGFR and ERBB2. Other membrane receptors in the SOX4 transcriptional network include: Frizzled family members FZD3, FZD5 and FZD8, the Hedgehog receptor PTCH-1, the Notch ligand DLL1, TRAIL decoy receptor TNFRSF10D, and other growth factor receptors, such as FGFR1 and IGF2R. DAVID analysis also revealed transmembrane ($P = 5.59$ x 10^{-10}) and protein phosphorylation/dephosphorylation ($P = 3.5$ x 10^{-18}/6.6 x 10^{-7}) as enriched annotations.

SOX4 is up regulated in response to numerous external ligands from TGFβ (158) and BMP-6 (203) to parathyroid hormone and progesterone (64). Previous work has shown that SOX4 directly signals form the IL-5Rα through cytoplasmic association with Syntenin (58). These findings suggest that SOX4 can modulate cellular input signals transcriptionally through feedback loops that regulate the levels of transmembrane receptors. We propose this is the case with EGFR and although the effects seem to be cell line specific, treatment of the RWPE-1 cell line with EGF caused up regulation of EGFR protein that was abrogated when SOX4 was eliminated with siRNA.

Other regulatory networks controlling receptor transcription may exist in addition to positive feedback loops. An alternate hypothesis is that there are transcription independent functions for SOX4, as has been shown for the transcription factor p53 (143). Interestingly, one of p53's DNA-independent roles lies in regulation of apoptosis and previous reports have suggested that the glycine rich region of SOX4 (amino acids 152-227) can induce apoptosis independent of DNA binding (83). Future experiments investigating the SOX4 protein-protein interaction network will partially address this question, identifying other cytoplasmic and nuclear proteins that function in the DNA independent regulation of receptor signaling events.

SOX4 influences downstream transcriptional responses through the control of transcription factors. DAVID analysis identified transcription ($P = 3.7 \times 10^{-18}$) as an enriched biological annotation, identifying 23 transcription factors that are direct target genes of SOX4 (**TABLE 6**). Transcriptional factors regulated by SOX4 include the Forkhead protein FOXA1, the glandular epithelial Ets factor ELF5, the homeodomain tumor suppressor NKX3.1 and the cell cycle regulator RBL1. This evidence suggests

SOX4 regulates signaling events both at the external input level and the internal output or transcription level. This regulation could be direct, as with the IL-5Rα, or through transcriptional regulation of SOX4 activation and repression targets.

Interestingly, we demonstrated that SOX4 can activate two developmental signaling pathways that are important both in cancer and normal development. The NOTCH pathway is activated when cell surface ligands, such as DLL1, bind and stimulate cleavage of the NOTCH receptor on neighboring cells (24). SOX4 transcriptional targets include the NOTCH ligand DLL1, cell surface protease ADAM10, and NOTCH response gene HES2. We demonstrated that over expression of SOX4 caused increased NOTCH activation in prostate cancer cells, suggesting a mechanism by which upregulation of DLL1, ADAM10, or both by SOX4 activates the NOTCH pathway. In the prostate the NOTCH pathway is vital for normal epithelial development (192) and can promote metastases to the bone (104). SOX4 may represent a novel component of the NOTCH pathway in the prostate. The WNT-β-catenin signaling network is highly conserved and has pronounced roles in cancers of the colon and breast (145). SOX proteins, including SOX4, SOX17 (167) and SOX7 (68), can directly bind β-catenin and modulate transcriptional activity. In the case of SOX4, we demonstrated a cooperative effect with constitutively active β-catenin to activate the TOP-Flash reporter plasmid. SOX4 transcriptional targets include upstream Frizzled receptors and TCF/LEF sites were significantly enriched in SOX4 peaks. SOX4 could function by either increasing pathway activation through upregulation of membrane receptors or modulation of the downstream β-catenin transcriptional response. Interestingly, SOX4 may represent a novel link between the WNT-β-catenin and NOTCH pathways. Upregulation of the

WNT pathway in breast cancer cells causes transformation that is dependent on the NOTCH pathway and SOX4 target gene DLL1 (13). This data suggests that SOX4 is a key regulator of both the NOTCH and WNT network and may facilitate cross talk between pathways.

6.3 DNA Binding and Transcriptional Cofactors

Here, we have generated DNA binding specificity data for SOX4, which improved computation analyses for SOX4 specific binding sites. The PBM array allowed SOX4 to bind every possible 10-mer sequence, but specificity was only computed for an 8-mer binding site. Our data confirmed the known SOX family core-binding motif and add new specificity at the 1st, 7th and 8th positions. Crystal structure evidence from SOX2 binding to the FGF4 enhancer demonstrated the importance of the core binding motif, each of the seven nucleotides (AACAAAG) is directly contacted by an amino acid from two of the three alpha-helices present in the HMG DNA binding domain (152). SOX4 bound the same core sequence as SOX2, but our PBM data showed subtle preferences for alternate bases at each position except the two core adenine nucleotides in the 4th and 5th position. Interestingly, it is between these two nucleotides that SOX2 was shown to induce the extreme, 45° DNA bending HMG domain proteins are known for (152). Since these two nucleotides may be vital to SOX4's DNA binding activity it is unsurprising that alternate bases are not tolerated at these critical residues. Recently, the same PBM arrays have been applied to the entire mouse superfamily of homeodomain transcription factors and binding site specificities calculated (19). Similar to the shared SOX core binding motif, homeodomain subfamilies such as the HOX family, share a near identical core recognition motif with most of the binding site variation located around the core. The

specificity for SOX4 is most likely enhanced outside of the central core motif at positions such as the 8th nucleotide. Alternatively, *in vivo* DNA characteristics such as the level of DNA compaction, the local chromatin landscape and interactions with other transcription factors may also influence the sequences that SOX4 can recognize.

Transcription factors often function as one part of multiprotein regulatory complexes with both DNA binding and non-DNA binding cofactors. The SOX family of transcription factors is no different, and 11 members of this 20 protein family have been reported to directly interact with other transcription factors (196). SOX2 has been demonstrated to require an interaction with PAX6 to transcriptionally activate genes involved in lens development (86). We used CONFAC software to identify other transcription factor binding sites located within the SOX4 binding peaks (**TABLE 5**). Interestingly, like SOX2 we identified binding sites for another paired box factor PAX5, two E2F cell cycle regulatory factors and TCF proteins involved in responding to WNT signaling.

The enrichment of SMAD4 sites is particularly interesting in light of the GSEA results that SOX4 regulates many TGFβ target genes, including Tenascin C. We hypothesize that SOX4 may physically interact with SMAD4 in response to TGFβ signals. While it is possible to perform *in vitro* co-binding studies using the PBM array to determine how binding site preferences are influenced by protein-protein interactions, it is not possible to test this hypothesis in LNCaP cells because they do not express the TGFβ1 receptor. Currently, evidence points to a role for SOX4 in modulating other transcriptional programs via hierarchical regulation of 23 downstream transcription factors and protein-protein interactions that influence transcriptional output. Future

studies focused on identifying functional binding partners and mapping their interactions will integrate the SOX4 transcriptional network with other signaling pathways, providing a global picture of where SOX4 lies in the maintenance of cellular homeostasis.

6.4 SOX4 and Cancer

Cancer is a complex disease, highlighted by an ever-increasing number of cellular alterations that allow unrestrained cell growth. In a landmark paper, Hanahan and Weinberg generalized the molecular alterations needed for cancer formation into six separate categories: (1) self sufficiency in growth signals, (2) evasion of apoptosis, (3) limitless replicative potential, (4) insensitivity to anti-growth signals, (5) sustained angiogenesis and (6) invasion and metastatic ability (69). Based on the target genes we identified, SOX4 can contribute to cancer progression in several ways, influencing four of the six carcinogenic hallmark categories. SOX4 target genes include regulators of pivotal prostate cancer signaling networks of differentiation, cell survival, and apoptosis.

The phosphatase PTEN and transcription factor NKX3.1 are prostate cancer tumor suppressors that negatively regulate the PI3K-AKT pathway (103). Mice heterozygous for NKX3.1 and PTEN, only in the prostate, develop prostate adenocarcinomas and metastases to the lymph node with high frequency (2), implicating the importance of the PI3K-AKT pathway in prostate tumors. Our data suggests that SOX4 can promote self-sufficiency in growth signals through regulation of growth factor receptors. SOX4 can increase expression of EGFR, FGFRL1, and IGF2R, potentially activating the PI3K-AKT pathway and enhancing proliferative signals in tumors. SOX4 represses anti-growth signals through repression of NKX3.1. NKX3.1 suppresses the cell cycle through negative regulation of androgen signaling and stabilization of p53 (103).

Through regulation of the pro-survival/pro-growth signaling axis, SOX4 functions to suppress inhibitory signals thereby promoting the cell cycle and tumor growth.

The role of SOX4 in the apoptosis pathway is less clear as conflicting evidence has shown that both eliminating (112, 148) or over expressing SOX4 at high levels can promote apoptosis (83). Recent evidence suggests this paradox may be due to SOX4's regulation of p53. In response to DNA damage SOX4 induces p53 stabilization, and is critical for transcription of p53 target genes that mediate cell cycle arrest (140). Furthermore, SOX4 can activate expression of PUMA, a gene critical for the p53 apoptotic response (84) and repress NKX3.1 which can promote p53 dependent cell cycle arrest (103). Both genes provide a mechanism through which either loss or increased expression of SOX4 can promote apoptosis. Therefore, heightened SOX4 tumor levels may require p53 inactivation for malignant transformation. Interestingly, SOX4 induced colony formation in the p53 compromised RWPE-1 cell line (112) but not in a WT p53 cell line (103). There is clearly a cellular balance that must be found and while elevated SOX4 levels may require alternate cellular alterations to negate apoptotic functions, it at least plays an indirect role through the pro-survival pathways it regulates.

Finally, SOX4 may promote metastasis and tissue invasion via two independent methods. Dedifferentiation is an integral part of the metastatic process as cells undergo the Epithelial-Mesenchymal Transition (EMT) to begin invasion. This process is highlighted by cells dedifferentiating into a mesenchymal-like cell type with motile properties, followed by the reverse process when a suitable environment is found elsewhere in the body. SOX4 inhibits terminal differentiation via repression of transcriptions factors, such as NKX3.1, and activation of MLL and MLL3, two histone

H3 K4 methyltransferases that induce activation of HOX gene expression (127). MLL methyltransferase complexes can also facilitate E2F activation of S-phase promoters, driving the cell cycle forward. MLL is a critical oncogene that is often translocated or amplified in myeloid leukemogenesis, thus activation of MLL suggests a mechanism for the role of SOX4 in this disease (36). SOX4 may indirectly function to hold cells in an undifferentiated state, repressing progression to terminal differentiation through activation of MLL and MLL3.

Recently, both SOX4 and an activation target of SOX4, Tenascin C were shown to enhance metastasis of breast cancer cells to the lung (184), as has the TGFβ pathway, which activates their expression (138). Other metastasis-associated SOX4 target genes include Integrin α_v and RAC1. RAC1 was recently shown to control nuclear localization of β-catenin in response to WNT signals (200). SOX4 can directly stimulate the invasive and migratory properties of cells, but through its ability to enhance survival and proliferation via the PI3K-AKT pathway, SOX4 combines to enhance the overall metastatic ability of prostate cancer cells.

SOX4's role in cancer progression of multiple tissues can also be attributed to its key role in the response to developmental signaling pathways, such as WNT, NOTCH, HEDGEHOG, and TGFβ networks. Preliminary evidence points to an ability to activate signaling from the NOTCH receptor through transcriptional activation of ADAM10, a protease required for NOTCH cleavage and activation, DLL1, the NOTCH receptor ligand, as well as activation of the NOTCH target gene HES2 (24). SOX4 can modulate the expression of TGFβ pathway response genes (**TABLE 7**). However, currently we do not know how SOX4 affects these networks. Interestingly, the TGFB1 receptor is

frequently lost in prostate stromal cells, causing an upregulation of WNT3A ligand (107). In the WNT pathway, SOX4 activates expression of 3 Frizzled receptors and TCF sites are enriched in SOX4 binding peaks. SOX4 therefore may represent a novel link between these networks, modulating signals to provide cross talk and feedback among developmental signaling pathways. Through regulation of cell growth, survival, apoptotic and developmental signaling networks, SOX4 can maintain and promote cellular states conducive to cancer formation and progression.

6.5 SOX4 Regulates Components of RISC and the Small RNA Pathway

MiRNAs are a small noncoding RNA species that regulate the translation and stability of mRNA messages for hundreds of down stream target genes via partial complementarity to short sequences in the 3' untranslated regions of mRNAs. The RNA-Induced Silencing Complex (RISC), which is composed of AGO1 or AGO2, TRBP and DICER processes miRNAs from precursors (pre-miRNA) to their mature form, cleaves target mRNAs, and participates in translational inhibition (150). RNA Helicase A (RHA/DHX9) interacts with RISC and participates in loading of small RNAs into the complex (157). We observed that three components of RISC: DICER, AGO1 and RHA/DHX9, are high-confidence direct targets of SOX4 (**APPENDIX C**). We confirmed these data by qPCR (**FIGURE 10C**). DICER has been independently observed to be over expressed in prostate cancers (8). In addition, we observed that Toll-like receptor 3 (TLR3), which binds double-stranded RNAs, induces gene silencing, and can induce apoptosis (159), was induced 2.8-fold upon over expression of SOX4. As TLR3 was not detected by ChIP-chip, this induction may be indirect. However, we can

not exclude the possibility that SOX4 may directly regulate TLR3 from a distal or intronic enhancer.

Our observation that SOX4 targets three genes important in small RNA processing is of particular interest in light of the role SOX4 has in development and cancer progression. MiRNAs have been implicated in numerous physiological processes from development to oncogenesis. They can act both as suppressors of breast cancer metastasis via targeting of Tenascin C and SOX4 (184) and as a promoter of breast cancer metastasis (117). The finding that SOX4 can affect expression of multiple components of the RISC complex may provide insight into why long term loss of SOX4 induces widespread apoptosis (112, 148). Unfortunately, the NimbleGen array did not tile the 695 promoters of human miRNA transcripts present in the Sanger Institute's miRBase database (65). In the only literature report of SOX4 as a target of miRNA regulation, mir-335 was demonstrated to target the SOX4 3' UTR and elicit a roughly 2-fold repression of SOX4 mRNA (184). The SOX4 3'UTR is almost twice as long as the SOX4 coding sequence and is highly conserved throughout mammals. Although three independent miRNA target prediction algorithms predict potential miRNAs that regulate SOX4 (65, 95, 105), to date only mir-335 has been experimentally validated. Nevertheless, experiments are underway to identify miRNAs that regulate SOX4.

6.6 Conclusions and Future Directions

Transcription factors are critical cellular regulators of gene expression that are the driving force behind deciding which genes are expressed throughout the lifetime of the cell. There are transcription-independent functions for transcription factors, however initial experiments must be focused on nuclear transcriptional roles. Identifying genomic

binding sites and therefore regulatory target genes lends new insights into the gene networks a transcription factor regulates, allowing inferences about the outcomes of perturbations of any number of given factors. SOX4 is a highly conserved transcription factor crucial for mammalian development and important in cancer progression. We have identified, for the first time, 3,600 genomic binding sites mapping to 3,470 genes possibly under the transcriptional control of SOX4 in LNCaP prostate cancer cells. We have used PBMs to determine the DNA binding preferences for SOX4 and confirmed that these sequences are bound by SOX4 *in vivo*. These data lend valuable insights into the cellular functions for SOX4.

SOX4 plays an important role in the regulation of transmembrane receptor signaling and modulation of cellular transcriptional programs that influence key developmental pathways such as the WNT-β-catenin, NOTCH, TGFβ, HEDGEHOG, and the PI3K-AKT signaling cascade. Due to the high number of receptors and transcription factors SOX4 regulates, one major role of SOX4 could be to adapt and adjust cellular responses to external stimuli that target these developmental pathways, maintaining cells in a predefined state of differentiation. *In vivo* mouse genetic studies that have demonstrated the loss of discrete cell populations following either loss or over expression of SOX4 support this conclusion. In the hematopoietic lineage, SOX4 expression is required for the survival of maturing B-cells (163), whereas down regulation of SOX4 is required for myelination and terminal differentiation of oligodendrocytes (146). It has not been determined at which stage of differentiation SOX4 functions, but it is known that SOX4 is downstream of the master pluripotency factor SOX2, and upstream of terminal differentiation. Evidence from

immunofluorescent imaging in the intestinal crypt demonstrated that SOX4 expression is restricted to the transit-amplifying cells that reside downstream of the stem cell population and is lost in the terminally differentiated cells that migrate away from the crypt and onto the villi (167). This evidence suggests that SOX4 is involved in maintaining a transit-amplifying cell phenotype that allows differentiation programs to be initiated, but ultimately SOX4 must be down regulated for terminal differentiation to occur. To date, genetic alterations have demonstrated developmental roles in each tissue tested, suggesting SOX4 plays a physiological role in multiple cell types required for specific stages of tissue development. The diverse types of cancers SOX4 over expression has been identified in supports this conclusion. Future experiments in our lab will determine role of SOX4 in the development of the mouse prostate through prostate-specific homozygous deletion of SOX4. This will also allow *in vivo* validation of SOX4 target genes by either single gene qPCR or microarray analysis of $SOX4^{-/-}$ versus $SOX4^{+/+}$ prostate tissues.

We have identified downstream transcriptional targets of SOX4 but have not addressed the question of how SOX itself is regulated transcriptionally, translationally or through post-translational modifications. Studies suggest that the SOX proteins, in particular subgroup members, can regulate each other transcriptionally. Excellent starting candidates to address this question would be other SOX group C members, such as SOX11 and SOX12, both of which show a high degree of overlapping function and expression patterns in the developing mouse (46). In addition, our results showed SOX12 mRNA was significantly altered on both microarray platforms. Another hypothesis is that SOX4 is upregulated by numerous external ligands and stimuli.

Analysis of key signaling cascades and downstream transcription factors may identify common factors and regulatory elements in the SOX4 promoter. Ongoing experiments are already focused on identifying miRNAs that target the SOX4 3'UTR and those that are transcriptionally activated by SOX4. In parallel, our lab proposes to perform high-throughput mass spectrometry experiments that will identify both protein binding partners for and post-translational modifications for SOX4.

In vivo studies are critical to verify SOX4 target genes and to study the pathways SOX4 affects in a living model system. To date, each study that has over expressed or deleted SOX4, either in specific tissues or the whole animal, have reported severe, tissue altering phenotypes. Experiments are currently underway to engineer a prostate specific deletion of SOX4 as well as a prostate specific, inducible SOX4 system. Currently, we have one male mouse that contains LoxP sites flanking both copies of SOX4 as well as the CRE recombinase gene under the control of the Probasin promoter. Probasin is specifically expressed in the male prostate and seminal vesicles and is detectable at the onset of puberty (201). SOX4 deletion will not occur until puberty, around two weeks of age, however, in humans, the prostate undergoes rapid, androgen dependent expansion at the onset of puberty (74). We hypothesize that loss of SOX4 during this, and later developmental processes will severely impact prostate development. Prostate specific over expression of SOX4 will address whether SOX4 alone can drive prostate tumorigenesis. Importantly, these new model mice will allow in vivo studies of SOX4 including target gene validation and preclinical testing of drugs that target SOX4 regulatory networks.

Integration of the data presented here and future work will facilitate the creation of a global cellular SOX4 interaction network, and provide insights into the signaling pathways that SOX4 regulates. As more data emerges, it is clearer than ever that SOX4 has an important role in normal mammalian development. Understanding the entire regulatory network, and how uncontrolled SOX4 expression can contribute to cancer progression is critical for disease prevention and treatment. These data have implications not only for prostate cancer but many other cancers as well.

Appendix A: SOX4 Binding Sites and Regulatory Genes

Tables denoting the location of the 3,600 SOX4 bound peaks and the 3,470 target genes

can be found on the Moreno Lab website:

http://morenolab.whitehead.emory.edu/pubs/SOX4-network/Scharer et al - Supplemental

Table 1.txt

Appendix B: Illumina Genes with Significant Expression Changes

Tables indicating the 1,766 genes altered by SOX4 over expression detected by the

Illumina platform can be found on the Moreno Lab website:

http://morenolab.whitehead.emory.edu/pubs/SOX4-network/Scharer et al - Supplemental

Table 2.txt

Appendix C: 282 High-Confidence SOX4 Target Genes

Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change	Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change
BICD2	23299		-5.03	IFITM1	8519	-3.03	
NKX3-1	4824		-4.53	IFIT3	3437	-2.93	
PDK1	5163		-4.02	SNFT	55509	-2.32	
FLOT2	2319		-3.78	OAZ3	51686	-2.29	
DAZAP2	9802		-3.20	SURB7	9412	-2.24	
GTF3C2	2976		-3.12	CAPZA1	829	-2.19	
MKNK2	2872		-3.08	DNAL1	83544	-2.13	
AMACR	23600		-2.97	HIST1H2BG	8339	-2.10	
GART	2618		-2.94	S100A11	6282	-2.09	
KIF5C	3800		-2.90	RPL21	6144	-1.99	
VAMP3	9341		-2.87	CAST	831	-1.96	
NFIX	4784		-2.83	RPL7A	6130	-1.95	
ELF5	2001		-2.65	HIST1H2AI	8329	-1.94	
CBR1	873		-2.57	SUHW4	54816	-1.93	
ARPC1A	10552		-2.48	UXT	8409	-1.89	
FOXA1	3169		-2.47	C10orf104	119504	-1.89	
DECR1	1666		-2.47	FRG1	2483	-1.88	
GRB7	2886		-2.41	CETN3	1070	-1.86	
GOLSYN	55638		-2.34	RPL14	9045	-1.84	
CDK2	1017	1.69	-2.25	TIMM8B	26521	-1.83	
SURF6	6838		-2.19	LACTB2	51110	-1.82	
SRF	6722		-2.03	MGMT	4255	-1.80	
RLBP1L1	157807		-1.97	C9orf9	11092	-1.79	
TPD52	7163		-1.94	HIST1H4H	8365	-1.79	
LOC130355	130355	-2.36	-1.93	C3orf28	26355	-1.78	
RAB14	51552		-1.75	GBA3	57733	-1.78	
HSPA1A	3303		-1.69	LOC391356	391356	-1.76	
FYCO1	79443		1.55	RRP15	51018	-1.76	
ZNF281	23528		1.71	RPL7	6129	-1.73	
PPT2	9374		1.79	GAGE2	2574	-1.72	
DICER1	23405		1.83	HMGN3	9324	-1.71	
BTN3A3	10384	1.85	1.97	FAIM	55179	-1.71	
RB1CC1	9821		1.97	C6orf203	51250	-1.71	
EIF2C1	26523		1.98	RP11-529I10.4	25911	-1.70	
MORF4L1	10933		2.07	KRT10	3858	-1.68	
ISG20	3669	-1.72	2.17	PPIAL4	164022	-1.68	
TMEM184B	25829		2.28	NSUN5B	155400	-1.67	
RYK	6259	1.54	2.42	PMS2	5395	-1.67	
PRDM2	7799		2.48	RPL37A	6168	-1.66	
DUSP5	1847		2.48	TPRKB	51002	-1.66	
ENPP4	22875		2.53	C4orf27	54969	-1.65	
NCOA4	8031		2.64	SHFM1	7979	-1.63	
ENPP5	59084		2.68	ZNRD1	30834	-1.63	
ADAM10	102	1.55	2.81	C10orf35	219738	-1.62	

Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change	Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change
CHCHD8	51287	-1.61		EIF2AK3	9451	1.52	
CLDN14	23562	-1.60		RHBDD1	84236	1.52	
FAM24B	196792	-1.60		KIFC2	90990	1.52	
PMS2L1	5379	-1.60		C11orf82	220042	1.52	
RPL18	6141	-1.60		XPC	7508	1.52	
CD163L1	283316	-1.59		MKNK1	8569	1.53	
PDDC1	347862	-1.59		SLC20A1	6574	1.53	
IFI44	10561	-1.58		MUM1	84939	1.53	
TCEAL1	9338	-1.57		CMTM6	54918	1.53	
ALKBH6	84964	-1.57		WDR5	11091	1.53	
HIST1H1C	3006	-1.56		ZNF165	7718	1.53	
SUMO4	387082	-1.56		ADCY6	112	1.53	
CHMP2A	27243	-1.56		CCDC126	90693	1.53	
CINP	51550	-1.56		G6PC3	92579	1.54	
SLC6A9	6536	-1.55		DDX23	9416	1.55	
NHP2L1	4809	-1.55		MCM4	4173	1.55	
PSMD13	5719	-1.55		HIATL1	84641	1.56	
FLJ10490	55150	-1.55		SLC12A8	84561	1.57	
RAC1	5879	-1.54		DERL1	79139	1.57	
BXDC5	80135	-1.54		NEDD4	4734	1.57	
ERMAP	114625	-1.54		ST6GAL1	6480	1.57	
REM2	161253	-1.54		SLC31A1	1317	1.58	
LSM1	27257	-1.53		MAGED1	9500	1.58	
SF3B1	23451	-1.53		DHX9	1660	1.59	
EAPP	55837	-1.53		DOLPP1	57171	1.59	
ERN1	2081	-1.53		GIT1	28964	1.61	
CFL2	1073	-1.52		PDIA6	10130	1.61	
LEPRE1	64175	-1.52		FGFRL1	53834	1.61	
ISCA1	81689	-1.52		SLC3A2	6520	1.62	
MRPS18C	51023	-1.51		BCR	613	1.62	
S100A13	6284	-1.51		RNF145	153830	1.62	
TMEM126B	55863	-1.50		GBA	2629	1.62	
FAM3D	131177	1.50		ARAF	369	1.62	
DIDO1	11083	1.50		DLL1	28514	1.63	
IGF2R	3482	1.51		INADL	10207	1.63	
GCNT1	2650	1.51		TKT	7086	1.64	
ELAC2	60528	1.51		LARS2	23395	1.64	
LRRC8D	55144	1.51		SLC12A9	56996	1.64	
FLVCR2	55640	1.51		TTC14	151613	1.52	
PPM1A	5494	1.52		PHLDA1	22822	1.52	
GRK6	2870	1.52		APBB3	10307	1.64	
TTC14	151613	1.52		PCSK7	9159	1.64	
PHLDA1	22822	1.52		MVK	4598	1.64	

Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change	Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change
HMGCR	3156	1.65		TROAP	10024	1.83	
ALOXE3	59344	1.65		RHOT2	89941	1.83	
TRAM2	9697	1.65		PAG1	55824	1.84	
SLC39A14	23516	1.65		AGRN	375790	1.85	
C4orf18	51313	1.65		CLCN7	1186	1.85	
SUMF2	25870	1.65		CSAD	51380	1.87	
PIGT	51604	1.65		SLC27A2	11001	1.87	
ITGAV	3685	1.65		TNFRSF10D	8793	1.88	
MASK-BP3	404734	1.66		C12orf35	55196	1.90	
ZNF184	7738	1.66		NIN	51199	1.92	
ARMC5	79798	1.66		C7orf53	286006	1.92	
C14orf173	64423	1.67		PLCH1	23007	1.94	
PTGFRN	5738	1.68		C6orf85	63027	1.96	
GALC	2581	1.68		FHOD3	80206	1.96	
ANKRD26	22852	1.68		C3orf64	285203	1.96	
PNKD	25953	1.69		FSTL1	11167	1.98	
AGPAT6	137964	1.70		FAM135A	57579	1.98	
NLRX1	79671	1.70		ARID2	196528	1.99	
DSC2	1824	1.71		NOMO3	408050	2.00	
POFUT1	23509	1.71		ACAD10	80724	2.01	
MANSC1	54682	1.72		PTCH1	5727	2.03	
SORT1	6272	1.72		CBLN2	147381	2.03	
PILRB	29990	1.73		KIAA0319L	79932	2.03	
LMBRD2	92255	1.73		ARSA	410	2.04	
TMEM43	79188	1.74		PDE4D	5144	2.05	
TMEM16F	196527	1.74		MLL3	58508	2.06	
KIAA1344	57544	1.75		KIAA0195	9772	2.08	
TMEM8	58986	1.75		ELOVL2	54898	2.09	
FZD5	7855	1.75		ACVR2A	92	2.10	
SLC39A8	64116	1.76		SLC1A4	6509	2.11	
DHCR24	1718	1.76		PSD3	23362	2.12	
VIPR1	7433	1.77		IGSF9	57549	2.16	
TES	26136	1.77		C9orf100	84904	2.19	
ADCK2	90956	1.78		ATG16L2	89849	2.19	
ARFGAP1	55738	1.78		PARP10	84875	2.20	
DAGLB	221955	1.79		RGMA	56963	2.20	
SERINC2	347735	1.79		ENTPD4	9583	2.21	
BFAR	51283	1.79		PIGO	84720	2.21	
RBL1	5933	1.80		ACCN3	9311	2.23	
TMEM29	29057	1.80		IL17RB	55540	2.26	
METTL4	64863	1.80		C18orf54	162681	2.31	
TM7SF2	7108	1.81		TNC	3371	2.31	
ATAD1	84896	1.81		POLH	5429	2.34	
SON	6651	1.81		RIF1	55183	2.36	
PTPRA	5786	1.82		ABCA2	20	2.38	

Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change	Symbol	Entrez ID	Illumina Fold Change	Affymetrix Fold Change
TMEM168	64418	1.82		MEX3B	84206	2.56	
PIK4CA	5297	2.39		MLL	4297	2.56	
DSG2	1829	2.39		ECT2	1894	2.57	
SEC61A1	29927	2.40		SMA4	11039	2.59	
LCOR	84458	2.41		C7orf20	51608	2.60	
GZF1	64412	2.42		PCDHB3	56132	2.82	
CEP250	11190	2.43		COL23A1	91522	2.89	
FLJ14082	80092	2.47		HPX	3263	3.40	
C9orf102	56959	2.50		CNTNAP2	26047	3.48	
ZDHHC21	340481	2.51		SLC37A3	84255	4.33	

Appendix D: Enrichment Scores of K-mers Bound by SOX4

Tables denoting the sequences bound by SOX4 and their respective enrichment scores can be found on the Moreno Lab website:

http://morenolab.whitehead.emory.edu/pubs/SOX4-network/Scharer et al - Supplemental Table 4.txt

Appendix E: Primers and Oligos

All primers and oligos described in this dissertation can be found on the Moreno Lab website:

http://morenolab.whitehead.emory.edu/pubs/SOX4-network/Scharer et al - Supplemental Table 7.txt

Appendix F: Materials and Methods

I. AURORA KINASE INHIBITORS SYNERGIZE WITH PACLITAXEL TO INDUCE APOPTOSIS IN OVARIAN CANCER CELLS

1.1 *Tumor samples, RNA isolation, Microarray Hybridization and Normalization*

A detailed explanation of patient samples and microarray hybridization and normalization techniques is described elsewhere (128). The complete dataset is available at the NCBI GEO website (http://www.ncbi.nlm.nih.gov/geo/index.cgi, accession number GSE7463) and at the author's website http://morenolab.whitehead.emory.edu.

1.2 *Cell Culture and Drug Treatment*

PTX10 and 1A9 cells were cultured in RPMI media (Mediatech, Herndon, VA) supplemented with 10% fetal bovine serum and grown in 5% CO2 at 37°C. Two days before treatment 1.5×10^5 cells were seeded in each well of a 6-well plate (Corning, Corning, NY). On day one of treatment combinations of 15 ng/mL paclitaxel (Sigma-Aldrich, St. Louis, MO) and either Dimethyl Sulfoxide (DMSO) control or the indicated concentration of VE-465 (Vertex Pharmaceuticals, Abingdon, United Kingdom) were added to 2 mL of fresh RPMI and incubated for 96 hours prior to FACS analysis or Caspase 3/7 activity assays.

1.3 *Fluorescence Activated Cell Sorting (FACS) Analysis*

Following drug treatment, cells were washed from the plate in media, centrifuged at 3000 RPM to pellet and washed once with cold PBS. Pellets were resuspended and fixed in 70% Ethanol/PBS at -20°C overnight. On the day of analysis, pellets were washed once with PBS and digested with 500 μl of 0.1 mg/mL PBS/RNaseA (Sigma-Aldrich, St.

Louis, MO) by incubating at 37°C for 15 minutes. DNA content was assessed by staining with 500 μl of 25 μg/mL PBS/Propidium Iodide (Sigma-Aldrich, St. Louis, MO). Cell suspensions were transferred to 5 mL collection tubes for FACS analysis. Samples were processed using a Becton Dickson FACSCalibur analyzer (Becton Dickson, San Jose, CA) and data analyzed using the FlowJo software package (Tree Star, Ashland, OR).

1.4 *Drug Treatment and Caspase Assay*

One day before drug treatment, each well of a white-walled, 96 well luminometer plate (Nalge Nunc International, Rochester, NY) was coated with a 1:4 dilution of BD matrigel matrix (BD biosciences, Bedford, MA) and RPMI media. The plates were incubated at room temperature for one hour and excess matrigel was removed before 4800 cells were seeded in each well in triplicate. On day one of treatment, cells were treated with or without 15 ng/mL paclitaxel (Sigma-Aldrich, St. Louis, MO) plus varying concentrations and combinations of VE-465 (Vertex Pharmaceuticals, Abingdon, United Kingdom), or with 50 μM Z-VAD (EMD Chemicals, San Diego, CA). Z-VAD is a general caspase inhibitor and was used as a negative control to block caspase activity and apoptosis. Control cells were left untreated. Three independent biological replicates were performed, luminescence measured, and data analyzed.

The Caspase-Glo™ 3/7 Assay (Promega, Madison, WI) lyophilized substrate (DEVD-aminoluciferin powder) was resuspended in Caspase- Glo™ 3/7 lysis buffer and equilibrated to room temperature. Forty-eight or 72 hours after cell treatment, the Caspase- Glo™ 3/7 reagent was added in a 1:1 volume ratio to each well of the 96 well luminometer plate. Immediately following the addition of the reagent, the contents of the

wells were gently mixed with a plate shaker at 500 RPM for 30 seconds. After one hour

incubation, the luminescence was measured with a Synergy HT plate reader (BioTek

Instruments, Winooski, VT). Culture medium was used as a blank and "no-cell

background" values were determined.

1.5 *Immunofluorescence*

PTX10 and 1A9 cells were grown on cover slips (Fisher Scientific, Hampton, NH) in 6-

well culture dishes (Corning, Corning, NY). Cells were washed 3 times with cold PBS

and fixed in 4% paraformaldehyde for 15 minutes at room temperature, permeablized on

ice for 2 minutes in 0.5% Tween-20/PBS and blocked in 5% nonfat dry milk (NFDM) for

30 minutes at room temperature. Mitotic cells were stained with anti-phospho-Histone H3

Serine 10 (Upstate, Charlottesville, VA) with 5% NFDM at a 1:200 dilution for 2 hours

at 4°C. Secondary antibody of anti-Rabbit AlexaFluor 488 (Molecular Probes, Eugene,

OR) was applied at a 1:400 dilution for 45 minutes at room temperature. Cells were

washed 3 times in PBS and stained with TOPro (Molecular Probes, Eugene, OR) at a

concentration of 3 μg/μl for 15 minutes to reveal the nucleus. Cover slips were mounted

on slides and visualized using a Zeiss Axiovert 35 fluorescence microscope.

1.6 *Western Blot*

60% confluent cells were lysed in lysis buffer (0.137 M NaCl, 0.02 M TRIS pH 8.0, 10%

Glycerol, and 1% NP-40), 50 μg total lysate separated by SDS-PAGE electrophoresis and

transferred to nitrocellulose for immunoblotting. Immunoblots were probed with an

antibody to Aurora-A (Abcam Inc., Cambridge, MA), Aurora-B (GenScript, Piscitaway,

NJ), phospho-AurB (Cell Signaling, Danvers, MA), p53 (Santa Cruz Biotechnology,

Santa Cruz, CA) and phospho(S315)p53 (Cell Signaling, Danvers, MA). To ensure equal

loading blots were then probed with a monoclonal antibody to PP2A, catalytic subunit (BD Biosciences, San Jose, CA).

1.7 *Tissue Microarray Analysis*

TMA sections were stained at the WCI Tissue and Pathology Core Facility http://www.pathology.emory.edu/WCIPathCore/ with H&E and with AurA antibody (1:300 dilution, Abcam, Cambridge, MA). Staining was scored on a four level scale (0 = no staining, 1 = weak staining, 2 = moderate staining, 3 = intense staining) by a GU pathologist

II. GENOME-WIDE PROMOTER ANALYSIS OF THE SOX4 TRANSCRIPTIONAL NETWORK IN PROSTATE CANCER CELLS

2.1 *Cell Culture and Stable Cell Line Construction*

All cell lines were cultured as described by ATCC except LNCaP cells, which were cultured with T-Medium (Invitrogen, Carlsbad, CA). HA tagged SOX4 was cloned into the pHR-UBQ-IRES-eYFP-ΔU3 lentiviral vector (gift from Dr. Hihn Ly, Emory University) and stable cells isolated as previously described (123).

2.2 *ChIP Assay*

ChIP assays were performed as previously described (123). Briefly, two 90% confluent P150s of both RWPE-1-FLAG-SOX4 or RWPE-1, or LNCaP-YFP and LNCaP-YFP/HA-SOX4 or RWPE-1-YFP and RWPE-1-YFP/HA-SOX4 cells were formaldehyde fixed and lysed. Chromatin suspensions were sonicated on a Branson Sonifier 250 with the settings: Duty – 50%, Output Control - 3. Sonication was performed on ice for four rounds of 8 four-second bursts. Anti-HA 12CA5, Anti-Flag-M2 (Sigma-Aldrich, St.

Louis, MI) or mouse IgG was used to immunoprecipitate protein-DNA complexes overnight at 4°C and collected using Dynal M280 sheep anti-mouse IgG beads (Invitrogen, Carlsbad, CA) for 2 hours. Dynal beads were washed, protein-DNA complexes eluted and DNA purified. Purified DNA was blunted and whole genome amplification performed by ligation-mediated PCR as described previously (153). To confirm the ChIPOTLe predicted ChIP-chip peaks, PCR primers were designed to span the genomic coordinates of the predicted peak and each target site was confirmed in three independent ChIP assays. All PCR primers used in ChIP-PCR can be found in **APPENDIX E**.

2.3 *ChIP-chip Analysis*

To determine the direct *SOX4* target genes on a global scale we performed ChIP assays in triplicate from the LNCaP cell line stably expressing *SOX4* and in duplicate from a control cell line that expressed YFP alone. Immunoprecipitated and input DNA were subjected to whole genome amplification, Cy3/Cy5 fluorescent labeling, and hybridization to the NimbleGen 25K human promoter array set. Input and immunoprecipitated DNA isolated from LNCaP-YFP and LNCaP-YFP/HA-SOX4 cells were amplified using linker-mediated PCR as described previously (153). Amplified DNA was labeled and hybridized in triplicate by NimbleGen Systems, Inc to their human 25K promoter array. Raw hybridization data was Z-score normalized and ratios of IP to Input DNA were determined for each sample. ChIPOTle software was used to determine enriched peaks using a 500 bp sliding window every 50 bp as previously described (123). NimbleGen microarray data are available from the GEO database, accession number GEO11915.

2.4 *Plasmid Luciferase Assays*

PCR fragments representing the binding sites in the *EGFR*, *ERBB2* and *TLE1* promoter

sequences were cloned in front of the pGL3-promoter luciferase construct (Promega,

Fitchburg, WI). Primers sequences used can be found in **APPENDIX E**. LNCaP cells

were transfected with 100 ng of a TK-Renilla construct, 500 ng of pGL3-promoter vector

alone and with cloned inserts, as well as 500 ng of either a *SOX4* or vector expression

construct. Dual Luciferase assays were performed 48 hours post transfection according to

the manufacturer's guidelines (Promega, Fitchburg, WI). All assays were performed in

triplicate on separate days.

2.5 *Quantitative Real Time PCR*

LNCaP cells were plated in 6-well culture dishes and grown to 90% confluency before

transfection with 1 μg of *SOX4* plasmid or vector control using Lipofectamine-2000

(Invitrogen, Carlsbad, CA). 24 hours post-transfection total RNA was harvested using the

RNeasy kit (Qiagen, Valencia, CA) and reverse transcription performed using Superscript

II reverse transcriptase (Invitrogen, Carlsbad, CA). QPCR was performed using SYBR

Green I (Invitrogen, Carlsbad, CA) on a Bio-Rad iCycler using 18s or β-actin as a control

and data analyzed using the delta-Ct method (114). All primers used in this study are

listed in **APPENDIX E**.

2.6 *Microarray Analysis*

Total RNA was isolated from three independent experiments of either vector control or

SOX4 transfected LNCaP cells as described above. Each transfection was performed in

triplicate and each sample was hybridized in duplicate creating six data points for each

condition. Total RNA was submitted to the Winship Cancer Institute DNA Microarray

Core facility (http://microarray.cancer.emory.edu/). All samples demonstrated RNA

integrity (RIN) of 8.3 or greater using an Agilent 2100 Bioanalyzer. RNA was hybridized

to the Illumina Human6 v2 Expression Beadchip that query roughly 47,000 transcripts

with 48,701 probes, and after normalization significantly changed probes were calculated

using Significance Analysis of Microarrays (SAM) software (188). Settings for SAM

were: two-class unpaired (X4 vs. Vector control), Imputation engine – 10 Nearest

Neighbor, Permutations – 500, RNG seed – 1234567, Delta – 1.316, Fold Change – 1.5,

False discovery rate – 0.749%. Microarray data are available in the GEO database,

accession number GEO11915.

2.7 *Beta-catenin Luciferase Assay*

90% confluent LNCaP cells were transfected with a combination of the following

plasmids: TK Renilla – 10ng, TOP-Flash Firefly – 100ng, vector control – 100ng, SOX4

vector – 100, 200, or 300 ng, β-catenin A33Y mutant – 100 ng. 24 hrs post-transfection

cells were lysed and dual luciferase assay performed according to the manufacturers

instructions (Promega). All assays were performed in triplicate.

2.8 *Electromobility Shift Assays*

Recombinant GST-SOX4-DNA binding domain (SOX4-DBD) corresponding to amino

acids 1-340 was cloned into the pGEX4T-1 plasmid (gift from Dr. Anita Corbet, Emory

University) and purified with Glutathione Sepharose 4B beads (GE Healthcare Bio-

Sciences AB, Uppsala, Sweden). Double-stranded, 35-mer oligos containing either a

core *SOX4* binding site (AACAAAG) or a mutated site (CCTGGCA) were end-labeled

using T4 Polynucleotide Kinase. 1, 2, 4 and 8 μg of GST-SOX4-DBD were incubated in

PBS with 50,000 cpm of probe, a mutated binding site or unlabeled competitor oligos.
All oligo sequences can be found in **APPENDIX E**. Binding reactions were carried out
at 4°C for 30 minutes, complexes loaded onto a non-denaturing acryalmide gel and
exposed overnight at -80°C

2.9 *Immunoblotting*

Cells were lysed in lysis buffer (0.137M NaCl, 0.02M TRIS pH 8.0, 10% Glycerol, and
1% NP-40), 50 µg total lysate separated by SDS-PAGE electrophoresis and transferred to
nitrocellulose for immunoblotting. Immunoblots were probed with polyclonal rabbit
SOX4 antisera described previously (112) and DICER1 (Santa Cruz Biotech., Santa Cruz,
CA). To control for equal loading Immunoblots were probed with a monoclonal antibody
to Protein phosphatase 2A (PP2A) catalytic subunit (BD Biosciences, San Jose, CA).

Appendix G: Literature Cited

1. **Aaboe, M., K. Birkenkamp-Demtroder, C. Wiuf, F. B. Sorensen, T. Thykjaer, G. Sauter, K. M. Jensen, L. Dyrskjot, and T. Orntoft.** 2006. SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. Cancer Res **66:**3434-42.

2. **Abate-Shen, C., W. A. Banach-Petrosky, X. Sun, K. D. Economides, N. Desai, J. P. Gregg, A. D. Borowsky, R. D. Cardiff, and M. M. Shen.** 2003. Nkx3.1; Pten mutant mice develop invasive prostate adenocarcinoma and lymph node metastases. Cancer Res **63:**3886-90.

3. **ACS.** 2008. Cancer Facts & Figures 2008.

4. **Ahn, S. G., G. H. Cho, S. Y. Jeong, H. Rhim, J. Y. Choi, and I. K. Kim.** 1999. Identification of cDNAs for Sox-4, an HMG-Box protein, and a novel human homolog of yeast splicing factor SSF-1 differentially regulated during apoptosis induced by prostaglandin A2/delta12-PGJ2 in Hep3B cells. Biochem Biophys Res Commun **260:**216-21.

5. **Ahn, S. G., H. S. Kim, S. W. Jeong, B. E. Kim, H. Rhim, J. Y. Shim, J. W. Kim, J. H. Lee, and I. K. Kim.** 2002. Sox-4 is a positive regulator of Hep3B and HepG2 cells' apoptosis induced by prostaglandin (PG)A(2) and delta(12)-PGJ(2). Exp Mol Med **34:**243-9.

6. **Akahane, D., T. Tauchi, S. Okabe, K. Nunoda, and K. Ohyashiki.** 2008. Activity of a novel Aurora kinase inhibitor against the T315I mutant form of BCR-ABL: in vitro and in vivo studies. Cancer Sci **99:**1251-7.

7. **Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter.** Molecular Biology of the Cell, 4 ed. Garland.

8. **Ambs, S., R. L. Prueitt, M. Yi, R. S. Hudson, T. M. Howe, F. Petrocca, T. A. Wallace, C. G. Liu, S. Volinia, G. A. Calin, H. G. Yfantis, R. M. Stephens, and C. M. Croce.** 2008. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Res **68:**6162-70.

9. **Anand, S., S. Penrhyn-Lowe, and A. R. Venkitaraman.** 2003. AURORA-A amplification overrides the mitotic spindle assembly checkpoint, inducing resistance to Taxol. Cancer Cell **3:**51-62.

10. **Andersson, A., C. Ritz, D. Lindgren, P. Eden, C. Lassen, J. Heldrup, T. Olofsson, J. Rade, M. Fontes, A. Porwit-Macdonald, M. Behrendtz, M. Hoglund, B. Johansson, and T. Fioretos.** 2007. Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. Leukemia **21:**1198-203.

11. **Araki, K., K. Nozaki, T. Ueba, M. Tatsuka, and N. Hashimoto.** 2004. High expression of Aurora-B/Aurora and Ipll-like midbody-associated protein (AIM-1) in astrocytomas. J Neurooncol **67:**53-64.

12. **Avilion, A. A., S. K. Nicolis, L. H. Pevny, L. Perez, N. Vivian, and R. Lovell-Badge.** 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev **17:**126-40.

13. **Ayyanan, A., G. Civenni, L. Ciarloni, C. Morel, N. Mueller, K. Lefort, A. Mandinova, W. Raffoul, M. Fiche, G. P. Dotto, and C. Brisken.** 2006. Increased Wnt signaling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism. Proc Natl Acad Sci U S A **103:**3799-804.

14. **Bader, G. D., I. Donaldson, C. Wolting, B. F. Ouellette, T. Pawson, and C. W. Hogue.** 2001. BIND--The Biomolecular Interaction Network Database. Nucleic Acids Res **29:**242-5.

15. **Bartlett, J. A., M. J. Fath, R. Demasi, A. Hermes, J. Quinn, E. Mondou, and F. Rousseau.** 2006. An updated systematic overview of triple combination therapy in antiretroviral-naive HIV-infected adults. Aids **20:**2051-64.

16. **Bayliss, R., T. Sardon, I. Vernos, and E. Conti.** 2003. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. Mol Cell **12:**851-62.

17. **Beiter, K., E. Hiendlmeyer, T. Brabletz, F. Hlubek, A. Haynl, C. Knoll, T. Kirchner, and A. Jung.** 2005. beta-Catenin regulates the expression of tenascin-C in human colorectal tumors. Oncogene **24:**8200-4.

18. **Berdnik, D., and J. A. Knoblich.** 2002. Drosophila Aurora-A is required for centrosome maturation and actin-dependent asymmetric protein localization during mitosis. Curr Biol **12:**640-7.

19. **Berger, M. F., G. Badis, A. R. Gehrke, S. Talukder, A. A. Philippakis, L. Pena-Castillo, T. M. Alleyne, S. Mnaimneh, O. B. Botvinnik, E. T. Chan, F. Khalid, W. Zhang, D. Newburger, S. A. Jaeger, Q. D. Morris, M. L. Bulyk, and T. R. Hughes.** 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell **133:**1266-76.

20. **Berger, M. F., A. A. Philippakis, A. M. Qureshi, F. S. He, P. W. Estep, 3rd, and M. L. Bulyk.** 2006. Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. Nat Biotechnol **24:**1429-35.

21. **Bergsland, M., M. Werme, M. Malewicz, T. Perlmann, and J. Muhr.** 2006. The establishment of neuronal properties is controlled by Sox4 and Sox11. Genes Dev **20:**3475-86.

22. **Billiard, J., R. A. Moran, M. Z. Whitley, M. Chatterjee-Kishore, K. Gillis, E. L. Brown, B. S. Komm, and P. V. Bodine.** 2003. Transcriptional profiling of human osteoblast differentiation. J Cell Biochem **89:**389-400.

23. **Bischoff, J. R., L. Anderson, Y. Zhu, K. Mossie, L. Ng, B. Souza, B. Schryver, P. Flanagan, F. Clairvoyant, C. Ginther, C. S. Chan, M. Novotny, D. J. Slamon, and G. D. Plowman.** 1998. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. Embo J **17:**3052-65.

24. **Bolos, V., J. Grego-Bessa, and J. L. de la Pompa.** 2007. Notch signaling in development and cancer. Endocr Rev **28:**339-63.

25. **Bolton, M. A., W. Lan, S. E. Powers, M. L. McCleland, J. Kuang, and P. T. Stukenberg.** 2002. Aurora B kinase exists in a complex with survivin and INCENP and its kinase activity is stimulated by survivin binding and phosphorylation. Mol Biol Cell **13:**3064-77.

26. **Bourguignon, L. Y., H. Zhu, L. Shao, and Y. W. Chen.** 2000. CD44 interaction with tiam1 promotes Rac1 signaling and hyaluronic acid-mediated breast tumor cell migration. J Biol Chem **275:**1829-38.

27. **Boyer, L. A., T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, and R. A. Young.** 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell **122:**947-56.

28. **Buck, M. J., A. B. Nobel, and J. D. Lieb.** 2005. ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. Genome Biol **6:**R97.

29. **Cannistra, S. A.** 2004. Cancer of the ovary. N Engl J Med **351:**2519-29.

30. **Carmena, M., and W. C. Earnshaw.** 2003. The cellular geography of aurora kinases. Nat Rev Mol Cell Biol **4:**842-54.

31. **Carvajal, R. D., A. Tse, and G. K. Schwartz.** 2006. Aurora kinases: new targets for cancer therapy. Clin Cancer Res **12:**6869-75.

32. **Chaboissier, M. C., A. Kobayashi, V. I. Vidal, S. Lutzkendorf, H. J. van de Kant, M. Wegner, D. G. de Rooij, R. R. Behringer, and A. Schedl.** 2004. Functional analysis of Sox8 and Sox9 during sex determination in the mouse. Development **131:**1891-901.

33. **Chellaiah, M. A., R. S. Biswas, S. R. Rittling, D. T. Denhardt, and K. A. Hruska.** 2003. Rho-dependent Rho kinase activation increases CD44 surface expression and bone resorption in osteoclasts. J Biol Chem **278:**29086-97.

34. **Chen, J., S. Jin, S. K. Tahir, H. Zhang, X. Liu, A. V. Sarthy, T. P. McGonigal, Z. Liu, S. H. Rosenberg, and S. C. Ng.** 2003. Survivin enhances Aurora-B kinase activity and localizes Aurora-B in human cells. J Biol Chem **278:**486-90.

35. **Cheung, M., M. Abu-Elmagd, H. Clevers, and P. J. Scotting.** 2000. Roles of Sox4 in central nervous system development. Brain Res Mol Brain Res **79:**180-91.

36. **Chowdhury, T., and H. J. Brady.** 2008. Insights from clinical studies into the role of the MLL gene in infant and childhood leukemia. Blood Cells Mol Dis **40:**192-9.

37. **Cichy, J., and E. Pure.** 2003. The liberation of CD44. J Cell Biol **161:**839-43.

38. **Crosio, C., G. M. Fimia, R. Loury, M. Kimura, Y. Okano, H. Zhou, S. Sen, C. D. Allis, and P. Sassone-Corsi.** 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. Mol Cell Biol **22:**874-85.

39. **Damber, J.-E., and G. Aus.** Prostate cancer. The Lancet **371:**1710-1721.

40. **Dennis, G., Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki.** 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol **4:**P3.

41. **Denny, P., S. Swift, N. Brand, N. Dabhade, P. Barton, and A. Ashworth.** 1992. A conserved family of genes related to the testis determining gene, SRY. Nucleic Acids Res **20:**2887.

42. **Dhanasekaran, S. M., T. R. Barrette, D. Ghosh, R. Shah, S. Varambally, K. Kurachi, K. J. Pienta, M. A. Rubin, and A. M. Chinnaiyan.** 2001. Delineation of prognostic biomarkers in prostate cancer. Nature **412:**822-6.

43. **Du Bois, A., and J. Pfisterer.** 2005. Future options for first-line therapy of advanced ovarian cancer. Int J Gynecol Cancer **15 Suppl 1:**42-50.

44. **Du, J., and G. J. Hannon.** 2002. The centrosomal kinase Aurora-A/STK15 interacts with a putative tumor suppressor NM23-H1. Nucleic Acids Res **30:**5465-75.

45. **Du, Y., S. E. Spence, N. A. Jenkins, and N. G. Copeland.** 2005. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. Blood **106:**2498-505.

46. **Dy, P., A. Penzo-Mendez, H. Wang, C. E. Pedraza, W. B. Macklin, and V. Lefebvre.** 2008. The three SoxC proteins--Sox4, Sox11 and Sox12--exhibit overlapping expression patterns and molecular properties. Nucl. Acids Res.**:**gkn162.

47. **Ernst, T., M. Hergenhahn, M. Kenzelmann, C. D. Cohen, M. Bonrouhi, A. Weninger, R. Klaren, E. F. Grone, M. Wiesel, C. Gudemann, J. Kuster, W. Schott, G. Staehler, M. Kretzler, M. Hollstein, and H. J. Grone.** 2002. Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. Am J Pathol **160:**2169-80.

48. **Evans, R., C. Naber, T. Steffler, T. Checkland, J. Keats, C. Maxwell, T. Perry, H. Chau, A. Belch, L. Pilarski, and T. Reiman.** 2008. Aurora A kinase RNAi and small molecule inhibition of Aurora kinases with VE-465 induce apoptotic death in multiple myeloma cells. Leuk Lymphoma **49:**559-69.

49. **Eyers, P. A., E. Erikson, L. G. Chen, and J. L. Maller.** 2003. A novel mechanism for activation of the protein kinase Aurora A. Curr Biol **13:**691-7.

50. **Fan, Z., P. J. Beresford, D. Y. Oh, D. Zhang, and J. Lieberman.** 2003. Tumor suppressor NM23-H1 is a granzyme A-activated DNase during CTL-mediated apoptosis, and the nucleosome assembly protein SET is its inhibitor. Cell **112:**659-72.

51. **Fang, G.** 2002. Checkpoint protein BubR1 acts synergistically with Mad2 to inhibit anaphase-promoting complex. Mol Biol Cell **13:**755-66.

52. **Farr, C. J., D. J. Easty, J. Ragoussis, J. Collignon, R. Lovell-Badge, and P. N. Goodfellow.** 1993. Characterization and mapping of the human SOX4 gene. Mamm Genome **4:**577-84.

53. **Fevre-Montange, M., J. Champier, A. Szathmari, A. Wierinckx, C. Mottolese, J. Guyotat, D. Figarella-Branger, A. Jouvet, and J. Lachuer.** 2006. Microarray analysis reveals differential gene expression patterns in tumors of the pineal region. J Neuropathol Exp Neurol **65:**675-84.

54. **Foster, J. W., M. A. Dominguez-Steglich, S. Guioli, G. Kowk, P. A. Weller, M. Stevanovic, J. Weissenbach, S. Mansour, I. D. Young, P. N. Goodfellow, and et al.** 1994. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature **372:**525-30.

55. **Friedman, R., C. Bangur, E. Zasloff, L. Fan, T. Wang, Y. Watanabe, and M. Kalos.** 2004. Molecular and Immunological Evaluation of the Transcription Factor SOX-4 as a Lung Tumor Vaccine Antigen. J Immunol **172:**3319-3327.

56. **Fu, J., M. Bian, Q. Jiang, and C. Zhang.** 2007. Roles of Aurora kinases in mitosis and tumorigenesis. Mol Cancer Res **5:**1-10.

57. **Fu, S., W. Hu, J. J. Kavanagh, and R. C. Bast, Jr.** 2006. Targeting Aurora kinases in ovarian cancer. Expert Opin Ther Targets **10:**77-85.

58. **Geijsen, N., I. J. Uings, C. Pals, J. Armstrong, M. McKinnon, J. A. Raaijmakers, J. W. Lammers, L. Koenderman, and P. J. Coffer.** 2001. Cytokine-specific transcriptional regulation through an IL-5Ralpha interacting protein. Science **293:**1136-8.

59. **Giannakakou, P., D. L. Sackett, Y. K. Kang, Z. Zhan, J. T. Buters, T. Fojo, and M. S. Poruchynsky.** 1997. Paclitaxel-resistant human ovarian cancer cells have mutant beta-tubulins that exhibit impaired paclitaxel-driven polymerization. J Biol Chem **272:**17118-25.

60. **Glover, D. M., M. H. Leibowitz, D. A. McLean, and H. Parry.** 1995. Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. Cell **81:**95-105.

61. **Goepfert, T. M., Y. E. Adigun, L. Zhong, J. Gay, D. Medina, and W. R. Brinkley.** 2002. Centrosome amplification and overexpression of aurora A are early events in rat mammary carcinogenesis. Cancer Res **62:**4115-22.

62. **Goldsworthy, M., A. Hugill, H. Freeman, E. Horner, K. Shimomura, D. Bogani, G. Pieles, V. Mijat, R. Arkell, S. Bhattacharya, F. M. Ashcroft, and R. D. Cox.** 2008. Role of the transcription factor sox4 in insulin secretion and impaired glucose tolerance. Diabetes **57:**2234-44.

63. **Goodwin, G. H., C. Sanders, and E. W. Johns.** 1973. A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. Eur J Biochem **38:**14-9.

64. **Graham, J. D., S. M. Hunt, N. Tran, and C. L. Clarke.** 1999. Regulation of the expression and activity by progestins of a member of the SOX gene family of transcriptional modulators. J Mol Endocrinol **22:**295-304.

65. **Griffiths-Jones, S., H. K. Saini, S. van Dongen, and A. J. Enright.** 2008. miRBase: tools for microRNA genomics. Nucleic Acids Res **36:**D154-8.

66. **Gritsko, T. M., D. Coppola, J. E. Paciga, L. Yang, M. Sun, S. A. Shelley, J. V. Fiorica, S. V. Nicosia, and J. Q. Cheng.** 2003. Activation and overexpression of centrosome kinase BTAK/Aurora-A in human ovarian cancer. Clin Cancer Res **9:**1420-6.

67. **Gubbay, J., J. Collignon, P. Koopman, B. Capel, A. Economou, A. Munsterberg, N. Vivian, P. Goodfellow, and R. Lovell-Badge.** 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature **346:**245-50.

68. **Guo, L., D. Zhong, S. Lau, X. Liu, X. Y. Dong, X. Sun, V. W. Yang, P. M. Vertino, C. S. Moreno, V. Varma, J. T. Dong, and W. Zhou.** 2008. Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells. Mol Cancer Res **6:**1421-30.

69. **Hanahan, D., and R. A. Weinberg.** 2000. The hallmarks of cancer. Cell **100:**57-70.

70. **Hannak, E., M. Kirkham, A. A. Hyman, and K. Oegema.** 2001. Aurora-A kinase is required for centrosome maturation in Caenorhabditis elegans. J Cell Biol **155:**1109-16.

71. **Harrington, E. A., D. Bebbington, J. Moore, R. K. Rasmussen, A. O. Ajose-Adeogun, T. Nakayama, J. A. Graham, C. Demur, T. Hercend, A. Diu-Hercend, M. Su, J. M. Golec, and K. M. Miller.** 2004. VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo. Nat Med **10:**262-7.

72. **Harris, B. Z., and W. A. Lim.** 2001. Mechanism and role of PDZ domains in signaling complex assembly. J Cell Sci **114:**3219-31.

73. **Hata, T., T. Furukawa, M. Sunamura, S. Egawa, F. Motoi, N. Ohmura, T. Marumoto, H. Saya, and A. Horii.** 2005. RNA interference targeting aurora kinase a suppresses tumor growth and enhances the taxane chemosensitivity in human pancreatic cancer cells. Cancer Res **65:**2899-905.

74. **Hayward, S. W., and G. R. Cunha.** 2000. The prostate: development and physiology. Radiol Clin North Am **38:**1-14.

75. **Herzog, T. J.** 2004. Recurrent Ovarian Cancer: How Important Is It to Treat to Disease Progression?
10.1158/1078-0432.CCR-04-0683. Clin Cancer Res **10:**7439-7449.

76. **Hirota, T., N. Kunitoku, T. Sasayama, T. Marumoto, D. Zhang, M. Nitta, K. Hatakeyama, and H. Saya.** 2003. Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells. Cell **114:**585-98.

77. **Honda, R., R. Korner, and E. A. Nigg.** 2003. Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. Mol Biol Cell **14:**3325-41.

78. **Hoser, M., S. L. Baader, M. R. Bosl, A. Ihmer, M. Wegner, and E. Sock.** 2007. Prolonged Glial Expression of Sox4 in the CNS Leads to Architectural Cerebellar Defects and Ataxia. J. Neurosci. **27:**5495-5505.

79. **Hoser, M., M. R. Potzner, J. M. Koch, M. R. Bosl, M. Wegner, and E. Sock.** 2008. Sox12 deletion in the mouse reveals non-reciprocal redundancy with the related Sox4 and Sox11 transcription factors. Mol Cell Biol.

80. **Hsu, J. Y., Z. W. Sun, X. Li, M. Reuben, K. Tatchell, D. K. Bishop, J. M. Grushcow, C. J. Brame, J. A. Caldwell, D. F. Hunt, R. Lin, M. M. Smith, and C. D. Allis.** 2000. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell **102:**279-91.

81. **Hu, W., J. J. Kavanagh, M. Deaver, D. A. Johnston, R. S. Freedman, C. F. Verschraegen, and S. Sen.** 2005. Frequent overexpression of STK15/Aurora-A/BTAK and chromosomal instability in tumorigenic cell cultures derived from human ovarian cancer. Oncol Res **15:**49-57.

82. **Hunt, S. M., and C. L. Clarke.** 1999. Expression and hormonal regulation of the Sox4 gene in mouse female reproductive tissues. Biol Reprod **61:**476-81.

83. **Hur, E. H., W. Hur, J. Y. Choi, I. K. Kim, H. Y. Kim, S. K. Yoon, and H. Rhim.** 2004. Functional identification of the pro-apoptotic effector domain in human Sox4. Biochem Biophys Res Commun **325:**59-67.

84. **Jeffers, J. R., E. Parganas, Y. Lee, C. Yang, J. Wang, J. Brennan, K. H. MacLean, J. Han, T. Chittenden, J. N. Ihle, P. J. McKinnon, J. L. Cleveland,**

and G. P. Zambetti. 2003. Puma is an essential mediator of p53-dependent and - independent apoptotic pathways. Cancer Cell **4:**321-8.

85. **Joukov, V., A. C. Groen, T. Prokhorova, R. Gerson, E. White, A. Rodriguez, J. C. Walter, and D. M. Livingston.** 2006. The BRCA1/BARD1 Heterodimer Modulates Ran-Dependent Mitotic Spindle Assembly. Cell **127:**539-52.

86. **Kamachi, Y., M. Uchikawa, A. Tanouchi, R. Sekido, and H. Kondoh.** 2001. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. Genes Dev **15:**1272-86.

87. **Karanam, S., and C. S. Moreno.** 2004. CONFAC: Automated Application of Comparative Genomic Promoter Analysis to DNA Microarray Datasets. Nucleic Acids Res **32:**W475-84.

88. **Katayama, H., K. Sasai, H. Kawai, Z. M. Yuan, J. Bondaruk, F. Suzuki, S. Fujii, R. B. Arlinghaus, B. A. Czerniak, and S. Sen.** 2004. Phosphorylation by aurora kinase A induces Mdm2-mediated destabilization and inhibition of p53. Nat Genet **36:**55-62.

89. **Katayama, H., H. Zhou, Q. Li, M. Tatsuka, and S. Sen.** 2001. Interaction and feedback regulation between STK15/BTAK/Aurora-A kinase and protein phosphatase 1 through mitotic cell division cycle. J Biol Chem **276:**46219-24.

90. **Keen, N., and S. Taylor.** 2004. Aurora-kinase inhibitors as anticancer agents. Nat Rev Cancer **4:**927-36.

91. **Kel, A. E., E. Gossling, I. Reuter, E. Cheremushkin, O. V. Kel-Margoulis, and E. Wingender.** 2003. MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res **31:**3576-9.

92. **Kim, B. E., J. H. Lee, H. S. Kim, O. J. Kwon, S. W. Jeong, and I. K. Kim.** 2004. Involvement of Sox-4 in the cytochrome c-dependent AIF-independent apoptotic pathway in HeLa cells induced by Delta12-prostaglandin J2. Exp Mol Med **36:**444-53.

93. **Kimmins, S., C. Crosio, N. Kotaja, J. Hirayama, L. Monaco, C. Hoog, M. van Duin, J. A. Gossen, and P. Sassone-Corsi.** 2007. Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. Mol Endocrinol **21:**726-39.

94. **Knox, R. J., F. Friedlos, D. A. Lydall, and J. J. Roberts.** 1986. Mechanism of cytotoxicity of anticancer platinum drugs: evidence that cis- diamminedichloroplatinum(II) and cis-diammined-(1,1- cyclobutanedicarboxylato)platinum(II) differ only in the kinetics of their interaction with DNA. Cancer Res **46:**1972-9.

95. **Krek, A., D. Grun, M. N. Poy, R. Wolf, L. Rosenberg, E. J. Epstein, P. MacMenamin, I. da Piedade, K. C. Gunsalus, M. Stoffel, and N. Rajewsky.** 2005. Combinatorial microRNA target predictions. Nat Genet **37:**495-500.

96. **Kufer, T. A., H. H. Sillje, R. Korner, O. J. Gruss, P. Meraldi, and E. A. Nigg.** 2002. Human TPX2 is required for targeting Aurora-A kinase to the spindle. J Cell Biol **158:**617-23.

97. **Kunitoku, N., T. Sasayama, T. Marumoto, D. Zhang, S. Honda, O. Kobayashi, K. Hatakeyama, Y. Ushio, H. Saya, and T. Hirota.** 2003. CENP-A phosphorylation by Aurora-A in prophase is required for enrichment of Aurora-B at inner centromeres and for kinetochore function. Dev Cell **5:**853-64.

98.   **Lapointe, J., C. Li, J. P. Higgins, M. Van De Rijn, E. Bair, K. Montgomery, M. Ferrari, L. Egevad, W. Rayford, U. Bergerheim, P. Ekman, A. M. DeMarzo, R. Tibshirani, D. Botstein, P. O. Brown, J. D. Brooks, and J. R. Pollack.** 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A **101:**811-6.

99.   **Lee, C. J., V. J. Appleby, A. T. Orme, W. I. Chan, and P. J. Scotting.** 2002. Differential expression of SOX4 and SOX11 in medulloblastoma. J Neurooncol **57:**201-14.

100.   **Lefebvre, V., B. Dumitriu, A. Penzo-Mendez, Y. Han, and B. Pallavi.** 2007. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol **39:**2195-214.

101.   **Lefebvre, V., P. Li, and B. de Crombrugghe.** 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J **17:**5718-33.

102.   **Leggate, D. R., J. M. Bryant, M. B. Redpath, D. Head, P. W. Taylor, and J. P. Luzio.** 2002. Expression, mutagenesis and kinetic analysis of recombinant K1E endosialidase to define the site of proteolytic processing and requirements for catalysis. Mol Microbiol **44:**749-60.

103.   **Lei, Q., J. Jiao, L. Xin, C. J. Chang, S. Wang, J. Gao, M. E. Gleave, O. N. Witte, X. Liu, and H. Wu.** 2006. NKX3.1 stabilizes p53, inhibits AKT activation, and blocks prostate cancer initiation caused by PTEN loss. Cancer Cell **9:**367-78.

104.   **Leong, K. G., and W. Q. Gao.** 2008. The Notch pathway in prostate development and cancer. Differentiation **76:**699-716.

105.   **Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge.** 2003. Prediction of mammalian microRNA targets. Cell **115:**787-98.

106.   **Li, J., H. Shen, K. L. Himmel, A. J. Dupuy, D. A. Largaespada, T. Nakamura, J. D. Shaughnessy, Jr., N. A. Jenkins, and N. G. Copeland.** 1999. Leukaemia disease genes: large-scale cloning and pathway predictions. Nat Genet **23:**348-53.

107.   **Li, X., V. Placencio, J. M. Iturregui, C. Uwamariya, A. R. Sharif-Afshar, T. Koyama, S. W. Hayward, and N. A. Bhowmick.** 2008. Prostate tumor progression is mediated by a paracrine TGF-beta/Wnt3a signaling axis. Oncogene **27:**7118-30.

108.   **Li, Z., O. S. Kustikova, K. Kamino, T. Neumann, M. Rhein, E. Grassman, B. Fehse, and C. Baum.** 2007. Insertional mutagenesis by replication-deficient retroviral vectors encoding the large T oncogene. Ann N Y Acad Sci **1106:**95-113.

109.   **Liao, Y. L., Y. M. Sun, G. Y. Chau, Y. P. Chau, T. C. Lai, J. L. Wang, J. T. Horng, M. Hsiao, and A. P. Tsou.** 2008. Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. Oncogene.

110.   **Lin, Y. G., A. Immaneni, W. M. Merritt, L. S. Mangala, S. W. Kim, M. M. Shahzad, Y. T. Tsang, G. N. Armaiz-Pena, C. Lu, A. A. Kamat, L. Y. Han, W. A. Spannuth, A. M. Nick, C. N. Landen, Jr., K. K. Wong, M. J. Gray, R.**

L. Coleman, D. C. Bodurka, W. R. Brinkley, and A. K. Sood. 2008. Targeting aurora kinase with MK-0457 inhibits ovarian cancer growth. Clin Cancer Res **14:**5437-46.

111. **Littlepage, L. E., H. Wu, T. Andresson, J. K. Deanehan, L. T. Amundadottir, and J. V. Ruderman.** 2002. Identification of phosphorylated residues that affect the activity of the mitotic kinase Aurora-A. Proc Natl Acad Sci U S A **99:**15440-5.

112. **Liu, P., S. Ramachandran, M. Ali-Seyed, C. D. Scharer, N. Laycock, W. B. Dalton, H. Williams, S. Karanam, M. W. Datta, D. Jaye, and C. S. Moreno.** 2006. Sex-Determining Region Y Box 4 is a Transforming Oncogene in Human Prostate Cancer Cells. Cancer Res **46:**4011-9.

113. **Liu, Q., S. Kaneko, L. Yang, R. I. Feldman, S. V. Nicosia, J. Chen, and J. Q. Cheng.** 2004. Aurora-A abrogation of p53 DNA binding and transactivation activity by phosphorylation of serine 215. J Biol Chem **279:**52175-82.

114. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods **25:**402-8.

115. **Lund, A. H., G. Turner, A. Trubetskoy, E. Verhoeven, E. Wientjens, D. Hulsman, R. Russell, R. A. DePinho, J. Lenz, and M. van Lohuizen.** 2002. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nat Genet **32:**160-5.

116. **Luo, J., D. J. Duggan, Y. Chen, J. Sauvageot, C. M. Ewing, M. L. Bittner, J. M. Trent, and W. B. Isaacs.** 2001. Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. Cancer Res **61:**4683-8.

117. **Ma, L., J. Teruya-Feldstein, and R. A. Weinberg.** 2007. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature **449:**682-8.

118. **Magee, J. A., T. Araki, S. Patil, T. Ehrig, L. True, P. A. Humphrey, W. J. Catalona, M. A. Watson, and J. Milbrandt.** 2001. Expression profiling reveals hepsin overexpression in prostate cancer. Cancer Res **61:**5692-6.

119. **Marumoto, T., S. Honda, T. Hara, M. Nitta, T. Hirota, E. Kohmura, and H. Saya.** 2003. Aurora-A kinase maintains the fidelity of early and late mitotic events in HeLa cells. J Biol Chem **278:**51786-95.

120. **Marumoto, T., D. Zhang, and H. Saya.** 2005. Aurora-A - a guardian of poles. Nat Rev Cancer **5:**42-50.

121. **Mavropoulos, A., N. Devos, F. Biemar, E. Zecchin, F. Argenton, H. Edlund, P. Motte, J. A. Martial, and B. Peers.** 2005. sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. Dev Biol **285:**211-23.

122. **Maxwell, C. A., J. J. Keats, A. R. Belch, L. M. Pilarski, and T. Reiman.** 2005. Receptor for hyaluronan-mediated motility correlates with centrosome abnormalities in multiple myeloma and maintains mitotic integrity. Cancer Res **65:**850-60.

123. **McCabe, C. D., D. D. Spyropoulos, D. Martin, and C. S. Moreno.** 2008. Genome-wide analysis of the homeobox C6 transcriptional network in prostate cancer. Cancer Res **68:**1988-96.

124. **McCracken, S., C. S. Kim, Y. Xu, M. Minden, and N. G. Miyamoto.** 1997. An alternative pathway for expression of p56lck from type I promoter transcripts in colon carcinoma. Oncogene **15:**2929-37.

125. **McGowan, E. M., and C. L. Clarke.** 1999. Effect of overexpression of progesterone receptor A on endogenous progestin-sensitive endpoints in breast cancer cells. Mol Endocrinol **13:**1657-71.

126. **Meraldi, P., R. Honda, and E. A. Nigg.** 2002. Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53-/- cells. Embo J **21:**483-92.

127. **Milne, T. A., S. D. Briggs, H. W. Brock, M. E. Martin, D. Gibbs, C. D. Allis, and J. L. Hess.** 2002. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell **10:**1107-17.

128. **Moreno, C. S., L. Matyunina, E. B. Dickerson, N. Schubert, N. J. Bowen, S. Logani, B. B. Benigno, and J. F. McDonald.** 2007. Evidence that p53-Mediated Cell-Cycle-Arrest Inhibits Chemotherapeutic Treatment of Ovarian Carcinomas. PLoS ONE **2:**e441.

129. **Moreno, C. S., S. Ramachandran, D. Ashby, N. Laycock, C. A. Plattner, W. Chen, W. C. Hahn, and D. C. Pallas.** 2004. Signaling and Transcriptional Changes Critical for Transformation of Human Cells by SV40 Small Tumor Antigen or PP2A B56gamma Knockdown. Cancer Res **64:**6978-6988.

130. **Morrish, B. C., and A. H. Sinclair.** 2002. Vertebrate sex determination: many means to an end. Reproduction **124:**447-57.

131. **Musacchio, A., and K. G. Hardwick.** 2002. The spindle checkpoint: structural insights into dynamic signalling. Nat Rev Mol Cell Biol **3:**731-41.

132. **Naylor, M. J., S. R. Oakes, M. Gardiner-Garden, J. Harris, K. Blazek, T. W. Ho, F. C. Li, D. Wynick, A. M. Walker, and C. J. Ormandy.** 2005. Transcriptional changes underlying the secretory activation phase of mammary gland development. Mol Endocrinol **19:**1868-83.

133. **Nissen-Meyer, L. S., R. Jemtland, V. T. Gautvik, M. E. Pedersen, R. Paro, D. Fortunati, D. D. Pierroz, V. A. Stadelmann, S. Reppe, F. P. Reinholt, A. Del Fattore, N. Rucci, A. Teti, S. Ferrari, and K. M. Gautvik.** 2007. Osteopenia, decreased bone formation and impaired osteoblast development in Sox4 heterozygous mice. J Cell Sci **120:**2785-95.

134. **Ohashi, S., G. Sakashita, R. Ban, M. Nagasawa, H. Matsuzaki, Y. Murata, H. Taniguchi, H. Shima, K. Furukawa, and T. Urano.** 2006. Phospho-regulation of human protein kinase Aurora-A: analysis using anti-phospho-Thr288 monoclonal antibodies. Oncogene **25:**7691-702.

135. **Orr, G. A., P. Verdier-Pinard, H. McDaid, and S. B. Horwitz.** 2003. Mechanisms of Taxol resistance related to microtubules. Oncogene **22:**7280-95.

136. **Otsuki, Y., M. Tanaka, S. Yoshii, N. Kawazoe, K. Nakaya, and H. Sugimura.** 2001. Tumor metastasis suppressor nm23H1 regulates Rac1 GTPase by interaction with Tiam1. Proc Natl Acad Sci U S A **98:**4385-90.

137. **Ouchi, M., N. Fujiuchi, K. Sasai, H. Katayama, Y. A. Minamishima, P. P. Ongusaha, C. Deng, S. Sen, S. W. Lee, and T. Ouchi.** 2004. BRCA1 phosphorylation by Aurora-A in the regulation of G2 to M transition. J Biol Chem **279:**19643-8.

138. **Padua, D., X. H. Zhang, Q. Wang, C. Nadal, W. L. Gerald, R. R. Gomis, and J. Massague.** 2008. TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. Cell **133:**66-77.
139. **Pan, X., H. Li, P. Zhang, B. Jin, J. Man, L. Tian, G. Su, J. Zhao, W. Li, H. Liu, W. Gong, T. Zhou, and X. Zhang.** 2006. Ubc9 interacts with SOX4 and represses its transcriptional activity. Biochem Biophys Res Commun **344:**727-34.
140. **Pan, X., J. Zhao, W. N. Zhang, H. Y. Li, R. Mu, T. Zhou, H. Y. Zhang, W. L. Gong, M. Yu, J. H. Man, P. J. Zhang, A. L. Li, and X. M. Zhang.** 2009. Induction of SOX4 by DNA damage is critical for p53 stabilization and function. Proc Natl Acad Sci U S A.
141. **Pearson, C. A., D. Pearson, S. Shibahara, J. Hofsteenge, and R. Chiquet-Ehrismann.** 1988. Tenascin: cDNA cloning and induction by TGF-beta. Embo J **7:**2977-82.
142. **Penzo-Mendez, A., P. Dy, B. Pallavi, and V. Lefebvre.** 2007. Generation of mice harboring a Sox4 conditional null allele. Genesis **45:**776-80.
143. **Pietsch, E. C., S. M. Sykes, S. B. McMahon, and M. E. Murphy.** 2008. The p53 family and programmed cell death. Oncogene **27:**6507-21.
144. **Pingault, V., N. Bondurand, K. Kuhlbrodt, D. E. Goerich, M. O. Prehu, A. Puliti, B. Herbarth, I. Hermans-Borgmeyer, E. Legius, G. Matthijs, J. Amiel, S. Lyonnet, I. Ceccherini, G. Romeo, J. C. Smith, A. P. Read, M. Wegner, and M. Goossens.** 1998. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet **18:**171-3.
145. **Polakis, P.** 2000. Wnt signaling and cancer. Genes Dev **14:**1837-51.
146. **Potzner, M. R., C. Griffel, E. Lutjen-Drecoll, M. R. Bosl, M. Wegner, and E. Sock.** 2007. Prolonged Sox4 expression in oligodendrocytes interferes with normal myelination in the central nervous system. Mol. Cell. Biol.**:**MCB.00339-07.
147. **Poulat, F., P. de Santa Barbara, M. Desclozeaux, S. Soullier, B. Moniot, N. Bonneaud, B. Boizet, and P. Berta.** 1997. The human testis determining factor SRY binds a nuclear factor containing PDZ protein interaction domains. J Biol Chem **272:**7167-72.
148. **Pramoonjago, P., A. S. Baras, and C. A. Moskaluk.** 2006. Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. Oncogene.
149. **Ramachandran, S., P. Liu, A. N. Young, Q. Yin-Goen, S. D. Lim, N. Laycock, M. B. Amin, J. K. Carney, F. F. Marshall, J. A. Petros, and C. S. Moreno.** 2005. Loss of HOXC6 Expression Induces Apoptosis in Prostate Cancer Cells. Oncogene **24:**188-98.
150. **Rana, T. M.** 2007. Illuminating the silence: understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol **8:**23-36.
151. **Reichling, T., K. H. Goss, D. J. Carson, R. W. Holdcraft, C. Ley-Ebert, D. Witte, B. J. Aronow, and J. Groden.** 2005. Transcriptional profiles of intestinal tumors in Apc(Min) mice are unique from those of embryonic intestine and identify novel gene targets dysregulated in human colorectal tumors. Cancer Res **65:**166-76.
152. **Remenyi, A., K. Lins, L. J. Nissen, R. Reinbold, H. R. Scholer, and M. Wilmanns.** 2003. Crystal structure of a POU/HMG/DNA ternary complex

suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes Dev **17:**2048-59.

153. **Ren, B., F. Robert, J. J. Wyrick, O. Aparicio, E. G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T. L. Volkert, C. J. Wilson, S. P. Bell, and R. A. Young.** 2000. Genome-wide location and function of DNA binding proteins. Science **290:**2306-9.

154. **Reppe, S., E. Rian, R. Jemtland, O. K. Olstad, V. T. Gautvik, and K. M. Gautvik.** 2000. Sox-4 messenger RNA is expressed in the embryonic growth plate and regulated via the parathyroid hormone/parathyroid hormone-related protein receptor in osteoblast-like cells. J Bone Miner Res **15:**2402-12.

155. **Rhodes, D. R., T. R. Barrette, M. A. Rubin, D. Ghosh, and A. M. Chinnaiyan.** 2002. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res **62:**4427-33.

156. **Rhodes, D. R., J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, T. Barrette, A. Pandey, and A. M. Chinnaiyan.** 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc Natl Acad Sci U S A **101:**9309-14.

157. **Robb, G. B., and T. M. Rana.** 2007. RNA helicase A interacts with RISC in human cells and functions in RISC loading. Mol Cell **26:**523-37.

158. **Ruebel, K. H., A. A. Leontovich, Y. Tanizaki, L. Jin, G. A. Stilling, S. Zhang, K. Coonse, B. W. Scheithauer, M. Lombardero, K. Kovacs, and R. V. Lloyd.** 2008. Effects of TGFbeta1 on gene expression in the HP75 human pituitary tumor cell line identified by gene expression profiling. Endocrine.

159. **Salaun, B., I. Coste, M. C. Rissoan, S. J. Lebecque, and T. Renno.** 2006. TLR3 can directly trigger apoptosis in human cancer cells. J Immunol **176:**4894-901.

160. **Scharer, C. D., N. Laycock, A. O. Osunkoya, S. Logani, J. F. McDonald, B. B. Benigno, and C. S. Moreno.** 2008. Aurora kinase inhibitors synergize with paclitaxel to induce apoptosis in ovarian cancer cells. J Transl Med **6:**79.

161. **Scharer, C. D., C. D. McCabe, M. Ali-Seyed, M. F. Berger, M. L. Bulyk, and C. S. Moreno.** 2009. Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells. Cancer Res **69:**709-17.

162. **Schepers, G. E., R. D. Teasdale, and P. Koopman.** 2002. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. Dev Cell **3:**167-70.

163. **Schilham, M. W., M. A. Oosterwegel, P. Moerer, J. Ya, P. A. de Boer, M. van de Wetering, S. Verbeek, W. H. Lamers, A. M. Kruisbeek, A. Cumano, and H. Clevers.** 1996. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature **380:**711-4.

164. **Seki, A., J. A. Coppinger, C. Y. Jang, J. R. Yates, and G. Fang.** 2008. Bora and the kinase Aurora a cooperatively activate the kinase Plk1 and control mitotic entry. Science **320:**1655-8.

165. **Sen, S., H. Zhou, and R. A. White.** 1997. A putative serine/threonine kinase encoding gene BTAK on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines. Oncogene **14:**2195-200.
166. **Sinclair, A. H., P. Berta, M. S. Palmer, J. R. Hawkins, B. L. Griffiths, M. J. Smith, J. W. Foster, A. M. Frischauf, R. Lovell-Badge, and P. N. Goodfellow.** 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature **346:**240-4.
167. **Sinner, D., J. J. Kordich, J. R. Spence, R. Opoka, S. Rankin, S. C. Lin, D. Jonatan, A. M. Zorn, and J. M. Wells.** 2007. Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. Mol Cell Biol **27:**7802-15.
168. **Sock, E., S. D. Rettig, J. Enderich, M. R. Bosl, E. R. Tamm, and M. Wegner.** 2004. Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. Mol Cell Biol **24:**6635-44.
169. **Stewart, S., and G. Fang.** 2005. Destruction box-dependent degradation of aurora B is mediated by the anaphase-promoting complex/cyclosome and Cdh1. Cancer Res **65:**8730-5.
170. **Stolt, C. C., P. Lommes, R. P. Friedrich, and M. Wegner.** 2004. Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. Development **131:**2349-58.
171. **Stolt, C. C., P. Lommes, E. Sock, M. C. Chaboissier, A. Schedl, and M. Wegner.** 2003. The Sox9 transcription factor determines glial fate choice in the developing spinal cord. Genes Dev **17:**1677-89.
172. **Stolt, C. C., S. Rehberg, M. Ader, P. Lommes, D. Riethmacher, M. Schachner, U. Bartsch, and M. Wegner.** 2002. Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. Genes Dev **16:**165-70.
173. **Subramanian, A., H. Kuehn, J. Gould, P. Tamayo, and J. P. Mesirov.** 2007. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics **23:**3251-3.
174. **Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov.** 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A **102:**15545-50.
175. **Sudbeck, P., and G. Scherer.** 1997. Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. J Biol Chem **272:**27848-52.
176. **Sun, L., A. M. Hui, Q. Su, A. Vortmeyer, Y. Kotliarov, S. Pastorino, A. Passaniti, J. Menon, J. Walling, R. Bailey, M. Rosenblum, T. Mikkelsen, and H. A. Fine.** 2006. Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. Cancer Cell **9:**287-300.
177. **Takahashi, K., and S. Yamanaka.** 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell **126:**663-76.

178. **Takahashi, T., M. Futamura, N. Yoshimi, J. Sano, M. Katada, Y. Takagi, M. Kimura, T. Yoshioka, Y. Okano, and S. Saji.** 2000. Centrosomal kinases, HsAIRK1 and HsAIRK3, are overexpressed in primary colorectal cancers. Jpn J Cancer Res **91:**1007-14.

179. **Talantov, D., A. Mazumder, J. X. Yu, T. Briggs, Y. Jiang, J. Backus, D. Atkins, and Y. Wang.** 2005. Novel genes associated with malignant melanoma but not benign melanocytic lesions. Clin Cancer Res **11:**7234-42.

180. **Tam, P. P., M. Kanai-Azuma, and Y. Kanai.** 2003. Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. Curr Opin Genet Dev **13:**393-400.

181. **Tanaka, T., M. Kimura, K. Matsunaga, D. Fukada, H. Mori, and Y. Okano.** 1999. Centrosomal kinase AIK1 is overexpressed in invasive ductal carcinoma of the breast. Cancer Res **59:**2041-4.

182. **Tang, C. J., C. Y. Lin, and T. K. Tang.** 2006. Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. Dev Biol **290:**398-410.

183. **Tatsuka, M., H. Katayama, T. Ota, T. Tanaka, S. Odashima, F. Suzuki, and Y. Terada.** 1998. Multinuclearity and increased ploidy caused by overexpression of the aurora- and Ipl1-like midbody-associated protein mitotic kinase in human cancer cells. Cancer Res **58:**4811-6.

184. **Tavazoie, S. F., C. Alarcon, T. Oskarsson, D. Padua, Q. Wang, P. D. Bos, W. L. Gerald, and J. Massague.** 2008. Endogenous human microRNAs that suppress breast cancer metastasis. Nature **451:**147-52.

185. **Terada, Y., M. Tatsuka, F. Suzuki, Y. Yasuda, S. Fujita, and M. Otsu.** 1998. AIM-1: a mammalian midbody-associated protein required for cytokinesis. Embo J **17:**667-76.

186. **Trapman, J., and A. O. Brinkmann.** 1996. The androgen receptor in prostate cancer. Pathol Res Pract **192:**752-60.

187. **Trieselmann, N., S. Armstrong, J. Rauw, and A. Wilde.** 2003. Ran modulates spindle assembly by regulating a subset of TPX2 and Kid activities including Aurora A activation. J Cell Sci **116:**4791-8.

188. **Tusher, V. G., R. Tibshirani, and G. Chu.** 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A **98:**5116-21.

189. **van de Wetering, M., M. Oosterwegel, K. van Norren, and H. Clevers.** 1993. Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. Embo J **12:**3847-54.

190. **Varambally, S., J. Yu, B. Laxman, D. R. Rhodes, R. Mehra, S. A. Tomlins, R. B. Shah, U. Chandran, F. A. Monzon, M. J. Becich, J. T. Wei, K. J. Pienta, D. Ghosh, M. A. Rubin, and A. M. Chinnaiyan.** 2005. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. Cancer Cell **8:**393-406.

191. **Wagner, T., J. Wirth, J. Meyer, B. Zabel, M. Held, J. Zimmer, J. Pasantes, F. D. Bricarelli, J. Keutel, E. Hustert, and et al.** 1994. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. Cell **79:**1111-20.

192. **Wang, X. D., C. C. Leow, J. Zha, Z. Tang, Z. Modrusan, F. Radtke, M. Aguet, F. J. de Sauvage, and W. Q. Gao.** 2006. Notch signaling is required for normal prostatic epithelial cell proliferation and differentiation. Dev Biol **290:**66-80.

193. **Wang, X. X., R. Liu, S. Q. Jin, F. Y. Fan, and Q. M. Zhan.** 2006. Overexpression of Aurora-A kinase promotes tumor cell proliferation and inhibits apoptosis in esophageal squamous cell carcinoma cell line. Cell Res **16:**356-66.

194. **Weiss, J., J. J. Meeks, L. Hurley, G. Raverot, A. Frassetto, and J. L. Jameson.** 2003. Sox3 is required for gonadal function, but not sex determination, in males and females. Mol Cell Biol **23:**8084-91.

195. **Welsh, J. B., L. M. Sapinoso, A. I. Su, S. G. Kern, J. Wang-Rodriguez, C. A. Moskaluk, H. F. Frierson, Jr., and G. M. Hampton.** 2001. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res **61:**5974-8.

196. **Wilson, M., and P. Koopman.** 2002. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr Opin Genet Dev **12:**441-6.

197. **Wilson, M. E., K. Y. Yang, A. Kalousova, J. Lau, Y. Kosaka, F. C. Lynn, J. Wang, C. Mrejen, V. Episkopou, H. C. Clevers, and M. S. German.** 2005. The HMG Box Transcription Factor Sox4 Contributes to the Development of the Endocrine Pancreas. Diabetes **54:**3402-9.

198. **Wotton, D., R. A. Lake, C. J. Farr, and M. J. Owen.** 1995. The high mobility group transcription factor, SOX4, transactivates the human CD2 enhancer. J Biol Chem **270:**7515-22.

199. **Wright, E. M., B. Snopek, and P. Koopman.** 1993. Seven new members of the Sox gene family expressed during mouse development. Nucleic Acids Res **21:**744.

200. **Wu, X., X. Tu, K. S. Joeng, M. J. Hilton, D. A. Williams, and F. Long.** 2008. Rac1 activation controls nuclear localization of beta-catenin during canonical Wnt signaling. Cell **133:**340-53.

201. **Wu, X., J. Wu, J. Huang, W. C. Powell, J. Zhang, R. J. Matusik, F. O. Sangiorgi, R. E. Maxson, H. M. Sucov, and P. Roy-Burman.** 2001. Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. Mech Dev **101:**61-9.

202. **Yang, H., L. He, P. Kruk, S. V. Nicosia, and J. Q. Cheng.** 2006. Aurora-A induces cell survival and chemoresistance by activation of Akt through a p53-dependent manner in ovarian cancer cells. Int J Cancer.

203. **Ylostalo, J., J. R. Smith, R. R. Pochampally, R. Matz, I. Sekiya, B. L. Larson, J. T. Vuoristo, and D. J. Prockop.** 2006. Use of differentiating adult stem cells (marrow stromal cells) to identify new downstream target genes for transcription factors. Stem Cells **24:**642-52.

204. **Yuan, H., N. Corbi, C. Basilico, and L. Dailey.** 1995. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev **9:**2635-45.

205. **Zeitlin, S. G., R. D. Shelby, and K. F. Sullivan.** 2001. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol **155:**1147-57.

206. **Zeller, K. I., X. Zhao, C. W. Lee, K. P. Chiu, F. Yao, J. T. Yustein, H. S. Ooi, Y. L. Orlov, A. Shahab, H. C. Yong, Y. Fu, Z. Weng, V. A. Kuznetsov, W. K. Sung, Y. Ruan, C. V. Dang, and C. L. Wei.** 2006. Global mapping of c-Myc binding sites and target gene networks in human B cells. Proc Natl Acad Sci U S A **103:**17834-9.

207. **Zhang, D., T. Hirota, T. Marumoto, M. Shimizu, N. Kunitoku, T. Sasayama, Y. Arima, L. Feng, M. Suzuki, M. Takeya, and H. Saya.** 2004. Cre-loxP-controlled periodic Aurora-A overexpression induces mitotic abnormalities and hyperplasia in mammary glands of mouse models. Oncogene **23:**8720-30.

208. **Zhou, H., J. Kuang, L. Zhong, W. L. Kuo, J. W. Gray, A. Sahin, B. R. Brinkley, and S. Sen.** 1998. Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. Nat Genet **20:**189-93.