The Utilization of Mouse Models to Study Gene Functions: The Role of Foxn3 and Chd2 in Murine Development and Cancer

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

George Samaan

December 2011

Abstract

Murine model organisms are an essential tool in the scientific community quest to decipher the molecular etiology of human diseases. Currently, several methods are used to induce or reproduce human diseases in mouse models using advanced genetic engineering techniques to mutate the wild-type genes. We utilized the Baygenomics gene-trap method to study the effects of two mammalian genes: FOXN3 and CHD2. The Forkhead Box (FOX) family of transcription factors shares a common DNA-binding domain and has been associated with organ development, differentiation, cell growth and proliferation, and cancer. Meanwhile, the CHD (Chromodomain helicase DNA binding protein) family of proteins is known to be involved in chromatin remodeling and regulation of gene expression. Phenotypic analysis of Foxn3 mutant animals revealed its indispensible role in craniofacial and embryonic development, embryonic lethality, expression of bone morphogenetic proteins, and spontaneous development of cancers in heterozygous and homozygous mutant mice. Preliminary evaluation of molecular mechanisms of FOXN3 signifies deregulation of cell-cycle checkpoint proteins Cyclin-B1 and CDK2 as the underlying etiology of tumors. Chd2 mutant mice exhibit spontaneous thymic and splenic lymphomas and reduced lifespan which can be restored through Chd2 re-expression in the thymus. At the molecular level, CHD2 deficiency reduces Puma (p53-upregulated modulator of apoptosis) induction after DNA damage in mouse thymocytes and HCT116 cells. Additionally, CHD2 is enriched at the *Puma* locus after DNA damage. CHD2-deficient cells also exhibit global reduction of active transcription markers H3K9-Acetylated and H4K8-Acetylated.

Table of Contents

List of Figures

List of Tables

Chapter I

The Utilization of Animal Models to Study Genes in Development and Disease

Introduction to Model Organisms

A model organism is a class of living organisms extensively used to identify or perform in-depth studies on a biological process. Because many aspects of biology are similar in most or all organisms, it is typically easier to study a particular aspect of a biological process in one organism over the others. Selection of proper model organism to conduct research is of utmost importance; principal investigators must select the model organism depending on the echelon of biology pursued, as in macro- versus micro- biology. The most popular model organisms have strong advantages to perform experimental research, such as rapid development with short life cycles, small adult size, constant availability, ease of handling, and common usage by other scientists (1). Model organisms can be generally classified into three main categories: mammalian models, non-mammalian models, and plants. For this dissertation, we will focus heavily on mammalian models and some non-mammalian models; plant model organisms such as *Arabidopsis thaliana* will not be discussed. Non-mammalian model organisms include members of all three domains of modern taxonomy: Archaea, Bacteria, and Eukarya. Mammalian model organisms represent the Mammalia class within the Animalia kingdom of the Eukarya domain.

Non-mammalian Model Organisms

Non-mammalian model organisms spread over the three domain of living organisms and vary greatly in tissue complexity and even cellular components. However, they are very advantageous to study fundamental cellular processes that are

inherently conserved in higher organisms. On genetic basis, our understanding of gene function is typically inferred by comparing the phenotypic consequences of mutating the canonical “wild-type” gene in one organism and relating this phenotype to mutants of orthologous genes in a different organism (2). One of the great advantages to using non-mammalian model organisms is the availability of whole-genome sequencing for many of these species, thus enabling scientists to compare genomic sequences and conservation of orthologous genes. Many scientists work on specific organisms or genus to uncover specialized processes; nevertheless, most scientists choose one of the NIH (National Institute of Health) supported model organisms.

Escherichia coli. E. *coli* is a prokaryotic organism (Bacteria domain) widely used for basic scientific research. E. *coli* has been widely used to obtain basic biological processes including information on molecular biology and gene regulation, metabolic pathways and biochemistry (3). One of the major disadvantages to using a prokaryotic system is the lack of a nuclear membrane to compartmentalize the nucleus and transcriptional machinery. Nevertheless, the fast growth rate (0.5 hours / duplication) has made the utilization of prokaryotes for generating recombinant DNA and expressing recombinant proteins a tremendous asset for scientists. E. *coli* was used to synthesize human insulin through genetic engineering, and research utilizing E. *coli* remains active to find alternative energy resources (4).

Saccharomyces cerevisiae. S. *cerevisiae* (baker’s yeast, here forth referred to as yeast) is the simplest model organism of the Eukarya domain. Yeast is widely used to study basic eukaryotic cell biology and genetics, and the results obtained from these studies impact our understanding of cellular functions higher eukaryotes. The ease of

genetic manipulation of yeast and its genetic complexity allows scientists to analyze gene function (5). Yeast has been utilized to dissect an array of cellular functions ranging from replication, transcription, cell cycle regulation and mRNA splicing, to more complex processes such as cell signaling pathways, homologous recombination and G-protein coupled receptors (6). Yeast has a fairly simplified genetic makeup (~4000 genes) for eukaryotic organisms with rare horizontal transfer gain of genes or gene duplications (7).

Caenorhabditis elegans. C. *elegans* is a small, free-living soil nematode (round worm) that has been used as the simplest multi-cellular eukaryotic organism. C*. elegans* has organ systems, complex sensory systems and shows coordinated behavior, thus making it a unique, yet simple animal to study at the embryological, neurobiological and cellular levels. C*. elegans* can be grown and genetically manipulated with the speed and ease of a micro-organism, making it a perfect organism to for genetic studies and genomic manipulation (8). Furthermore, C*. elegans* is transparent, which allows for visualization using a dissecting microscope to monitor cell movement. C*. elegans* has been at the forefront of small RNA research and their regulation of gene expression (9). Initial discovery of the first microRNAs (miRNA) and their function as regulators of gene expression was discovered in C. *elegans*. Moreover, double-stranded RNA-induced gene silencing and the current understanding of the RNA interference (RNAi) machinery were also first discovered in C. *elegans* (10). This discovery unveiled a new mechanism for gene regulation that is conserved in higher eukaryotes. Gene-function studies in C. *elegans* remain easiest for the simple model of delivering double stranded RNA through microinjecting the worm or feeding the worm bacteria that expresses the dsRNA.

Drosophila melanogaster. D. *melanogaster*, commonly known as the fruit fly, is one of the most commonly used organisms for classical and molecular genetics research. Use of *Drosophila* as a model organism is alluring for many reasons including its easy-to-manipulate genetic system, relatively low cost and fast generation time, and biological complexity comparable to mammals (11). Many genes in the *Drosophila* genome have well conserved orthologues in mammals with highly conserved epigenetic mechanisms. *Drosophila* has been utilized to extensively study segmentation of its body plan and the control of homeotic genes in this segmentation (12). Embryology and development and oocyte polarity is another field where *Drosophila* has been used extensively to unravel the basic mechanisms behind these fundamental processes. *Drosophila* is also extensively used to study epigenetic transcriptional control, insulator functions, and chromatin modifications because of the ease of mapping of its polytene chromosome. *Drosophila* has long been used to study classical genetics and possess many unique genetic markers that are manifested in visible phenotypes that can be seen with the naked eye or under low-powered dissecting microscopes (12). The utilization of recessive balancer chromosomes along with genetic markers makes it easy to select for progeny with the desired homozygous mutation.

Other non-mammalian organisms. Other NIH sponsored non-mammalian organisms include organisms from the Archea domain. Retroviruses, which are not classified as living organisms, are also sponsored organisms by NIH, as well as the parasites from the *plasmodium* genus, which cause malaria. Lastly, ***Dictyostelium*** *discoideum* (social amoeba) and *Danio rerio* (zebrafish) are the other non-mammalian model

organisms utilized in modern research laboratories for various studies including reproduction.

Mammalian Model Organisms

Mammalian model organisms cover the mammalia class of the Animal kingdom and are heavily utilized in modern science to research and investigate human disease. Human subjects are not used for research purposes because of the obvious risk to life, except in highly regulated clinical trials. Therefore, animal models are employed in the study of human disease and they are chosen because of their similarity to humans on genetic, anatomical, and physiological basis (13). Experimental research of diseases is done on animal models because of their unlimited supply, ease of manipulation, and controlled settings. To obtain scientifically valid research, the conditions associated with an experiment must be closely controlled through manipulating only one variable while keeping all others constant. Scientists then monitor the consequences of changing the variable and record the observations to test their hypotheses. Since most experiments require an adequate number of subjects for testing, it is easier for scientists to use large numbers of animals to conduct their research. Animal models are used to study the genetic etiology of diseases, as well as diagnosis and treatments of such disease. Rodents are the most widely used NIH-sponsored mammalian model organisms due to their similar physiology to humans and high genomic homology; however some studies can use primates for more accurate studies of human diseases. The two rodent models most used in scientific research are the mouse (*Mus musculus*) and the rat (*Rattus norvegicus*). For the purposes of this dissertation, we will focus on the mouse as a mammalian model organism to decipher gene functions.

Mus Musculus. The common house mouse is one of the most commonly used animals to study genetic principles and human disease. Humans and mice diverged 75 million years ago, since then, their genomes have evolved and changed at a rate of nearly one substitution for every two nucleotides and by deletion and insertion mutations (14). Yet, ortholgous sequences can still be aligned to resemble the divergence and conservation of essential genes. The mouse genome is 14% smaller than the human genome in terms of euchromatic regions with 2.5Gb (19 autosomes, 1 XY chromosomes) compared to 2.9 Gb (22 autosomes, 1 XY chromosomes) (14). The human genome has undergone expansion since the divergence of humans and mice with an overall ratio of 0.91 mouse/human locus span (measured in relative size of orthologous landmarks in the genome). The human gene catalogue predicts 27,049 transcripts aggregated into 22,808 distinct predicted genes, while the mouse gene catalogue predicts 29,201 transcripts clustered in 22,011 distinct genes. Moreover, 80% of mouse genes have a strict 1:1 orthologue in humans, and the remaining genes belong to a gene family that has undergone expansion. This shows the high degree of homology between human and mouse genomes and the ability to correlate human and mouse gene functions. Since both genomes have been sequenced and comparative analysis of both genomes is established, the biomedical studies of human genes can be complemented by experimental manipulations of a corresponding mouse gene, allowing fast and reliable functional understanding of gene functions. In respect to the mouse genome, there are practical techniques for random mutagenesis along with hundreds of spontaneous mutations occurring in mice. However, the most important aspect of mouse biology is the ability to change the genome through transgenic, knockout and

knock-in techniques, allowing precise manipulation of the mouse genome (15). Thus, the utilization of mouse is unmatched as a model system for probing mammalian biology and human disease.

Methods for Inducing Human Disease.

Despite similar genetic homology between humans and mice, replicating human disease conditions in mouse is complex due to the heterogeneity and outbred nature of human alleles compared to the inbred strains of mice. Many human congenital and developmental diseases are manifested as a result of chromosomal deletions and aberrations of alleles that can eliminate multiple genes within the region (16). Moreover, human diseases can be modified through single nucleotide polymorphisms (SNPs) and copy number variants (CNVs), affecting penetrance and responsiveness to treatments. Such studies are hard to duplicate in murine systems compared to modern reverse genetics approaches which have undertaken the majority of active research. These reverse genetics studies rely on various methods to interfere with gene function to study the phenotype and effect of this gene on homeostasis. Selection of proper interference method can ultimately affect the expressivity of the phenotype and the correlation of the model to human diseases; therefore, we will discuss some of the current methods used in creating genetically-modified mice.

Large scale mutation screens. *N*-ethyl-*N*-nitrosourea (ENU) is an efficient mutagen used in the spermatogonia stem cells of mice to induce an array of different point mutations. The progeny mice can be identified based on alterations to the wild type phenotype and expansion of certain line based on the desired phenotype. On average, a mutation is predicted every 2.38 Mb of gene-encoding sequences and a functional

mutation is expected every 9.48 Mb, thus making mutagenesis dispersed and random (17). Although this technique is advantageous in generating multiple variable mutant progeny, there are disadvantages to this technique. First, it's difficult to establish the main causative mutation of the phenotype and requires multiple complementation crosses, and secondly, it's difficult to assess whether one gene mutation or two synergistic gene mutations contribute to the phenotype. One of the most successful ENU-generated and most utilized mouse models is the multiple intestinal neoplasia APC (Apc^{Min}) mouse model, which was the first mouse model to recapitulate the human Familial Adenomatous Polyposis (FAP) disease (18). Apc^{Min} mice are heterozygous for a germ-line mutation in *Apc*, and loss of heterozygosity due to somatic mutations renders the protein non-functional and activates the Wnt signaling pathway.

Other large scale mutations involve radiation induced (e.g. X-ray) chromosomal breaks and rearrangements (19). The main drawback to this mutagenesis system is the excessive gene translocations and deletions, and chromosomal rearrangements. On the other hand, this may simulate some human diseases where loci deletions and chromosome aberrations underlie the disease. IR-induced genomic instability is utilized as a method to induce cancers in mice mutant for DNA repair genes or any other gene to test the tumorogenic potential of the specific gene (20).

Transgenic mice. Transgenic mice are generated by injecting foreign genetic material either directly into the fertilized egg through 'pronuclear microinjection', or injecting the foreign DNA into nucleus of embryonic stem cells which are eventually injected into the blastocysts of a fertilized egg (21). The content of the foreign genetic material integrates into the genome of the embryo where it can modify the organism. Transgenic mice

expressing the exogenous Maloney leukemia virus were the first transgenic mice created (22). Transgenes have a powerful promoter (e.g. CMV) upstream of the coding sequence to drive efficient transcription of the gene. However, advancements of genetic engineering have made transgenic mice more useful in biomedical research. Tissue-specific expression of exogenous genes has become a more useful model system that depends on cloning a tissue-specific gene promoter upstream of the exogenous gene construct (23). More recently, cloning a tamoxifen response element in the promoter of exogenous genes induces gene expression through intravenous injection of the hormone tamoxifen (24). However, the main drawback to transgene technology is that integration of the foreign DNA into the genome occurs at random, and the number of copies integrated varies. Therefore, transgene technology cannot be used to manipulate endogenous genes in a predetermined manner.

Knock-Out and Knock-In mice. When a single gene is predetermined to trigger a specific disease, targeted mutagenesis of this gene is essential to elucidate molecular data of the etiology of the disease. Targeting a predetermined gene utilizes a molecular technique that is both cumbersome and inefficient: homologous recombination (HR) (19). The desired mutation is created through *in-vitro* gene targeting of Embryonic stem (ES) cells which are then microinjected into blastocysts and eventually a pseudo-pregnant female for embryo development. The mutations in ES cells will be propagated through all tissue of the adult mouse after a series of crossings and selections based on coat color. In general, targeting of desired gene can result in ‘Knock-out’ of gene resulting in loss of function in the null allele, gain of function through ‘Knock-in’ of a transgene, or point mutation through exchange of an exon with a mutant exon.

Targeting vector or foreign DNA material typically carries characteristic sequences and markers for selection, in addition to the mutation of interest. Homologous recombination requires sequences that are similar to the target region of interest where strand exchange can take place (25). The vector also contains a neomycin resistance gene within the targeting vector to positively select for ES cells that have undergone recombination. A Thymine Kinase (TK) gene is found at in the targeting vector just outside the HR exchange region, whereby if a cell doesn't undergo proper HR, TK will be expressed and the ES cell will die. Although expensive, HR-mediated gene targeting has become the most effective way of studying genes and their role in diseases.

Conditional gene modifications. HR mediated gene targeting sprung a new approach to targeting mutagenesis in mouse models. Since many genes are embryonic lethal if deleted, and many genes become inactivated in adults, a system that removes gene function at specific times or in adult organism became a more realistic rendition of human diseases (25). Scientists started targeting their genes by adding a flanking sequence termed 'LoxP sites' in their targeting vectors, in addition to the original sequence of interest. Cre recombinase is an enzyme that recognizes these LoxP sites and floxes them (deletes the intervening DNA depending on their orientation), which often leads to deletion of an exon within the target gene. This creates a conditional gene knock-out phenotype that depends on the expression of the Cre recombinase enzyme. To add specificity to the system, Cre recombinase transgenic mice can be crossbred with knock-out mice to create mice with both features in their genome. Regulation of the *Cre* recombinase gene can be done at the promoter by adding a tissue-specific promoter that allows expression of Cre in a tissue specific lineage rather than whole

genome, thus targeting mutagenesis to the specified tissue. Additionally, a tamoxifen response element can be used in the *Cre* promoter, thus allowing floxing of the gene upon injection of estrogen to the bloodstream of mice (24).

Mouse Models for Tumorogenesis.

There are many paradigms that underscore the translational efficacy of mouse models in understanding human disease. Despite having a high degree of genetic homology between both organisms, some studies that were carried out in mice don't translate into humans (26). There are well over 1000 mutant strains of mice and the majority replicates an inherited human genetic disease or a human disease phenotype. It's important to note that few models replicate all symptoms associated with complex human diseases. Diseases that are hard to replicate in mice, but may otherwise be studied in primates, include cognitive and neurodegenerative disorders (27). However, one of the greater successes of mouse models comes from studying tumorogenesis and the role of genes in promoting cancers. Neoplasia, abnormal proliferation of cells, is a hallmark biological process during the multistep development of neoplasm (tissue mass or cancer) (28). Mouse models have given us tremendous clues into identifying 'Driver mutations', a genetic modification that provides an advantage to cells to facilitate tumor formation and survival (29). Gradually, the neoplasia accumulates more 'passenger' mutations in other genes important for DNA repair, cell proliferation and death, angiogenesis and invasion, leading to neoplasm and metastasis. Mouse models have played a great role in understanding these processes and identifying critical gene mutations to drive tumorogenesis. Thus we will review some of the important models.

Breast cancer models. Breast cancer is one of the most diagnosed cancers in females and one of the primary sources of cancer-related deaths in the world. Mammary tissue in mammals undergo changes during pregnancy, lactation, and weaning of offspring, and therefore constantly undergoing angiogenesis, cell proliferation, and eventually senescence and apoptosis (30). Mammals share this phenomenon in all females, and studies on mice can ascertain the molecular mechanisms of the disease condition in a conserved manner. Transgenic mice that over-expressed the oncogene *Myc* in the mammary epithelium formed tumors in that region (31). The same group later reported on the synergistic roles of *H-Ras* and *Myc* in mammary tumors; these strains, in addition to others (ErbB2, TGFα) are still used to test drugs in mice that have breast cancer (30, 32, 33). Later studies revealed that Brca1 (breast cancer 1, early onset) inactivation in mouse mammary tissue leads to genomic instability and tumor formation. This gene was then found to be mutated in a great percentage of women who have early onset breast cancer, and became a marker for breast cancer (34). Subdermal injection of foreign tissue, also known as Xenograft transplantation of tumors, is a mechanism to study tumor invasion and metastasis in mice, which lead to a better understanding of mutations that aide in angiogenesis, epithelial-mesenchymal transition, transplantation in other organs, and other features of breast cancer metastasis (30). Mouse models remain a useful tool for pharmaceutical companies and their active research to test new therapies for breast cancer.

Colorectal cancer models. Colorectal cancer (CRC) is a heterogeneous disease that can be caused by accumulating somatic mutations due to aging, environmental and chemical exposure, or diet leading to sporadic CRC (35). CRC represents the third most

frequently diagnosed form of cancer. Familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC) are the most common types of hereditary CRC. Patients predisposed to FAP typically carry a germline mutation for the tumor suppressor gene adenomatous polyposis coli (*APC*) gene (36). Additionally, various somatic mutations of *APC* occur in about 85% of all sporadic CRC, and the majority of these mutations occur in the mutation cluster region (MCR) of the gene. Several mouse models that study early events of CRC have been developed and can be divided into three major pathways; the mismatch repair pathway, the non-Wnt pathway, and the Wnt pathway (37). The mismatch repair pathway was linked to HNPCC when mice deficient for DNA mismatch repair (MMR) proteins *Mhl1, Msh2,* and *Msh6* developed gastric, intestinal, and colorectal cancers (38). The non-Wnt pathway mediated mouse models examine the role of transforming growth factor beta (TGF-β) in CRC and inflammation-mediated carcinogenesis (37). Studies on mice with TGF-β mutations showed a tumor suppressor role in early CRC and an enhancer of invasion and metastasis in late CRC. The Wnt pathway mutant mouse models provided a link between hereditary FAP and *Apc* mutations. The multiple intestinal neoplasia APC (Apc^{Min}) mouse model was the first mouse model to recapitulate the human FAP disease and continues to be utilized as a model for synergistic activation of onco-genes in tumor progression (39). CRC mouse models can be use to study carcinogenic molecules and chemicals, chemotherapeutic agents, and the roles of everyday drugs like NSAIDs in CRC prevention and treatment.

Models for other cancer hallmarks. There are other mouse models of human cancer which can't be categorized into a specific organ or class. However, there are many

models that describe other fundamental processes of cancer formation, progression, or invasion, and these models cover a wide spectrum of genes. Hanahan and Weinberg recently characterized the six biological hallmarks of cancer to be: 'sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis' (28). Perhaps one gene above all else has been the centerfold of all cancer research, the tumor suppressor protein p53. Deregulation of p53 activity is seen in approximately 70% of all cancers, and these are conserved estimates (40). Because of the complexity of the p53 pathway, its interacting proteins, and its direct-induced genes, it's likely that p53 contributes to all six fundamental hallmarks of cancer, hence the term' guardian angel of the genome'. Several p53 mutant mouse models exist to study the extensive phenotypes caused by various mutations in p53. To mimic human mutations in p53 which are typically single nucleotide polymorphisms (SNP), recent work has shifted towards recreating these SNPs in mouse models (41-43). These studies may contribute to the molecular mechanisms of increased tumor susceptibility of p53 polymorphisms. Future studies may also look into the effect of polymorphism of p53 effectors and response genes in tumor susceptibility and responsiveness to therapeutic agents. ATM, p21, and MDM2 are p53 pathway genes that have been reported to have polymorphism that contributes to cancer susceptibility, while polymorphisms in the promoters of p53 target genes (*BAX, BCL2, PIG3*) have been reported to affect tumor susceptibility (43).

Chapter II

Eukaryotic Transcription Regulation, and The Forkhead Box Family of Proteins

Introduction to Eukaryotic Transcription Regulation

Transcription is the fundamental process of the cell by which genetic information stored in DNA is activated through the synthesis of a complementary messenger RNA. Transcription is a highly regulated process in Eukaryotic cells at hierarchical levels that include pre-initiation, elongation, termination and epigenetic chromatin structure. Eukaryotic RNA Polymerase II, which is responsible for transcription of protein-coding genes, is composed of multiple subunits but lacks the capacity to initiate transcription in a purified form (44). Instead, the RNA polymerase requires the activity of general transcription factors to recognize a DNA sequence and initiate transcription of all genes; these 'basal' transcription factors are TFIIB, D, E, F, and H (45). Schaffner and Chambon identified gene-specific 'enhancer' elements that bind 'activator proteins' to allow transcription of specific genes (46). Kornberg purified the 'mediator' complex that links the regulatory elements of specific gene expression with general transcription factors and RNA polymerase. This introduced the concept of *cis*- and *trans*-regulatory elements that control the activation of specific-gene transcription. Transcription factors, known as *trans*-regulatory elements, can bind to specific sequences in upstream regulatory regions and gene promoters, known as *cis*-regulatory elements to selectively activate gene expression in a specific manner.

Transcription Factors

Eukaryotic transcription regulation is mediated through transcription factors (*trans*-regulatory elements) that can act as either 'Activators' or 'Repressors' of

transcription. The primary domains of all transcription factors (TFs) include a DNA binding domain and trans-activation domain; however, many TFs include an additional ligand-binding domain (47). The majority of transcription factors possess DNA-binding domains that can recognize specific DNA sequences formed at the major grooves (48). DNA-binding domains can be generally broken down to one of five major classes: Basic domains (helix-loop-helix), helix-turn-helix (e.g. Homeodomains, winged/forkhead), Zinc finger domains, β-scaffold domains (e.g. MADs box), or steroid hormone receptor (49). These domains are amphipathic amino acids rich motifs with the DNA binding surface rich in aromatic amino acids. Transcription factors can homo- or hetero-dimerize and recruit the basal transcriptional machinery to initiate transcription; hetero-dimerization provides an additional layer of diversity in transcription regulation. While the majority of transcription factors activate transcription, some function as repressor protein by binding to DNA sequences that overlap activator-binding sites or by binding to sequences that overlap a transcription start site.

Transcription factors have a 'Transactivation domain' (TAD) that links DNA recognition with recruitment of the transcriptional machinery. TADs are enriched with conserved acidic and hydrophobic amino-acids that are critical for transactivation of transcription (50). Some transcription factors have a nine-amino-acid transactivation domain, 9aa TAD (51). TAD is a common motif to a large number of yeast (Gcn5, Gal4, Oaf1, and Pho4) and animal transcription factors including VP16, p53, NF-IL6, NFAT1, and NF-κB. The TAD domain of these transcription factors interacts directly with the general transcriptional cofactor TAF9 (TAFII31), a component of TFIID and the basal transcriptional machinery. The third domain of most transcription factors is a ligand-

binding domain or a sensory domain, which initiates transactivation of the TF, as in the case of Estrogen Receptor (ER) (52).

Regulation of Transcription Factors

Translation and nuclear localization

Activation of TFs is regulated at different levels in the cell to assure timely and tissue specific gene transcription. TFs, just like all functional genes, are transcribed in the nucleus and exported to the cytoplasm for translation into a functional protein. TFs are active in the nucleus and possess a nuclear localization signal (NLS), a sequence of positively charged amino acids exposed on the surface of the protein (53). NLS is then recognized by the Karyopherin family of proteins (include Importin and Transportin proteins) that transport the transcription factor through the nuclear pore into the nucleus in an ATP-dependent process (54). However, regulatory proteins can bind to transcription factors and sequester them in the cytoplasm by masking the NLS binding site, thus rendering the protein unable to translocate to the nucleus to activate transcription. Inhibitor of κB (IκB) binds and sequesters the transcription factor NF-κB in the cytoplasm, thus regulating it transcriptional activation function in a non-specific manner (55). Moreover, IκB achieves NF-κB sequestering in the cytoplasm by masking its NLS signal and preventing its translocation to the nucleus (56). Other TFs are sequestered by interactions with specific regulatory proteins. MDM2 is an E3 ubiquitin ligase protein and a negative regulator of the tumor suppressor TF p53; MDM2 binds to p53 in the cytoplasm and induces proteasome-mediated degradation of p53 (57). Beta-catenin, a component of the Wnt signaling pathway, is another transcription factor that

is sequestered in the cytoplasm by binding to the transmembrane protein E-cadherin, or binding to APC/Axin/GSK-3β complex to mediate protesaomal degradation (58).

Activation signals

Activation of TFs is a complex process that depends on the inherent domains of TFs (Ligand binding domain) and other regulatory proteins. In the case of steroid receptors (e.g. Estrogen, Glucocorticoids), binding of the steroid leads to dissociation of the cytoplasmic co-repressors (Src for ER, HSP70/90 for GR) and translocation to the nucleus, where the receptor dimerizes and binds to its response element and recruits additional co-activator proteins (59, 60). However, a more complex chain of regulatory proteins exists for the tumor suppressor protein p53. DNA damage triggers multiple kinases to phosphorylate p53 at Ser15 and Ser37 by ATM/ATR and DNA-PK (61). This phosphorylation weakens the affinity of MDM2 interactions with p53, thus promoting both the accumulation and activation of p53. Additionally, the checkpoint kinases Chk1 and Chk2 can phosphorylate p53 at Ser20, enhancing its tetramerization, stability and activity (62). This tetramerization of p53 occurs in the nucleus and masks the nuclear export signal (NES) to maintain p53 transactivation (63). Acetylation of p53 is a second modification mediated by p300 and CBP acetyltransferases, which enhances p53 DNA binding domain affinity to bind its response element (64). Another level of p53 regulation occurs through ARF, which binds to MDM2 and promotes its rapid degradation, thus stabilizing p53 and maintaining its transcriptional activity (65). In many instances, the transcription factor is found inside the nucleus, but is bound to a negative regulator. The Retinoblastoma protein (pRB) is a tumor suppressor protein that negatively regulates cell-cycle progression by binding to the E2F transcription factor and masking its

transactivation domain, thus inhibiting its transcriptional activation of genes that encode DNA replication proteins (CDC6, TK) (66). pRB is typically hypo-phosphorylated and bound to E2F; however, its interaction with E2F is reduced upon phosphorylation by checkpoint proteins Cyclin D/CDK4/6 and Cyclin E/ CDK2 to promote DNA replication.

Transcriptional regulatory complexes

The TAD domain of TFs allows for competitive interactions with proteins or protein complexes that can either activate or repress transcription. These protein complexes recruit chromatin remodeling proteins that modify the DNA to allow for DNA-binding domains of TFs to recognize their *cis*-regulatory sequences and promote gene-specific transcription. TCF is a transcription factor that recognizes its cognate Lef-1 site on target genes; however, TCF is typically bound to Groucho, a protein that interacts with HDAC1 to repress transcription (67). Once the Wnt signaling pathway component β-catenin translocates to the nucleus, it displaces Groucho from TCF and interacts directly with TCF, allowing for gene transcription. The oncoprotein c-Myc is a transcription factor that binds E-boxes of DNA to activate transcription of specific target genes (68). C-Myc heterodimerizes with MAX, another transcription factor to enhance the specificity of DNA interaction of c-Myc. Upon binding to the E-Box, c-Myc interacts with histone acetyltransferase (HAT) enzymes to modify the chromatin and enhance transcription. However, Mad family of proteins compete for binding with Max and C-myc and can thus form heterodimers that can recognize the E-Box sequences (69). Mad-containing dimers antagonize c-Myc dimers by interacting with the mSin3a protein, which in turn recruits histone deacetylase (HDAC) enzymes to compact chromatin and reduce transcription.

Transcription Factors and Signaling Pathways

Many transcription factors belong to conserved families of proteins that function in a signaling cascade to regulate gene expression. These signaling pathways are involved in multiple cellular and developmental processes and mutations to any component of these conserved pathways lead to an array of developmental disorders and cancers. Signaling pathways are used in multicellular organisms to communicate between cells and utilize ligands, growth factors, and receptors to initiate these changes. The mitogen-activated protein kinase/Erk (MAPK/Erk) signaling pathway is involved in cellular growth and differentiation, and is activated by a wide variety of receptors including the epidermal-growth factor receptor (EGFR) (70). EGFR is a cell-surface receptor that belongs to the Erb family of tyrosine kinase receptors; it's activated by binding of the EGF ligand and the transforming growth factor α (TGFα) as they both function in cell proliferation and differentiation (71). Activation of EGFR leads to sequential activation of an adaptor (GRB2), linking the receptor to a guanine nucleotide exchange factor (SOS) and transducing the signal to small GTP binding proteins (Ras), which in turn activates the core unit of the cascade composed of a MAPKKK (Raf), a MAPKK (MEK1/2), and MAPK (Erk). These proteins can then signal terminal transcription factors like Activating Protein 1 complex (c-Fos and c-Jun), c-Myc, and cAMP reponse element binding (CREB) to activate gene transcription (72).

The NF-κB transcription factor is an important regulator of cellular response to multiple stimuli, but most importantly those of cytokines and viral antigens, to mount an immune response and protect the organism (73). Amongst the transcriptional targets of NF-κB is the promoter of the kappa light chain, an essential component of antibodies

responsible for creating the diversity in the immune response. NF-κB activation is receptor mediated and involves different families including Receptor activator of nuclear factor kappa B (RNAK), a member of TNFR family, Toll-Like Receptors (TLR), and Interleukin Receptors (74). Ligand binding to these receptors activates one of two kinases, MEKK3 (MAPK/Erk Kinase Kinase 3) and NIK (NF-κB Inducing Kinase). Both kinases lead to activation of IKK to phosphorylate IκBα (Inhibitor of κB) and its degradation, thus exposing the NLS signal of NF-κB to translocate to the nucleus (75). Interleukin receptors can mediate another signaling pathway, the JAK/STAT pathway. Binding of Interleukins to their receptors leads to auto-phosphorylation of the JAK family adapter protein, which then proceeds to phosphorylate its target STAT family of proteins (76). This leads to dimerization of STAT proteins and translocation to the nucleus to activate transcription.

The Notch and Wnt signaling pathways are essential to eukaryotic embryonic development and tissue specificity. Notch is a transmembrane protein with an extra-cellular and intra-cellular domain. There are four different Notch family receptors (Notch1-4) that can interact on the cell surface with five ligands (belong to Jagged and Delta families of receptors) (77). The Notch pathway regulates cell fate determination of neighboring cells through lateral inhibition, depending on their ability to express either the receptors or the ligands. Upon ligand binding to the extracellular domain of Notch receptors, two different proteases get activated: the ADAM proteinase family to cleave the extracellular domain, and the γ-secretase complex to cleave the intracellular domain. The intracellular domain then translocate to the nucleus where it interacts with multiple proteins to activate transcription (78). Similarly to the Notch pathway, the Wnt

signaling pathway is activated by binding of the Wnt ligand to the Frizzled receptor, which forms a complex with LRP and Dishevelled. Binding of the Wnt ligand activates the protein Dishevelled to bind to the “destruction complex” composed of Axin, APC, GSK-3β and Beta-catenin. Dishevelled then phosphorylates Axin thus weakening the affinity of the destruction complex and GSK-3β ability to phosphorylate β-catenin and target it for destruction (67). B-catenin then enters the nucleus and activate transcription by binding to its Tcf/Lef1 sites.

Superclass Helix-turn-Helix

The Superclass Helix-Turn-Helix transcription factors share a unique DNA recognition domain. Family members of this Superclass include the Homeobox domain family (including Paired box), the Fork head / Winged helix family, the Heat Shock Factors, and the transcriptional enhancer factor (TEA) family (49). Perhaps the Homeotic box (HOX) family and the Forkhead Box (FOX)/Winged helix proteins family are the best known members of this superclass. The FOX family will be discussed in great details later in this chapter. Hox genes encode transcription factors defined by the DNA-binding domain termed the homeodomain that function independent of conserved signaling pathways (79). Hox genes are present in all animals and the homologues are highly conserved. Hox genes in vertebrates are located contiguously in clusters with the number of these clusters varying according to anatomic complexity. There are 4 Hox clusters in mammals (HOXA, B, C, and D) and 39 Hox genes have been identified in humans. Hox genes have been well studied in *Drosophila*, where they have been shown to be involved in normal temporo-spatial limb and organ development along the anterior-posterior (A-P) axis. There are three basic mechanisms utilized by Hox cluster

to regulate normal tissue development in vertebrates: spatial collinearity, posterior prevalence, and temporal collinearity (80). Spatial collinearity correlates the position of a Hox gene in the cluster with its tissue expression along the A-P axis. Normally, 3′ genes within the cluster are expressed in anterior tissues and 5′ genes in posterior tissues. Posterior prevalence postulates that Hox genes which are positioned more 5′ in the cluster will have a dominant phenotype to those more 3′. This is due to transcription of microRNA embedded along with HOX gene in the cluster; miRNA can then silence more upstream HOX genes. Temporal collinearity is similar to spatial collinearity but correlates temporal tissue order of expression of Hox gene with the cluster.

HOX proteins have been shown to effect essential cell processes like cell migration, cell cycle control, apoptosis, cell adhesion, and development. HOX proteins function at the 'high level executive' stage during development, which denotes targeting of 'executive genes' that encode transcription factors, cell-cell signaling, or morphogens (80). *Decapentaplegic* is an executive gene regulated by the *Drosophila* HOX gene *Ultrabithorax* and *Abdominal-A*, and it expression triggers cell shape changes in the gut that are required for normal visceral morphology. These two HOX proteins also regulate *distal-less* expression, a gene involved in limb formation. Meanwhile, Deformed is another HOX gene that regulates reaper expression, which in turn activates apoptosis during maintenance of the boundary of maxillary and mandibular segments of the head. HOX proteins are less well studied in higher mammalian development, but studies have shown a role in cell-cycle regulation and cancer. Hoxa10 appears to be activated during differentiation of mouse cultured myelomonocytic cells into monocytes followed by growth arrest; this was attributed to the induction of cell-cycle arrest protein p21 (81).

The Forkhead Box

The forkhead family is a large family of transcription regulators that share a structurally related DNA binding domain: the forkhead (82, 83). Originally, the fork head gene (fkh) was discovered as a mutation in Drosophila *melanogaster* which presented with an ectopic/displaced head in larvae and a spiked head phenotype in adult flies. Expression of the fkh gene was determined to be in the terminal domains of the D. *melanogaster* embryo which correlate to the anterior head and posterior tail of an adult fly (84). Protein sequence analysis indicated the lack of a homeo domain in the fkh protein, and the sequence showed no similarity to any known protein family. The closest discovered homologous protein factor that had been identified to play a role in head/tail development was the *spalt* (sal) gene of the 'region-specific homeotic genes' family (84). However, the fkh protein was not structurally related to the sal protein. Meanwhile, Lai and his group were able to clone the cDNA encoding Hepatocyte Nuclear Factor 3α (HNF3α) in rat liver (85). Sequence analysis of HNF3α revealed no similarity to the homeo domain protein HNF-1 or the Leucine-zipper protein C/EBP, which play a role as liver nuclear transcription factors. Lai and colleagues limited the DNA-binding region to amino acids 124-288 of the protein and predicted short α-helices based on the abundance of proline and glycine residues (85). Detlef Weigle and Herbert Jackle compared the Drosophila fkh protein with the mammalian HNF3α protein. In comparing the DNA-binding domains of both proteins, they noticed 86% identity and 92 % conservation in a 110 amino acids span of Drosophila and mammalian Fox proteins which contained the DNA binding motif (86). HNF3α was later renamed FOXA1 and the other HNF3β and HNF3γ proteins identified in the study were assigned to the FOXA

subfamily as FOXA2 and FOXA3 respectively. Subsequent crystallography of the FOXA1 protein revealed the structure of Forkhead proteins and its interaction domains.

FOX crystal structure and phylogeny

Clark et al first resolved the structure of HNF3γ (FOXA3) interacting with a 13 nucleotide region of the transthyretin (TTR) promoter previously used to identify HNF3 family of proteins. Crystal structure of the forkhead domain of FOXA3 shows the presence of three α-helices (H1, H2, and H3) at the N-terminus of the protein. There are also three β-strands (S1, S2, and S3) located towards the C-terminus of the FOXA3 protein (87). A polypeptide sequence connects S2 and S3 to form the first loop (W1), and another polypeptide sequence is found after S3 to form the second loop (W2). S1 is found between helices 1 & 2 and interacts with S2 and S3 to form an anti-parallel B-sheet. Different from homeo domains, The N-terminal region of Forkhead proteins interacts with C-terminal region via S1 to form a hydrophobic core. A closer look at the interaction of a FOXA3 monomer with TTR promoter shows interactions of H1 and H2's N-termini with the phosphate backbone of DNA likely in the minor grooves. H3 interacts with the major groove of DNA, and both wings border this interaction. The W1 and W2 loops flanking the H3 interacting with the major grove yields a butterfly shaped interaction, which has given the term "winged helix" to Forkhead proteins. Side-chains of the helices and wings interact with DNA phosphate and ribose groups; this includes multiple W2 arginine residues found to interact with the minor and major grooves of DNA at specific nucleotide positions. Other residues were shown to interact specifically with sense or anti-sense nucleotides of the DNA fragment (87).

After the discovery and classification of the HNF3 proteins as the first members of forkhead/winged helix proteins, many scientists began to discover other proteins containing the same domain among various chordates (88). FOX (Forkhead Box) was adapted as the nomenclature for such proteins, and phylogenetic analysis further classified FOX proteins into 15 subclasses (FoxA through FoxS) depending on conservation of the forkhead domain. Arabic numerals designate members of a subclass of FOX proteins which depends on protein sequence (88). Homeobox genes (Hox) are a class of important transcription factors involved in development. Functionally related Hox, Parahox, and Nkx genes are clustered together in chromosomal regions (89, 90). However, no such clustering occurs for the Fox family of genes except for a few gene clusters. A cluster of FoxL1, FoxC, FoxF and FoxQ1 was identified in the chordate amphioxus. In humans, two different clusters were identified on chromosomes 16q24.3 and 6p25. On chromosome 16q24.3, the *FOXL1*, *FOXC1* and *FOXF2* genes were found within a 70 kb span. At chromosome 6p25, the *FOXC2*, *FOXF1* and *FOXQ1* genes were found within a 325 kb span (90).

FOX proteins and development

FOX proteins differ from related Superclass helix-turn-helix HOX genes in their function as terminal transcriptional effectors rather than high-level executive proteins. FOX Proteins can function individually and redundantly as monomers or in an overlapping manner. FoxA1 and FoxA2 cooperate during endoderm differentiation into lung epithelium and during liver morphogenesis and hepatic specification (83, 91, 92). Moreover, FOXF1 and FOXF2 exhibit non-allelic and non-complementation functional overlap as homozygous Foxf2 mutants and Foxf1/Foxf2 heterozygotes displayed gut

and intestinal development problems including megacolon, colorectal muscle hypoplasia and agangliosis (91). Mutations in FOXO3a and FOXL2 result in premature ovarian failure (93). Foxc1 and Foxc2 have interactive and essential roles in cardiac morphogenesis because compound heterozygotes exhibit similar cardiac or renal phenotypes to single null mutants (94). Foxd1 and Foxd2 also exhibit an additive and synergistic role in renal development, and Foxc1 and Foxc2 also appear to affect renal development especially migration of renal cells (95). Foxc1, Foxc2, and Foxl2 are expressed in the neural-tube derived ocular tissue and mutations in these genes lead to abnormalities in ocular development. Many FOX genes (Foxd3, Foxc1, Foxc2, Foxl2) are expressed in the neural crest and mutations to these genes result in cell-migration defects and organ defectsFoxb1b deletion results in neural tube closure defects, reduced posterior body axis and embryonic lethality, while Foxh1 mutations affects left-right axis specifications (95). FOXE1 is involved in cranio-pharyngeal development and mutations lead to thyroid agenesis and other phenotypes. Foxi1 is responsible for inner-ear development, especially the vestibulum and cochlea (96). Foxn1 and Foxp3 mutations exhibit severe immune-system defects, as Foxn1 is highly expressed in skin and thymus tissue and maintains balance between proliferation and differentiation (95).

FOX proteins and transcription

Forkhead box proteins can bind DNA as monomers and independently regulate transcription by working as activators, repressors, or can have a dual function. FOXD2 over expression was shown to bind the RIα1b promoter and increase its sensitivity to cAMP, thus inducing RIα mRNA and protein level, which is involved in T-cell activation (97). Additionally, the FoxD2 avian homologue (CWH-2) was shown to act as

transcriptional repressor by binding to Qin/BF-1 promoter sites of genes involved in brain development and oncogenesis (98). In some cases, splice variants of the same gene can function as activators or repressors (99). The Foxm1 (Human homologue TRIDENT/FKHL16) protein is a winged helix protein specifically expressed in rapidly dividing cells and is involved in controlling mitosis and promoting S phase and G2/M phase progression (99, 100). The 396BP region upstream of the Foxm1 gene start codon is responsible for modulating and transducing growth factor stimuli to effect expression of Foxm1 (100). FOXM1a is a splice variant of the FOXM1 gene which binds the FOXM1 consensus sequence but does not trans-activate genes. However, FOXM1b and FOXM1c are involved in the activation of many genes (c-myc, c-fos, Cyclin-B, Cdc25B) involved in cell-cycle regulation, genomic stability, and tumorogenesis (99).

Additional diversity in FOX proteins regulation of transcription can be achieved by interacting with signal transduction protein complexes. The forkhead activin signal transducer-1 (FAST-1) or FOXH1 is a forkhead protein that interacts with human Smad2 and Smad4 to activate the Activin Response Element (ARE) (101). The activation of the ARE by the complex is dependent on the presence of a TGF-β stimulating peptide, which suggests the FOXH1-Smad complex functions as a signal transduction cascade in the TGF-β pathway. Seoane, et al presented more evidence supporting a role for FOX proteins in signal transduction when they showed a role for FOXO family of proteins in regulating $p21^{Cip1}$ through a TGF-β signaling pathway. FOXO proteins were shown to form a complex with Smad2, 3, and 4 and positively activate $p21^{Cip1}$ and relief its c-myc repression (102). Moreover, FOXG1 was shown to be a negative regulator of the Smad-FOXO induction of $p21^{Cip1}$ in neuroepithelial cells by interacting with Smad-

FOXO complex. FOXG1's role in progenitor neuro-epithelial cell development was previously studied, and it was shown to be a transcriptional repressor to differentiation signal stimulation (103). Foxg1 plays a similar role in the regulation of Foxh1 during lung development. Analysis of glioblastoma multiforme tumor samples and cell lines showed an increased expression of FOXG1, which suggests a role for the silencing of TGF-β mediated $p21^{Cip1}$ induction (102). Other FOX transcription factors function as terminal effectors for other conserved signaling pathways including Hedgehog, Insulin/IGF, Wnt signaling pathway, and MAPK/Erk (93).

FOXN Subfamily

The N subfamily of FOX proteins was first discovered when Li and colleagues identified the human T-cell leukemia virus enhancing factor (HTLF) as protein that bound the human T-cell leukemia virus long terminal repeat (HTLV-1 LTR) at the -117 to -155 region (104). The DNA binding domain of the HTLF resembled the forkhead protein which was newly characterized, and the HTLF was later identified as the second member of the N-Subfamily (FOXN2). The FOXN subfamily received more significance when Nehls and colleagues were able to clone the nude locus of mouse chromosome 11 (105). Characterization of the protein encoded by this region yielded a 648 amino acid protein with a DNA binding domain that resembled the forkhead domain (105). Interestingly, the newly discovered protein showed a 56% identity to the human T-cell leukemia virus enhancing factor (HTLF), which was earlier identified by Li and colleagues. The protein, later assigned FOXN1, was found to be mutated in nude mice which have rudimentary thymus and lack a T-cell induced immune defense (106). Down-regulation of hair keratins, abnormal morphogenesis of hair follicles, and failure to

develop a thymus from the ectoderm are characteristics of nude mice (106). FOXN1 expression was later shown to be regulated by Wnt4 of the Wnt pathway which utilizes the PI3K-Akt to phosphorylate FOXN1 in thymic epithelial cells (107). Pati and colleagues isolated a high-copy suppressor of the S. cerevisiae MEC1 G2-M checkpoint mutants that conferred increased survival of G2-M checkpoint mutants (108). The protein encoded by the cDNA was called Checkpoint suppressor 1, and characterization of the full length protein showed homology to the forkhead/winged helix DNA binding proteins (108). Moreover, CHES1 sequence showed high homology to the human T-cell leukemia virus enhancing factor (HTLF) with 51% identity and 69% conserved residues. Identification of members of the N subfamily continued when Gouge and colleagues used the forkhead domain of Foxg1 to screen a mouse' eye cDNA for new proteins utilizing low-stringency hybridization (109). They identified FoxN4, a 523 amino acid polypeptide where sequence analysis predicted an activation domain and a DNA binding domain. The later domain showed high similarity to HTLF/FOXN1, and in-situ hybridization analysis showed specific expression in neural tissue and exclusively in the retina after embryonic day E13.5 (109). Most recently, Katoh was able to clone and identify FOXN5 and FOXN6 who share a 69 amino acid FN56 domain in addition to the forkhead domain (110, 111). In a study on FoxN expression in Xenopous *laevis*, FoxN5 mRNA was shown to be maternally expressed, and FoxN5 transcripts are degraded during gastrulation and completely eliminated by neurlation (stage 15) (112). Interestingly, FOXN6 mRNA was shown to be expressed in breast cancer cell lines and primary breast cancers, and FOXN5 was associated with the human B-cell CLL/lymphoma 9-like (BCL9L) gene. A nuclear complex, which consists of BCL9L,

human T-cell leukemia virus enhancing factor, and beta-catenin, is shown to activate Wnt signaling pathway dependent expression of many genes including the oncogene Myc (113).

FOXN3 (Checkpoint Suppressor 1)

FOXN3 (also known as Checkpoint suppressor 1, FoxN3) belongs to the N subfamily of forkhead box transcription factors. Initial discovery and characterization of the protein was done by Pati and colleagues when they screened checkpoint-mutant *Saccharomyces cerevisiae* transfected with human cDNA for rescue of phenotype (108). *Tx23*, a human cDNA isolated from U118 glioblastoma cell line, was isolated as a high-copy suppressor of the S. cerevisiae MEC1 G2-M checkpoint mutants. The protein expressed by the tx23 cDNA, Checkpoint Suppressor 1 (CHES1), conferred increased survival of mec1-1 mutants to methyl methanesulfonate (MMS), UV radiation, and ionizing radiation which are DNA damage agents. Meanwhile, full length cDNA of the CHES1 protein restored G2-M checkpoint function in rad9 and rad24 mutants to wildtype levels. Analysis of the CHES1 protein encoded by the full length cDNA showed a novel protein that shares a common region with the forkhead/winged-helix family. The protein showed greatest homology to the human T-cell leukemia virus enhancing factor (HTLF). Interestingly, the 200 amino acids at the carboxy terminus of the CHES1 protein isolated were sufficient to restore the mec1-1 G2 checkpoint mutants in S. cerevisiae (108). However, the carboxy terminal did not encode for the forkhead DNA binding region.

The C-terminal portion of the human FOXN3 protein was found to interact with Sin3, a component of the Sin3/Rpd3 histone deacetylase complex (HDAC) in budding yeast (114). The Sin3/Rpd3 HDAC complex is targeted to specific promoter regions via Sin3 interactions with site-specific DNA-binding proteins (115). CHES1/FOXN3 appears to inhibit the activity of Sin3/Rpd3 as deletion of Sin3 in Rad9 deficient S. cerevisiae conferred resistance to UV damage similar to the results seen upon CHES1 introduction to Rad9 deficient S. cerevisiae (114). Other studies showed the C-terminus of FOXN3 represses transcription when targeted to a reporter promoter in cell lines derived from tumor tissues, which is consistent with its interaction with the Sin3/Rpd3-HDAC complex (116). Moreover, CHES1 interacts with the C-terminus region of the Ski-interacting protein, SKIP; this interaction varies from the well conserved SNW protein-binding domain of SKIP (116). SKIP is a well-conserved nuclear regulatory protein involved in pre-initiation of transcription, splicing and polyadenylation of RNA (116, 117). In addition to the Ski protein, SNW/SKIP interacts with many proteins including the tumor suppressor protein Retinoblastoma and receptor-regulated Smad2 & Smad3 proteins, which are TGF-β modulators (117). SKIP functions to recruit either activation or repression complexes to mediate multiple signaling pathways that are involved in cell proliferation and differentiation (118). CHES1's transcriptional repression of the GAL4 reporter promoter can be correlated to CHES1's interaction with SKIP, which has been shown to recruit many repression complexes, including mSin3a, HDAC1, HDAC2, and NcoR/SMRT (116, 119). Case reports on human patients with deletions of the chromosomal region containing *FOXN3* have shown phenotypes that include, dysmorphic features, delayed development, absent speech, auditory neuropathy,

craniofacial defects, dental anomalies, microcephaly (abnormally small head), skeletal defects, and hypotonia (120, 121).

The human *FOXN3/CHES1* has been mapped to 14q32.11 while the mouse Foxn3 was mapped to chromosome 12, region E (108). The human FOXN3 gene encodes a 490 amino acids protein with a predicted molecular weight of 54 kDa, while the mouse Foxn3 encodes a 457 amino acids protein with a predicted molecular weight of 50 kDa. Human FOXN3 has a splicing variant which is encoded by exon VI and adds an additional 66 bp (22 amino acids) to the gene transcript. Also, There is a variable leader exon which is used equally in both splice variants (112). In-situ hybridization analysis using FoxN3 probes in *Xenopus laevis* shows a persisitent expression in the eye, the branchial arches and the vagal ganglion.

The role of FoxN3 in mammalian development is not known due to the lack of data relating to animal models that are defective in FoxN3 expression. However, recent studies have shown the involvement of FoxN3 in craniofacial and eye development in *Xenopus laevis* (122). *X. laevis* larvae injected with FoxN3 antisense morpho-olino (MO) were shown to have reduced eye size, reduction and false positioning of craniofacial cartilages, increased apoptosis in the brain, and reduced food intake. Cranial nerves, including hypoglossal and branches of the trigeminal nerve, were either lost or severely deformed especially nerve branches that innervate the lower jaw. The severity of the phenotype progressed with the developmental stage of the *X. laevis* larvae beginning at stage 41 onward and ending at stage 50 when reduced food intake ultimately lead to death. The severity of phenotype was also antisense MO dose dependent with greater and a more severe phenotype being produced with complete reduction of FoxN3 (122).

This may suggest that nullyzygoisty for FoxN3 in higher mammals would be lethal while heterozygosity may demonstrate phenotype.

Foxn3 mouse model

In an effort to understand the role FoxN3 in mammalian development, we have developed a mutant mouse model for the Foxn3 gene. Foxn3 deficient mice were generated using the Baygenomics gene trap embryonic stem cell resource. The Baygenomics insertional mutagenesis strategy involves the use of a gene-trap cassette consisting of a splice-acceptor-βgeo cassette (β-galactosidase-neomycin fusion gene) and characterized using 5' RACE (Figure 1). We obtained one of the ES cell clones that had been characterized to have a gene trap insertion within the *Foxn3* gene for analysis of the gene trap insertion site. Using multiple intronic forward primers and a gene-trap specific reverse primer for PCR amplifications and sequencing, we determined the gene-trap insertion site to be in intron 3 (Figure 2). Foxn3-targeted ES cells were used for blastocyst injections using the microinjection services at the University of Massachusetts Medical School, Worcester. Germ line founder mice were generated and the colonies were expanded for the analysis of the mutant offspring. The working model for the Baygenomics gene traps encompasses the splicing of the 5' splice site of Foxn3's exon 3 (encoding the first 218 amino acids) into the 3' splice site in the gene trap (123). This allows for the in-frame fusion of the first three exons of *Foxn3* to the B-galacotsidase neomycin sequence and the expression of this fusion protein. The expression of this fusion protein is mandated for the blastocysts to survive neomycin selection. Upon efficient splicing of the Foxn3 gene into the gene trap, the remainder

exons will not be translated and thus the wild type protein will be haplo-insufficient in heterozygous progeny and completely lost in nullisomic progeny.

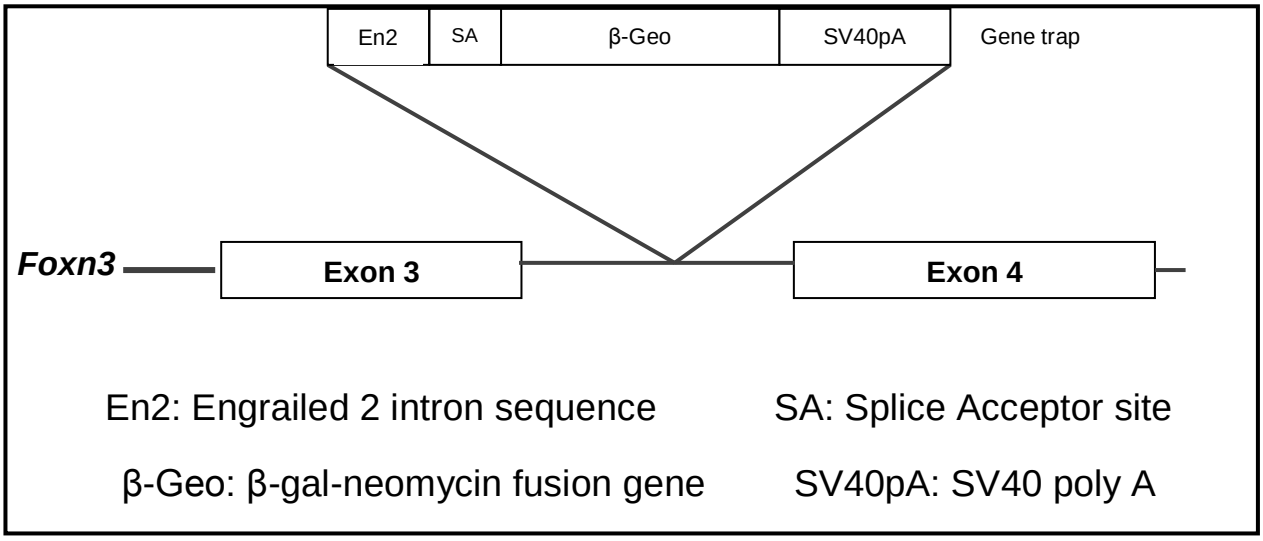

Figure 1. Schematic representation of the Baygenomics gene trap integrated into intron 3 of Foxn3 gene. **The gene trap cassette contains an Engrailed 2 intronic sequence followed by a splice acceptor site, an ATG-less B-galactosidase neomycin fusion gene and a ploy-A site.**

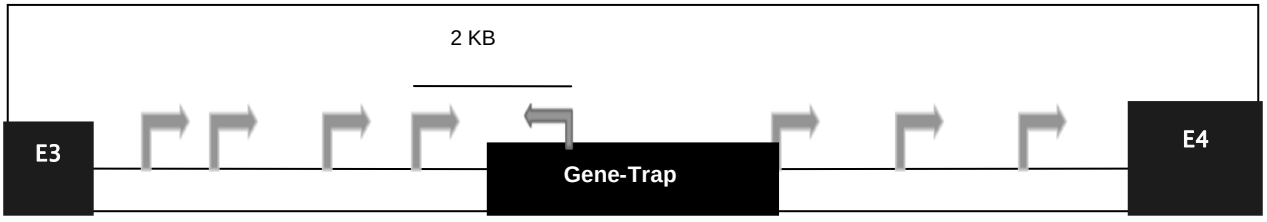

Figure 2. Schematic representation of the strategy utilized to configure the gene-trap insertion site in intron 3 of the Foxn3 gene. **Foxn3 Intron 3 forward primers were designed at 2kB space intervals and tested against one gene-trap specific reverse primer in WT and Heterozygous mice. Sequencing of the PCR product indicated that the gene trap was integrated within intron 3 (15,450 bases from the beginning of the intron) of the *Foxn3* gene**

Materials and Methods

Polymerase Chain Reaction (PCR)

A nested PCR strategy was used to genotype the tail DNA of mice. Initial genotyping of tail DNA was done by southern blot analysis, which is time consuming and cumbersome. Gene-trap specific primers spanning 350nt (gene-trap relative positions: 1665 forward primer and 2000 reverse primer) were initially used to distinguish between wildtype and mutant progeny. Intronic primer 3.33X was used as a forward primer and 984r was used as a gene-trap reverse primer to authenticate the insert site in the mutant progeny. To distinguish heterozygous from nullisomic progeny, Intronic forward primer 3.33x and intronic reverse primer 400 reverse (position relative to 400nt after gene trap end site) were used. The PCR reaction mixture consisted of 10X PCR buffer (50 mM KCl; 10 mM Tris-HCl final concentration), $MgCl_2$ (3mM final concentration), dNTP (200μM final concentration of each), forward and reverse primers (0.4 μM final concentration), Platinum Taq (0.5U final concentration), and ddH_20 to a final reaction volume of 25 μl. The PCR conditions were as follows: an initial denaturation step of 5 minutes at 94°C, followed by another denaturation step of 30 seconds, annealing for 30 seconds at 58°C, and extension for 90 seconds at 72°C. This was repeated for a total of 32 cycles. A final cycle of 72°C for 5 minutes was also included. The PCR products were run on a 1.5% agarose gel at 95 V for approximately 1 hour. Ethidium bromide was used to visualize the bands, and the gel was photographed using the Epichem^3 Darkroom and analyzed using LabWorks GelPro Application. The presence of a 350BP band with the gene-trap specific primers or the presence of a 1.5KB band in the insert primers indicated at least one gene trapped allele and thus heterozygousity. The lack of

a band in the two intronic primers indicated a nullisomic progeny, while the presence of a 1.2KB band indicated a heterozygous progeny.

PRIMERS

Primers used genotype analysis
GS1 (Forward): TGATGCTCTCTCTACCAGTGGGAT
GS2 (Reverse): ATACACAGCACCCACTCCTATCCA
TR1 (Reverse): CCCAACTGACCTTGGGCAAGAACATA

Primers used for RT PCR
Foxn3:
Forward: ATGCACCTACTGGGTGGAAGAACT
Reverse: GGCTCCGTTTCTCTTGAAGAAGGT
Foxn3-βgeo:
Forward: ATGCACCTACTGGGTGGAAGAACT
Reverse: TTTCCCAGTCACGACGTTGT
Bmp2:
Forward: TATGCTAGATCTGTACCGCAGGCA
Reverse: CAAGTTGGCTGCTGCAGGCTTTAT
Bmp4:
Forward: AGCCAACACTGTGAGGAGTTTCCA
Reverse: ACTGCAGGGCTCACATCGAAAGTT
Bmp7:
Forward: ACCGCAGCCGAATTCAGGATCTAT
Reverse: ACCATGAAGGGTTGCTTGTTCTGG
Runx2:
Forward: GTGGCCACTTACCACAGAGCTATT
Reverse: GATGAAATGCTTGGGAACTGCCTG

Rpl38 (*Ribosomal protein L38*):

Forward: TTCGGTTCTCATCGCTGTGAGTGT
Reverse: TCTTGACAGACTTGGCATCCTTCC

Micro Computed Tomography (mCT) Image analysis

High-resolution CT images were acquired using a MicroCAT™ II + SPECT instrument (Siemens Medical Solutions, Molecular Imaging, Knoxville, TN). Each image comprised 360 projections at 1 degree intervals and was acquired with X-ray source energy of 80 kVp. CT data were rendered using the Amira 3-D image analysis software package (Amira, Version 3.1: Mercury Computer Systems, Chelmsford, MA). Foxn3 deficient mice were isolated at 9.5, 11.5, 13.5, 18.5 day post coitum and 8,40 day old mice from various matings. Cranium was isolated and placed in Bouin's fixative or 10% formalin. Micro CT analysis were performed by Dr. Jon Wal at the University of Tennessee's Medical Center.

RNA isolation and RT-PCR protocol

Total RNA was extracted using Trizol according to the manufacturer's instructions (Invitrogen) from MEFS and various tissues from adult mice (3 months of age). For expression analysis of skulls, the brain tissue was removed from neonatal mice and the remainder of the tissues was homogenized. First strand cDNA synthesis was performed with 2µg of total RNA (pretreated with RNAse free DNAse) with random hexamers and M-MLV reverse transcriptase (Promega) for 1 hour at 42^{o}C followed by inactivation of the reverse transcriptase at 70^{0}C for 15 minutes. A similar amount (2µg) of total RNA (pre-treated with RNAse free DNAse) was subjected to the above mentioned conditions in the absence of M-MLV reverse transcriptase. For PCR assays,

2µl of the reaction mixture (from a total of 40 µl) obtained from the first strand cDNA synthesis reaction was used. The PCR conditions were 94^0C for 2 min, followed by 30 cycles at 94^0C for 30 sec, 58^0C for 30 sec and 72^0C for 10 sec. PCR products were resolved by agarose gel electrophoresis and stained with ethidium bromide.

Expression analysis of *Foxn3* during mouse development

Embryos obtained from timed matings from wild-type females and *Foxn3*$^{+/H}$ males were fixed with 2% paraformaldehyde and stained in a solution containing X-gal (2 mM $MgCl_2$, 0.01% sodium deoxycholate, 0.02% NP-40, 5 mM potassium ferricyanide, 5mM potassium ferrocyanide, 0.1 M phosphate buffer, pH 7.3). Genotypes of embryos were determined from genomic DNA isolated from yolk sacs.

Results

The results of this chapter were published in a manuscript titled “Foxn3 is essential for craniofacial development in mice and a putative candidate involved in human congenital craniofacial defects” (124).

Published results.

Genomic DNA isolated from ES cells were analyzed by PCR to confirm *Foxn3* disruption by using primers (spanning 2 kb intervals) that were specific for *Foxn3* intron 3 and the gene-trap sequences. After obtaining one specific band in heterozygous DNA, we confirmed the authenticity of the band by obtaining other forward primers in the region and running PCR analysis to confirm the band size. Similar analysis was done utilizing other reverse primers from the gene-trap (Figure 3). Sequencing of the PCR product indicated that the gene trap was integrated within intron 3 (15,450 bases from the beginning of the intron) of the *Foxn3* gene.

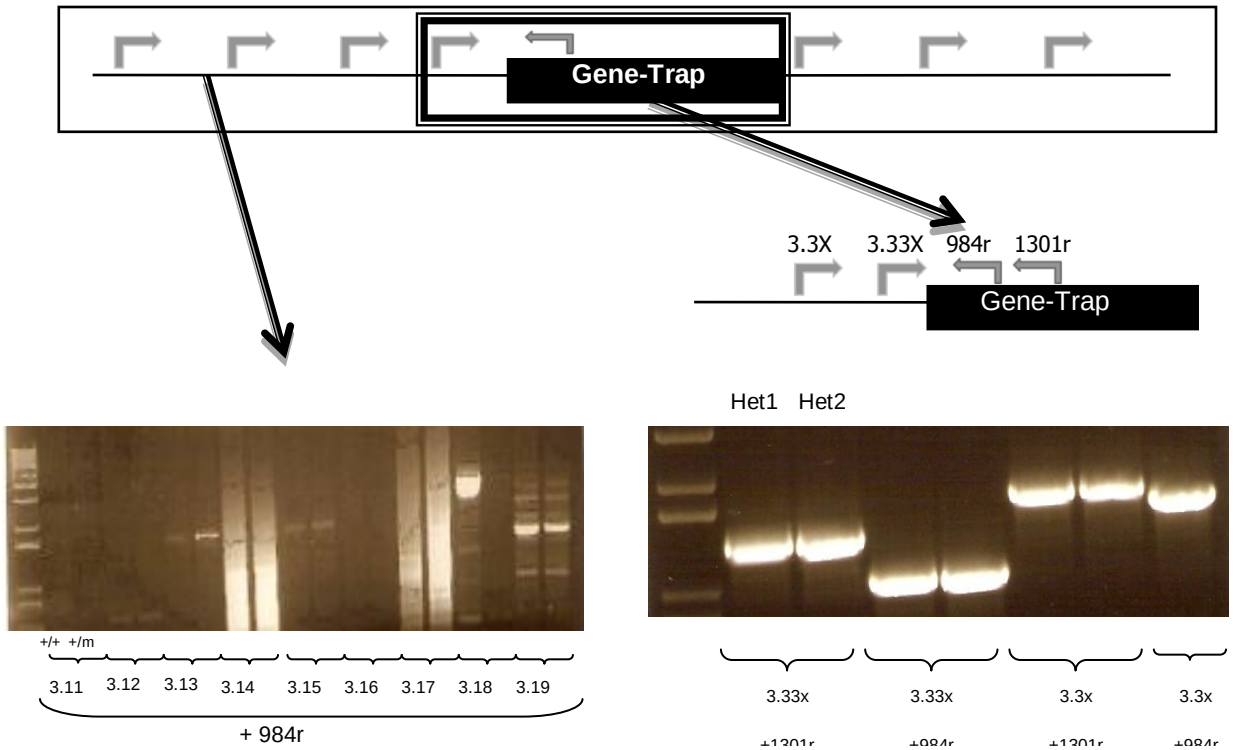

Figure 3. PCR analysis of multiple forward primers of intron 3 of the Foxn3 gene. **Left figure shows the banding pattern obtained from running one gene-trap specific reverse primer with multiple forward primers in WT and heterozygous mice. Right figure shows the isolated forward primers with variable gene-trap specific reverse primers.**

After confirming the insertion site of the gene trap, we designed an intronic reverse primer 400 BP after the insertion site. This reverse primer was designed to yield a 1.2KB band when run with a specific intronic forward primer. This product can only be made in the heterozygous mice because they still maintain a wildtype copy of the FoxN3 gene (Figure 4). The insertion of the gene trap extends this 1.2Kb product to approximately 10KB in the mutant chromosome, which is not amplified by a PCR analysis. Therefore, our nullizygous mice will not show a product for this specific PCR. This allows us to differentiate heterozygous versus nullizygous mice without utilizing a southern blot.

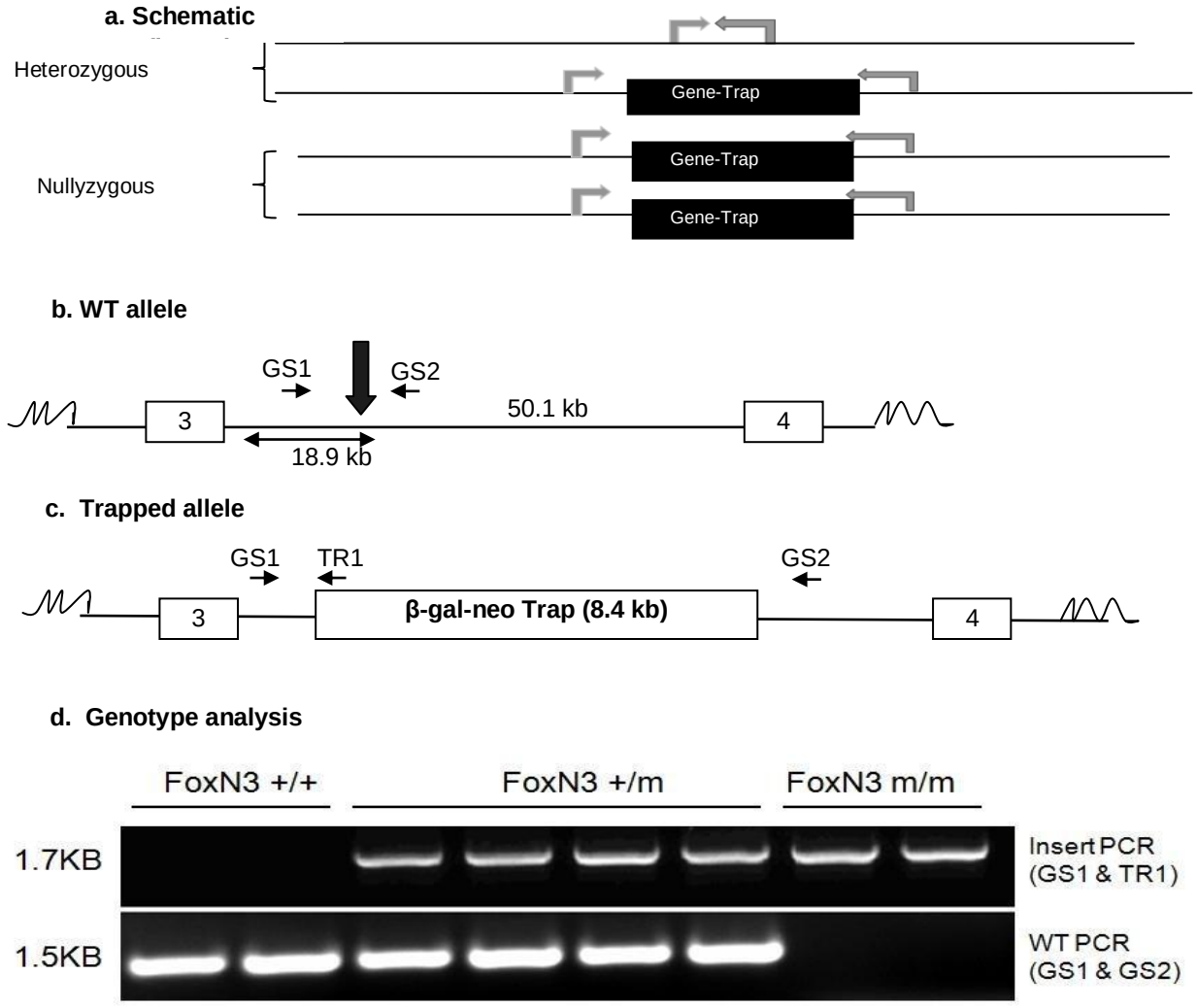

Figure 4. Schematic representation of the Foxn3 chromosome locus in wild type, heterozygous, and nullisomic mice. **A. schematic representation of forward and reverse primer pair (GS1 and GS2) used for genotyping of Foxn3 mutant mice. B. Wildtype allele of Foxn3 intron 3 showing relative positions of the primer pair and the gene trap. C. Mutant allele showing the Insert PCR primer pair (GS1 and TR1) used for identifying Foxn3 mutant mice. D. Agarose gel electrophorysis analysis of wildtype, heterozygous and nullisomic Foxn3 mice and the primer pair used for each genotype.**

At the protein level, the insertion of the gene trap was determined to be within the C-terminal portion of the Fork-head box domain and upstream of three C-terminal exons encoding the SIN3 binding domain. The truncated FoxN3 protein yields a FoxN3-βgeo fusion protein containing the first 218 amino acids of the 457 amino acids protein

(Figure 5). Once we were able to genotype the mouse colony and the offspring of matings, we analyzed the embryonic and post-natal lethality of FoxN3 mutant mice. Analysis of Heterozygous matings at neonatal day 1 shows 61% survival of the Nullisomic offspring which indicates some embryonic lethality (Table 1). Further analysis of these offspring during weaning (day 21) shows a 25% survival of the null offspring, which further confirms the lethality of the FoxN3 mutant mice.

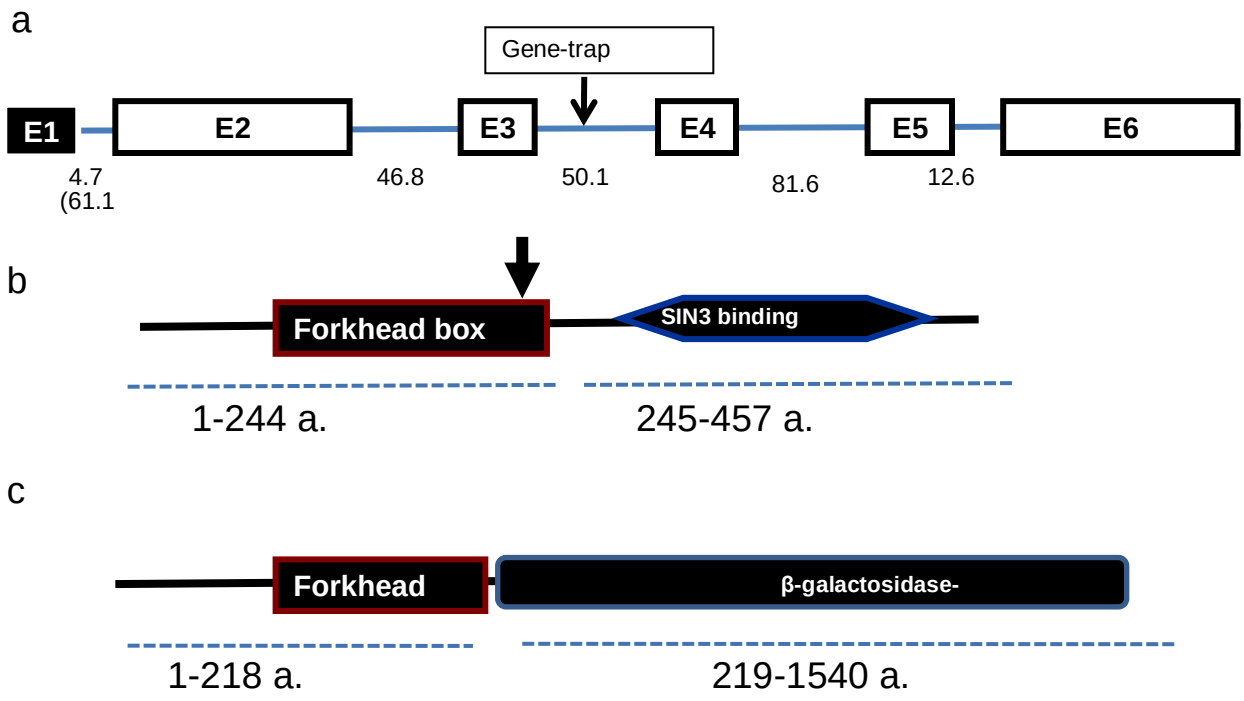

Figure 5. Schematic configuration of the FoxN3 gene. **A. Schematic representation of the FoxN3 gene showing its exons, introns, and gene-trap insertion site. B. NCBI analysis of the insertion site at the FoxN3 protein. Notice the loss of the C-terminal SIN3 binding domain as well as a portion of the Forkhead box. C.Schematic representation of the fusion protein containing the first 218 amino acids of FoxN3 and the C-terminal βgeo fusion protein.**

Table 1. Embryonic and postnatal lethality of FoxN3 mutant mice. **Neonatal pups and offspring at weaning were genotyped for FoxN3 disruption using PCR. The total number of intercrosses (n) analyzed for each group and the expected numbers of mice based on Mendelian ratio are indicated within parentheses.**

Developmental stage	Total	WT	+/m	m/m (expected)	% Survival of m/m
Neonatal day 1 (n=9)	71	21	38(42)	13(21)	61%
Weanlings at day 21 (n=7)	55	16	35(32)	4(16)	25%

We analyzed *Foxn3* expression patterns in embryos obtained from timed matings between wild type and *Foxn3* heterozygotes. *Foxn3* expression was limited to the craniofacial regions and the developing vertebral column during embryogenesis (Figure 6). Importantly, *Foxn3-β-gal* expression was specific only to the heterozygotes (as determined via the genotyping of the yolk sacs) and the wild type embryos did not show any staining for β-galactosidase. Analysis of *Foxn3* gene expression in tissues from adult mice indicated a differential expression pattern with the highest expression in the brain, eye, skeletal muscle, thymus and ovaries (Figure 6).

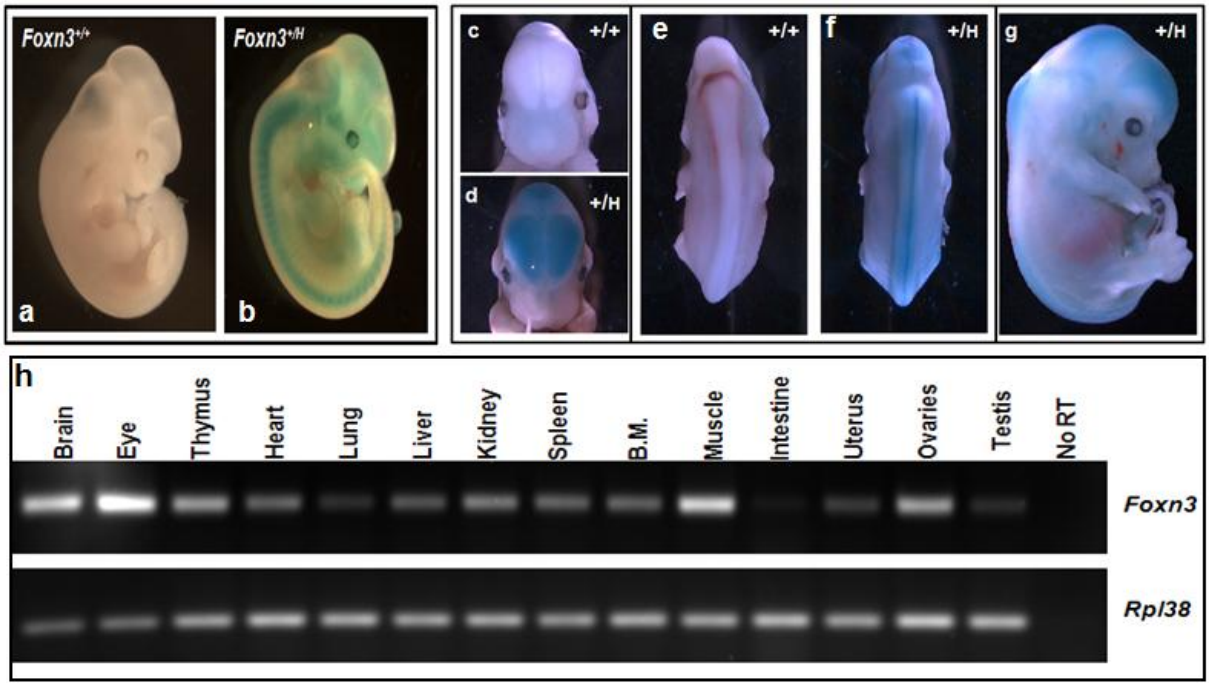

Figure 6. Expression analysis of Foxn3 during embryogenesis. **Embryos obtained from wild type and heterozygote intercrosses at E11.5 and E13.5 were fixed in 2% paraformaldehyde, stained with X-gal overnight and photographed. Representative photographs show *Foxn3* heterozygous embryos (stained blue in panel b b, panel c (bottom), panel f and panel g and control wild type littermate embryo Panel a, panel c (top) and panel e that were stained similarly. Panel h shows RT-PCR expression of Foxn3 in various adult tissues.**

Examination of the developing mice showed that homozygous mutants exhibited severe runting and reduction in head circumference, which is indicative of craniofacial developmental defects (Figure 7). The average body weights of the wild type and mutant animals at weaning were 13.5 $\pm$0.70g (FoxN3+/+, n=3), 12.9$\pm$ 0.70g (FoxN3+/m, n=6), and 4.7$\pm$0.43g (FoxN3m/m, n=4).

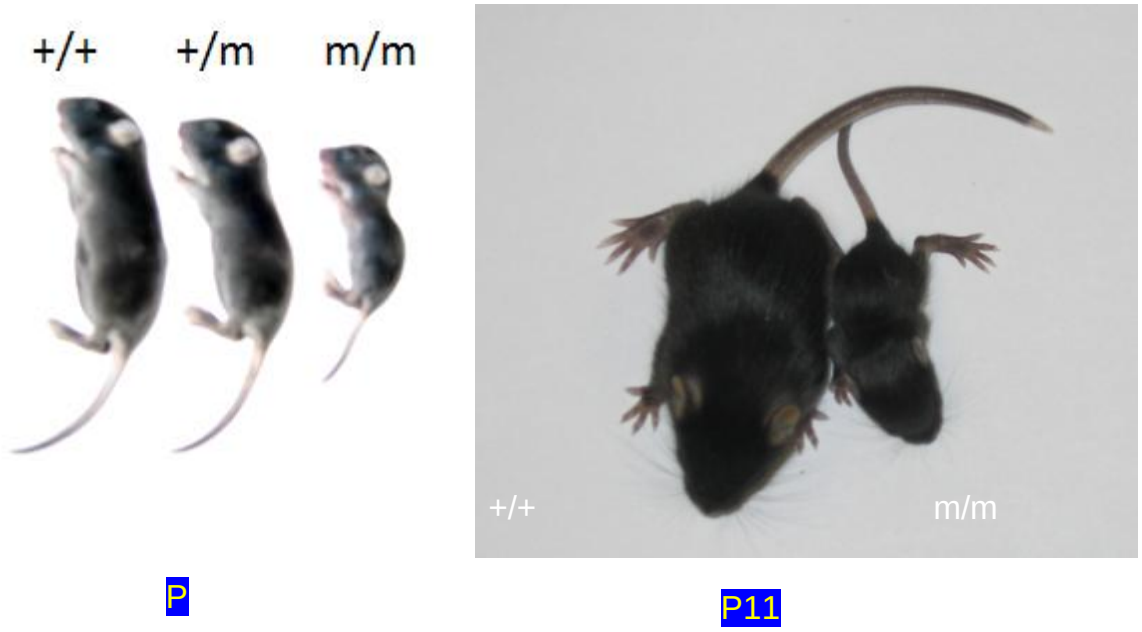

Figure 7. Photographs of FoxN3 WT and mutant mice at P8 and P11 showing the runting phenotype.

To further characterize the microcephaly phenotypes, we performed imaging studies of the mutant and wild type littermate controls at stages P2, P8, P10, and P11 using micro-computed tomography. A majority of the homozygous mutants and a minor subset of heterozygotes showed cranial vault defects of the skull, delayed suture closure and defects in the frontal, parietal and occipital bones of the skull (Figure 8 **b-d**). Subsequent histological analysis of calvariae obtained from 11 day old WT and homozygous mutant littermates indicated a drastic reduction in calvarial bone formation in the mutants (Figure 8 **d**) that was consistent with the mCT-scans that showed a reduced density in the homozygous mutants [compare Figure 8 **c** (bottom panel) and Figure 8 **d**]. Examination of the mandible showed reduced size and midline fusion that prevented further growth (Figure 8 **e**). There were apparent midline defects indicative of defective development of the basispehenoid, basioccipital, tympanic, and palatal bones Figure 8 **e**).

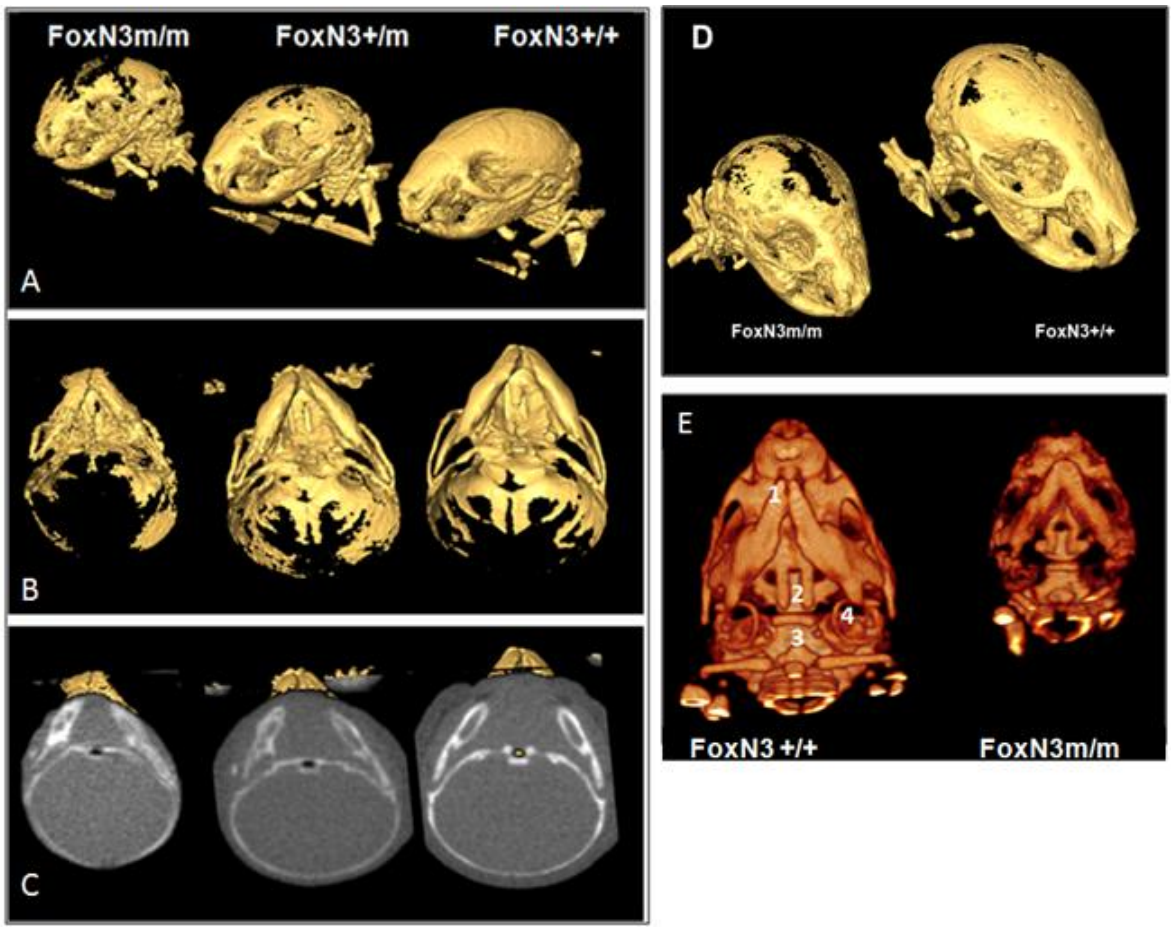

Figure 8. Craniofacial skeletal developmental defects in Ches1 mutants. **A-B: Isosurface rendering of micro-computed tomography (CT) images of the skull images of 8 day old WT and mutant pups are shown in panels A (lateral view) and B (coronal plane of the dorsal aspect). The mutants have reduced bone structure in the skull bones that appear as hollow structures because image thresh-holding was optimized for viewing the calcified structures. C: Coronal planar view through base of skull. D: CT scan comparison of 11 day old wt and homozygous mutant pups. E: Volume texture rendering of microCT images of 2 day old wild type and mutant littermate pups shows retarded growth of the mandible which appears to be fused at the midline prematurely. For comparison, the premaxillary area (1), basispehenoid (2), basioccipital (3), and tympanic bones (4) are indicated in the WT animal on the left. High-resolution CT images were acquired using a MicroCAT™ II + SPECT instrument (Siemens Medical Solutions, Molecular Imaging, Knoxville, TN). Each image comprised 360 projections at 1 degree intervals and were acquired with an x-ray source energy of 80 kVp. CT data were rendered using the Amira 3-D image analysis software package (Amira, Version 3.1: Mercury Computer Systems, Chelmsford, MA).**

We also performed histological analysis of the calvaria in 11 day old wildtype and null mutants which show a reduction in the thickness of the calvaria in the mutant animals (Figure 9).

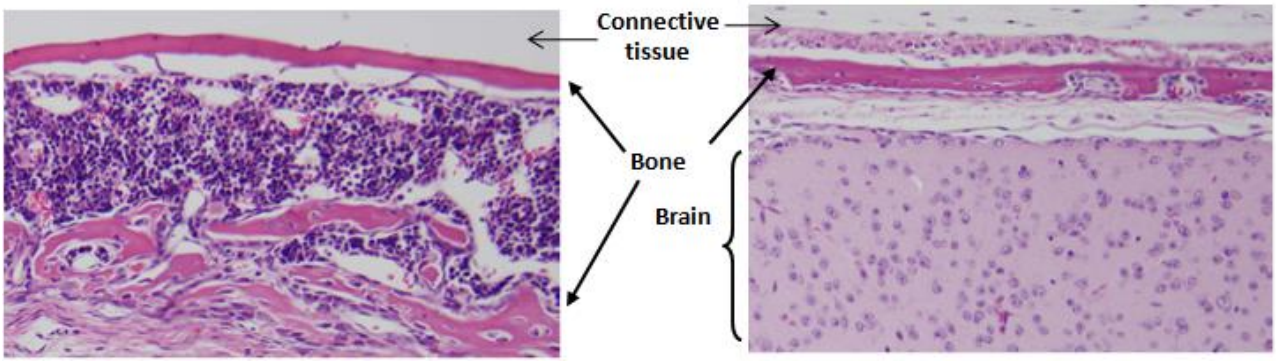

Figure 9. Histological assessment of calvaria in 11 day old wild type (left) and homozygous mutant (right) littermates. **Notice the calvaria is reduced in thickness in the mutant mouse. Saggital sections of the skull were made stained by hematoxylin and eosin and photographed at the same magnification (20X). Identical areas of the same width and length beginning from the connective tissue of the scalp are shown. The drastic reduction in the skull bone thickness of the mutant allows for visualization of a portion of brain in the mutants whereas the same are in the wild type consists only of the skull.**

Furthermore, a subset of mutant animals exhibited eye migration defects. When we compared 13.5 d.p.c wildtype and null mutant embryos, we noticed a cranium development problem, microencephaly, and eye migration problems and size reduction (Figure 10).

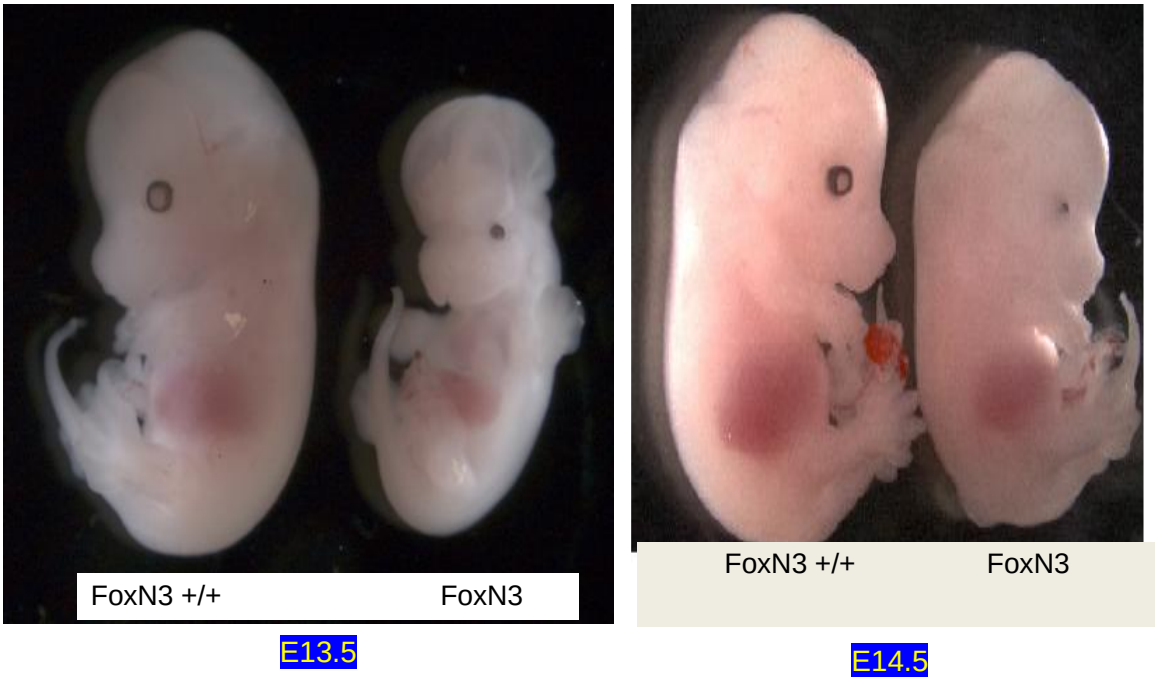

Figure 10. Compound microscope images of 13.5 and 14.5 d.p.c. wildtype and null embryos. **In the left figure, FoxN3 null embryo exhibits microencephaly phenotype along with eye size reduction and migration defect. In the right figure, FoxN3 null embryos show a great reduction in eye size that could be linked to an eye migration defect.**

In addition to the craniofacial defects, among the small fraction of the surviving homozygous mutants and the heterozygotes, a high percentage of the mutant mice also presented malocclusion due to elongated incisors. Examination of 5 month old adult mutant mice indicated an overall growth retardation of the mutants and scoliosis like phenotype (hunch-back spine) in comparison to WT littermates. To correlate the phenotypes of the FoxN3 mutant mice, we performed a comparative analysis of human patients with congenital disorders and deletions of the human *FoxN3* chromosomal locus (14q32.11) reported in the literature. Interestingly, there was a substantial overlap of the clinical features of human patients and the FoxN3 mutant mouse model further underlining the importance of FoxN3 in craniofacial development (Table 2. Phenotype comparison of human patients with 14q deletions and FoxN3 mutant mice.).

Table 2. Phenotype comparison of human patients with 14q deletions and FoxN3 mutant mice.

Clinical feature	**14q deletions (humans)**	**FoxN3 mutant mice phenotypes**
Growth Retardation	+	+
Microcephaly	+	+
Eye deviations	+	+/-
Eye lid opening defects	+	+/-
Dental abnormalities	+	+
Ear abnormalities	+	ND
Scoliosis	+	+
Jaw abnormalities	+	+
Renal abnormalities	+	ND

To examine the mechanism by which FoxN3 affects craniofacial development, we tested the expression of several genes known to regulate bone morphogenesis. Total RNA was extracted from cranium of P0 wildtype and homozygous mutant mice and cDNA was synthesized. FoxN3 levels were reduced in FoxN3 null mice compared to wildtype; however, there was leaky expression of the wildtype mRNA (Figure 11). This leaky expression could be beneficial as the FoxN3 gene may act as a hypomorphic allele, and the complete knockout of FoxN3 may cause the death of homozygous mutant mice during embryogenesis. Additionally, FoxN3 homozygous mutants showed reduced expression of the Bone morphogenetic proteins 2 and 4 (Bmp2 and Bmp4) compared to the wildtype littermates (Figure 11). Similar expression patterns were shown in mouse embryonic fibroblasts (MEFs).

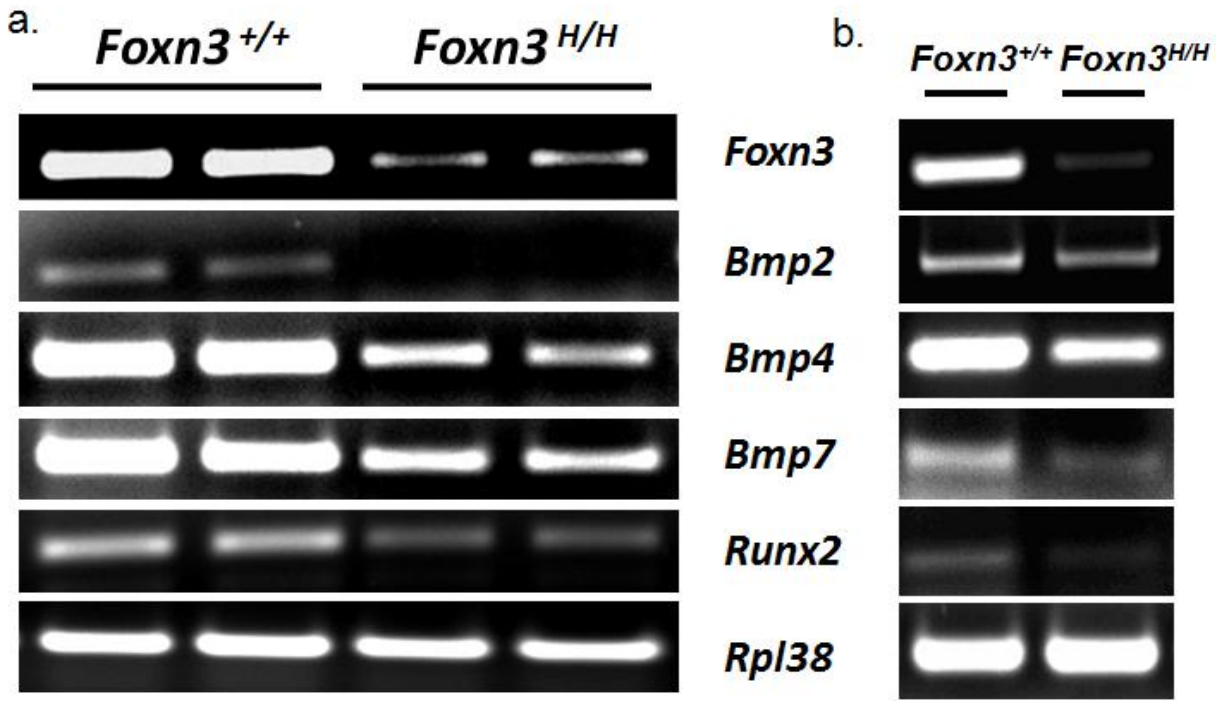

Figure 11. Reduced expression of osteogenic genes in Foxn3 mutant mice and cell lines. **Total RNA was isolated from neonatal (P1) skulls (a) and MEFs (b) of wild type and homozygous mutants. RT-PCR analysis was performed with primers specific for the indicated genes and the house-keeping gene encoding the ribosomal protein 38 (last panel).**

Discussion

Role of Foxn3 in Mammalian Development

The culmination of our data indicate that the *Foxn3* gene plays a crucial role in mouse development as the hypomorphic mutant mice displayed craniofacial developmental defects that lead to embryonic and postnatal lethality. The craniofacial developmental defects in the *Foxn3* mutants are also correlated with the craniofacial tissue specific expression of *Foxn3* during embryonic development. Furthermore, the phenotypes of the *Foxn3* mutant mice are consistent with recent studies that have shown the involvement of *Foxn3* in craniofacial and eye development in *X. laevis* (122). More importantly, the comparative analysis of the phenotypes of the *Foxn*3 mutant mouse model with the clinical features of human patients with deletion of the *FOXN3* locus shows a substantial overlap of the phenotypes that provide initial clues to the role of *FOXN3* in craniofacial development and underline its importance in mammalian development (Table 2).

The mouse model we have described was generated utilizing the gene-trap strategy and involves the random disruption of genes in mouse embryonic stem cells. There are many advantages and disadvantages to choosing the gene-trap approach over the conventional knockout strategy. The gene-trap is a cost efficient, high throughput technology that allows the phenotypic screening of many genes in a short amount of time (125). The Engrailed-2 intronic sequence simulates the intronic region present in a mouse intron, which allows for the spliceosome machinery to bind and recognize the sequence. The splice acceptor (SA) site adds the last exon in-frame to an ATG-less B-galacotsidase-neomycin gene followed by a stop codon and a ploy

adenylated tail. Upon efficient splicing of the gene into the gene trap, exons upstream of the gene-trap will be fused in-frame to the B-galacotsidase neomycin gene and express a fusion protein while the remainder exons will not be translated. Nevertheless, this fusion protein can cause a dominant-negative phenotype as reported by our group or may result in the loss of cytoplasm-nuclear shuffling (126). Although gene-trap based mutations can give rise to the dominant-negative phenotype, a classical homologous recombination based germ-line knockout can give rise to a similar phenotype as was the case in Terry Van Dyke's attempt to remove exons 2 and 3 from the Bub1 gene (127). However, an inefficient splicing of the upstream exon in to the downstream exon (via the circumvention of the splice acceptor present in the gene-trap) has been shown to occur in a minor percentage of gene trap mutants by our group and others (126, 128). Yet, this leakiness in the gene-trap allows for examining the functions of genes which will be lethal in null progeny (e.g. Bub1) (129). Another advantage for using the gene-trap is the tissue-specific reverseal of the gene trap by the expression of a Cre recombinase that clips two LoxP sites flanking the SA site. The Cre recombinase is driven by the binding of a tissue-specific protein to activate the Cre-gene promoter.

Interestingly, our data shows that the *Foxn3* gene trap is not completely efficient in disrupting the expression of Foxn3 as the homozygous mutant tissues do express the wild type mRNA albeit at low levels (Figure 11). This leaky expression of the native *Foxn3* gene in the homozygous mutants further explains the partial lethality and segmental craniofacial defects of the homozygous mutants. More importantly, the hypomorphic alleles generated by gene trap mutants provide a novel opportunity to study the role of gene function(s) at the organismal level that provide data on the

developmental mechanisms involving heterotypic cell-cell interactions. Such analyses are not possible with knock-out mouse models that suffer from disadvantages relating to embryonic lethality phenotypes or conditional mutants that are not amenable to functional studies involving cell and tissue specific heterotypic interactions during mammalian development.

Expression Analysis of Foxn3 in Developing Embryo

The β-galctosidase fusion protein present in the gene trap is advantageous as it allows visualization of the expression of Foxn3 during development. β-galactosidase breaks down X-Gal into 5-bromo-4 chloroindole which gives a bright blue color. Since the expression of β-galctosidase is correlated to *Foxn3* expression, color change will only occur in specific tissue that express *Foxn3*. Figure 6 shows a specific expression of Foxn3 in the developing cranium and neural tube (brain, and spinal cord) of E13.5 embryos. We were able to see high expression of Foxn3 in the eyes of E11.5 and E13.5 embryos (Figure 6, and data not shown). It was during the examination of these embryos that we noticed many of the phenotypes previously associated with Foxn3 mutations. Figure 10 shows stereomicroscope images of wildtype and *Foxn3* mutant E13.5 and E14.5 embryos. Comparison of littermate embryos shows microcephaly and cranium development defects in the mutant embryos compared to the wildtype embryos. Furthermore, the null embryos exhibit great reduction in eye size and eye migration defects which can be linked to cranium development defects. We could not attribute the reduction in eye size to reduction in cell proliferation, cell mass, or increased apoptosis; however, Schuff and his group indicated an increase in apoptosis in *FoxN3* depleted *X. laevis* (122).

Foxn3's importance in development is demonstrated by the embryonic and post-natal lethality of *Foxn3* mutant mice (Table 1). Matings of *Foxn3* heterozygous mice indicated 61% survival of the nullizygous offspring at day 1, which indicates developmental and embryonic lethality. However, at weaning (day 21), only 25% of nullizygous mice survived, which indicates more developmental and post-natal lethality in Foxn3 mutant mice. Most of the nullizygous mice died by post-natal day 3, and many of the nullizygous mice that survived exhibited a runting phenotype when compared to littermate heterozygous or wildtype mice (Figure 7). Measurement of the body mass of *Foxn3* null mice indicated a three fold reduction when compared to wildtype littermates. An interesting pattern that emerged from expansion of the null mice colony was the increased survival of pups born from homozygous null mutants; we observed a high number of mice that survived through weaning which contrasted with the post-natal lethality (data not shown). *Foxn3* expression of nullizygous postnatal day 1 mice generated from null crossings was much higher than the expression of similar mice obtained from heterozygous intercrosses (Figure 11a, and data not shown). This can be attributed to variable expressivity of a gene where individuals with the same genotype can display differing phenotypic characteristics (130).

The underlying etiology of the post-natal lethality in Foxn3 mice was not determined, but previous work on *X. laevis* and our observations on average body mass may yield clues. Schuff and colleagues reported lost or deformed cranial nerves in *X. laevis* including hypoglossal and branches of the trigeminal nerve innervate the lower jaw. Moreover, they reported on the death of *X. laevis* embryos by stage 50 for reduced food intake. The loss of innervation of the lower jaw signifies reduced orifice function

and thus reduced food intake. The loss of lower jaw movement has more severe effects in a developing mouse as it initially has to suckle from the mother to feed and later chew solid food to feed and develop. The runting phenotype in our *Foxn3* null mice can be explained by their reduced food intake due to reduced lower jaw function or misalignment of the jaw. We have also observed that some of our heterozygous and null mice die from abnormal growth of the incisors or malocclusion (data not shown).

Perhaps the most significant finding of my work was the role of *Foxn3* in mammalian craniofacial bone development. High resolution micro CT images of 8 day old and 11 day old wildtype and mutant pups shows reduced bone structure and density in the skull of *Foxn3* mutant pups compared to wildtype littermates (Figure 6 A,B). mCT analysis also show absence or reduced bone structure on the oremaxillary area, the basisphenoid bone, the basioccipital bone, and the tympanic bone in 2 day old Foxn3 mutant pups (Figure 6 E). We confirmed the bone density defect in *Foxn3* mutant pups by histological evaluations of calvarial cross-sections, and we noticed severe reduction of the thickness of the calvaria and skull bones in the mutant pups (Figure 9). The underlying etiology of the reduced bone density in *Foxn3* mutant mice could be attributed to variation in the osteoclast to osteoblast ratio in Foxn3 mutants or developmental defect in osteoblast differentiation leading to reduced calcification of cartilage (131). Therefore, we examined the expression profile of prototypical genes involved in osteogenesis that may have been deregulated in Foxn3 mutant mice. As shown in Figure 11, the expression of bone morphogeneic proteins 2,3,4,7 and Runx2 were all down regulated in Foxn3 mutant skulls compared to wildtype littermates. Even more, the down regulation of these genes is consistent in Foxn3 mutant mouse

embryonic fibroblasts (MEFs) generated from other heterozygous crosses. These genes were tested because they have been shown to play a role in osteogenesis and also contained a Forkhead protein consensus sequence of [(A/G)(T/C)AAA(C/T)A] in their proximal promoter (132, 133).

Craniofacial Development and FOXN3

The mouse skull is composed of frontal, parietal, interparietal and occipital bones that have specifically evolved to surround and protect the brain. These skeletal components receive lineage contributions from both the cranial neural crest cells and the paraxial mesoderm, both of which migrate to defined locations overlying the brain and subsequently differentiate into osteogenic and chondrogenic mesenchyme between embryonic stages E7.5 and E11.5 (132). The evolution of neural crest cells has been postulated as the foundation for the initial appearance and evolutionary expansion of vertebrates as it is unique to vertebrates (134). The skull vault primarily develops by intramembranous ossification, characterized by direct differentiation of the osteogenic mesenchyme into osteoblasts as opposed to endochondral ossification. During skull development, the margins of each bone anlage are populated by highly proliferative osteoprogenitor cells thereby maintaining calvarial expansion (135). By E15.5 the individual skull bones have acquired their basic structure and are separated by sutures which are composed of fibroblasts and skeletal mesenchyme. Calvarial bone growth and expansion is coordinated by the growth of the brain through continued production of the osteoprogenitors present within the suture (135). Several growth factors and transcription factors have been implicated in craniofacial and skeletal development (133, 136, 137). Of these, the fibroblast growth factors (FGF) and a number of

homeodomain proteins (Msx2, Dlx5, Engrailed 1, and Alx4) have been shown to regulate calvarial bone development (138, 139). Evidence indicates that the spatiotemporal expression of several of these growth and transcription factors and their interactions with each other regulate craniofacial skeletal development. The reduced expression of the osteogenic genes in the craniofacial tissues of the *Foxn3* homozygous mutants suggests that the Foxn3 protein activates the expression of genes necessary for osteogenesis and plays a regulatory role in the craniofacial skeletal development. However, the molecular mechanism(s) of the aberrant expression of osteogenic genes in the *Foxn3* mutants remain to be ascertained. Interestingly, the FOXN3 protein has been shown to associate with Sin3, a component of the Sin3/Rpd3 histone deacetylase complex (HDAC) in budding yeast and inhibit the activity of Sin3/Rpd3 HDAC complex. The Sin3 complex is targeted to specific promoter regions via Sin3 interactions with site-specific DNA-binding proteins. Our data suggests that the loss of osteogenic gene expression in the *Foxn3* mutants is potentially due to the loss of inhibition of Sin3 complex that can in turn lead to deacetylation of histones within the promoters of the osteogenic genes resulting in their repression. Nonetheless, a direct role for Foxn3 in the transcriptional activation of specific genes necessary for craniofacial development cannot be ruled out as the interaction studies have utilized heterologous systems (*FOXN3* is not coded in the yeast genome). Furthermore, the interaction of Foxn3 with the transcriptional adaptor SKIP suggests that its role in the transcriptional regulation of its target genes might involve and be dictated by its interacting partners in a cell type and tissue specific manner. It is noteworthy that that the transcriptional activation and

repression functions of other Fox proteins have been shown to be cell-context dependent (92).

Correlation of Foxn3 and Human Disease

Our assessment of the role of *Foxn3* in developmental disease significantly correlates with congenital and developmental defects in human patients with deletions of chromosomal region 14q32.11 (Table 2). The most observable phenotypes were the growth retardation, eye deviations, scoliosis, and jaw abnormalities including malocclusions. Moreover, we observed abnormal kidney morphology and function as we performed necropsies on *Foxn3* heterozygous and null adult mice, but we have not yet confirmed these observations by histological assessments. The remarkable overlap of phenotypes in Foxn3 mutant mice and human patients with *FOXN3* locus deletions signifies the conservation of the role of Foxn3 in mammalian development and may provide an underlying mechanism of human disease. Indeed, expression of bone morphogenic protein 4 (BMP4) in mice has been shown to increase bone fractures healing time and bone density (140). Moreover, Wright and colleagues showed that muscle derived stem cells are capable of undergoing osteogenic lineage differentiation when transfected with BMP4-encoding retrovirus (141). This suggests that ectopic expression of BMP4 may be utilized to treat patients with craniofacial bone development defects or such application may be used to treat human patients with facial bone fractures or breaks. However, we have not yet evaluated the underlying molecular mechanism that *FOXN3* utilizes to regulate bone morphogenic factors expression.

Chapter III

FOX Proteins, Cell-Cycle Regulation and Tumorgenesis

Previous studies on FOX proteins and cancer

The forkhead family of proteins has undergone an evolutionary expansion in higher eukaryotes and mammals compared to C. *elegans* and *Drosophila* (142). This allows adaptation of FOX proteins to the increased developmental and tissue complexity required in higher organisms. FOX members have been shown to play diverse roles in development, immune system, apoptosis, cell cycle control, cancer, and invasion of metastatic tumors (143). FOX proteins are terminal effectors for many signaling pathways, and mutations in FOX proteins perturb proper activation of these pathways to maintain homeostasis. Correlation of gene function can be made through *in-vivo* studies on rodent models because all human FOX genes have orthologues in mouse, except for FOXD family which has one member (Foxd1) in mouse but seven members in humans due to a recent duplication (144). Mutations in several Fox subfamilies (FoxA, FoxC, FoxL, FoxM, FoxO, and FoxP) have been shown to promote tumorogenesis in mouse models in a specific or redundant form. FoxC proteins are involved in the development of vascular system, and are correlated to cancer through a role in vascular angiogenesis (142). FoxL1 is expressed in the messenchyme of the GI tract and null mutation to Foxl1 reduced epithelial cell proliferation due to disrubtion of the Wnt pathway in this organ. Furthermore, $Apc^{+/Min}$;$Foxl1^{-/-}$ mice showed higher tumor load in the colon compared with $APC^{+/Min}$;$Foxl1^{+/+}$ and $Apc^{+/Min}$; $Foxl1^{+/-}$ mice thus showing synergistic effect of mutations in adenomatous polyposis coli (APC) gene and Foxl1. This suggests that some FOX genes can be deregulated during tumorogenesis but may not be able to initiate it on their own. There are a few FOX family members that

regulate cell-cycle processes and can thus be considered true tumor suppressor genes or proto-oncogenes (93).

FOXM family

FOXM1 is a proliferation-associated transcription factor that promotes S-phase entry and proper execution of mitosis in rapidly dividing cells (99, 145). FOXM1 gene is up-regulated in pancreatic cancer and basal cell carcinoma due to the transcriptional regulation by Sonic Hedgehog (SHH) pathway (142). FOXM1 has splice variant proteins that play a crucial role in mitosis and G1/S phase and G2-M cell-cycle progression through regulation of different proteins. Foxm1b transcription factor is essential for liver development, hepatoblast regeneration and hepatocyte mitosis during liver morphogenesis. Conditional deletion of Foxm1b results in diminished DNA replication and failure of hepatoblast to enter mitosis due to decreased levels of Polo-like kinase 1 and Aurora B kinase, which regulate proteins during mitosis (145). Foxm1b also regulates the induction of c-Myc, c-Fos, and cyclin-B, all associated with cell proliferation and S-phase entry (145). The splice variant FOXM1a binds the same DNA sequences as FOXM1b, but does not trans-activate genes (146). In human studies, the FOXM1 gene is located on the chromosomal locus 12p13, a region typically amplified in breast adenocarcinomas, head and neck squamous cell carcinomas, and cervical squamous carcinomas. This region also contains *CDKN1B* (encodes p27), so it remains possible that FOXM1 may not initiate tumorogenesis, but rather enhancing its progression (142).

FOXO family

The FOXO family is considered to function as a tumor suppressor family through promoting cell cycle arrest and apoptosis. FOXO family expression varies between cell types or organs. FoxO1 is highly expressed in adipose tissue, while FoxO4 is in muscle and FoxO3a in liver and FoxO6 is exclusively in brain. All FoxOs bind to the same FOXO recognition element 5'-(G/C)(T/A)AA(C/T)AA-3' and can essentially regulate the same set of genes through binding of this motif. The FOXO family plays a dual role as transcriptional activators and transcriptional repressors, and they exhibit nuclear shuttling by having nuclear localizing sequence (147, 148).Initial correlation of FOXO family with cancer was revealed in translocations in tumors especially in alveolar rhabdomyosarcoma where a translocation in (2;13) and (1;13) results in PAX3-FoxO1 and PAX7-FoxO1 fusion proteins (142). Since then, studies in mice have revealed that single FoxO mutant mice form mostly neoplasias while triple mutants (FoxO1, FoxO3A, FoxO4) can induce Myeloid-lymphoid leukaemia (MLL) and thymic lymphomas, thus exhibiting redundancy and compensation in the FoxO subfamily (149). FOXO family protein regulate cell cycle arrest and apoptosis by regulating the expression of cyclin-dependent kinase (Cdk) inhibitors $p27^{Kip1}$ and $p21^{Waf1}$ or repressing the expression of Cyclins D1 and D2 (150). FOXO family proteins are involved in glucose homeostasis and activation of multiple insulin-responsive genes; FOXO are targets for PI3K-Akt mediated phosphorylation to regulate glucose homeostasis (84). Phosphorylation of FOXO family protein by Akt protein kinases causes the retention of FOXO proteins in the cytoplasm, thus allowing for cell cycle to proceed (148).

FOXP family

The FOXP family members have been implicated to function as either oncogenes or tumor suppressor genes, depending on various contexts. FOXP family members have been shown to play roles in immune response regulation, B-cell development, tumorogenesis, and neuronal development (151-153). A well-known member of this family is FOXP2, which is associated with many developmental disorders most notably speech and expressive language impairments, but has not been associated with cancer (154). FOXP1 was the first member of this family to be cloned, and differently from FOXO family members, is expressed in all tissues. In several subtypes of B-cell non-Hodgkin lymphomas, high-level FOXP1 protein expression is achieved through recurrent chromosome translocations involving the *FOXP1* locus and is a marker for poor prognosis for patients (155). Alternative splicing leads to N-terminal truncations of FOXP1 and over expression of short oncogenic isoforms of FOXP1 in diffuse large B-cell lymphoma (DLBCL) (153). Alternatively, nuclear expression of FOXP1 in breast tumors was associated with improved patient survival, and thus was predicted to function as a tumor suppressor protein (156). FOXP1 and FOXP2 have also been shown to act as repressors on genes involved in lung morphogenesis (157). FOXP3 is the last member of the FOXP family, and it's been linked to the immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) (158). FOXP3 attenuates the expression of cytokine necessary in T cell development by repressing the activity of nuclear factor of activated T-cells (NFAT) (159). FOXP3 is a transcriptional repressor of the onco-protein c-Myc, and somatic inactivation or deletion of X-linked FOXP3 has been observed in patients with prostate cancer (160). Additionally, FOXP3 loss of

expression is observed in aggressive breast cancer in humans, and predisposes mice to breast cancer (93).

FOXN3 and cancer

Comparative real-time RT-PCR studies have shown that FoxN3 (Ches1) is under-expressed in 46% of human oral squamous cell carcinoma samples, and its expression was decreased by 15 fold in OSCC tissue with a concomitant over-expression of CDK1, NPM and NDRG1 (161). Ches1 and CDK1 were predicted to be linked in a regulatory network during tumorogenesis (161). Ches1 was also under-expressed in clear renal cell carcinomas (cRCC) compared to normal tissue (162). In a study of laryngeal carcinoma markers, Ches1, along with ADAM-12, CDK2, and KIF-14 were found to be good specific markers for this cancer (163). The down-regulation of *FOXN3* in malignant tissues, along with the finding that human FOXN3 recruits SKIP to repress transcription, indicate that certain oncogenes regulated by FOXN3 may be over-expressed in cancer cells due to decreased FOXN3 levels (116). To further investigate a possible role for FOXN3 in initiating tumorogenesis, we analyzed the heterozygous and some homozygous Foxn3-mutant mice for survival and tumor formation over two years time period.

Materials and Methods

Generation of *Foxn3* mutant mice

Foxn3 mutant mice were generated and maintained as previously described (124). Upon signs of morbidity, mice were euthanized and organs were harvested and saved in 10% formalin solution. Wild type, heterozygous mutant, and nullizygous mutant mice (n= varies) were irradiated with X-ray radiation (7 Gray, 2 Gray/week) beginning at 4 weeks old. Upon signs of morbidity, mice were euthanized and necropsy was performed to check for tumors.

Generation of Mouse Embryonic Fibroblasts.

MEFs were derived from 13.5-day-old *Foxn3+/+, Foxn3+/m* and Foxn3*m/m* embryos. After removal of the intestinal organs, each embryo was washed with PBS (Gibco) and minced using 18 gauge needles. Cells were cultured in Dulbecco's modified Eagle medium with high glucose (HyClone, South Logan, Utah, USA) supplemented with 2mM L-glutamine, 1X antibiotic-antimycotic, 1µl/ml of fungizone (Invitrogen, Carlsbad, CA) , and 15% heat-inactivated fetal bovine serum (Gibco). Primary cells were frozen in aliquots after the first passage.

Immortalization of Foxn3 MEFs

Foxn3 +/+, Foxn3 +/m, Foxn3m/m MEFs were immortalized by transforming the cells with SV-40 (simian vacuolating virus 40) large T antigen containing plasmid pBSSVD2005. Primary Foxn3 passage 2 MEFs were plated at 50% confluency in a 100mm plate and cells were transiently transfected with pBSSVD2005 using Turbofect reagent (Fermentas, Glen Burnie, MD) according to the manufacturer's instructions.

When the cells were just confluent, they were split at a dilution of 1/10. The cells were continued to be split at 1:10 dilutions till passage 5. This type of splitting ensured a strong negative selection against non-transformed cells.

Clonogenicity Assay

SV-40 immortalized Foxn3 +/+, Foxn3 +/m, Foxn3m/m MEFs ($2.5X10^3$) cells were seeded in complete media and treated with 10 J/m2 UV radiation or exposed to 4 Gy dose of X-ray radiation 24hrs after seeding. After ~8 -10 days, the plates were stained with coomassie blue and colonies were counted using a stereomicroscope. The colonies from the various treatments were normalized to the untreated control. For each treatment, the cells were plated in four 60mm2 plates. The whole experiment was repeated twice using two independently derived cell lines.

RNA Interference Transfection

RNAi oligonucleotides were obtained from IDT DNA (F-LUC-Si, FOXN3) and resuspended in nuclease-free water to a concentration of 40µM. HCT116 cells were plated in 60mm-well plates in DMEM media containing 10% fetal bovine serum to give 30–50% confluence. Transfection of the RNAi oligonucleotides was performed using Trifectin (IDT-DNA) to result in a final RNA concentration of 40 nm. The cells were harvested at different time points and lysed in BE lysis buffer plus 1mM PMSF (40mM NaCl, 20mM Tris-HCl pH 7.8, 2mM MgCl2, 0.5% NP40, Benzonase (Novagen, 250U/mL) for Western blot analysis. siRNA sequences were:

F-Luc si: 5'-rGrCrA rUrArU rCrArA rArGrC rArCrA rUrCrA rGrGrU rCrCA C-3'

5'-rGrUrG rGrArC rCrUrG rArUrG rUrGrC rUrUrU rGrArU rArUrG rCrCrU-3'

Foxn3: 5- rGrCrA rUrArU rCrArA rArGrC rArCrA rUrCrA rGrGrU rCrCA C -3

5'-rGrUrG rGrArC rCrUrG rArUrG rUrGrC rUrUrU rGrArA rUrGrC rCrU-3'

Western Blot

HCT116 cells were scraped in phosphate-buffered saline (PBS), pellet at 1'400 RPM, and lysed in BE lysis buffer plus 1mM PMSF (40mM NaCl, 20mM Tris-HCl pH 7.8, 2mM MgCl2, 0.5% NP40, Benzonase (Novagen, 250U/mL). Cell lysates were mixed with reducing Laemmli sample buffer, separated on 8% or 12% SDS–polyacrylamide gels and transferred onto PVDF membranes (Thermo scientific). Membranes were blocked at room temperature for 2 hours in TBS-Tween (100mM Tris-HCl, 200mM NaCl, 0.1% Tween 20) with 5% milk, and incubated with primary antibody overnight at 4^{o}C or at room temperature for 2 hours. After washing with TBS-T, membranes were incubated with horse radish peroxidase-conjugated secondary antibody for 1 hour at room temperature in TBS-T containing 2% milk. Signal was visualized with the Super Signal chemiluminescent reagent (Thermo scientific).

Reagents and Antibodies

Cell culture reagents were obtained from the following sources: fetal calf serum and Antibiotic-Antimycotic (Gibco, Invitrogen). Antibodies were obtained from the following sources: FoxN3 (Santa Cruz Biotechnology, Santa Cruz, CA; sc-54257), Cyclin-B1, Cyclin-D1, CDK2, CDK1 (CDC2), Beta-Tubulin (Gentex, Zeeland, MI).

Statistical analyses and survival curves

Standard error, mean and *P*-values were determined using the statistics software from Microsoft Excel. Kaplan–Meir survival curves were generated and analyzed with Prism 4 (GraphPad Software).

Results

3.1. Characterization of tumorogenesis in Foxn3 mice

The mutant offspring that survived beyond the perinatal stage did not show any overt developmental abnormalities except for a pronounced reduction in size at weaning but was less apparent in 3–4 months old mice. However, after 7–12 months of age, the heterozygous and homozygous mutant mice began to exhibit weight loss, lordokyphosis (hunch-back spine) and loss of vitality. Survival analysis showed drastic reduction in the lifespan of the *Foxn3* heterozygous (+/m) and homozygous (m/m) mutant mice (median lifespan of 80 weeks for heterozygous, median lifespan of 64 weeks for homozygous) (Figure 12). Histological examination of organs harvested from morbid mice showed that a majority of the mice were succumbing to splenic or thymic lymphomas with a minor fraction of adenomas and sarcomas (data not shown). The earliest incidence of lymphomas in the mutant mice was at 26 weeks of age. 50% of nullizygous mutants (8/16) exhibited spontaneous tumor formation and 43% in heterozygous mutants (13/30). In comparison, only 25% of wild-type mice were diagnosed with a tumor during the analysis period (4/12). Wild-type mice also develop lymphomas as a function of age, and such tumors account for about 5–20% incidence as reported for mice of various genetic backgrounds by others (164).

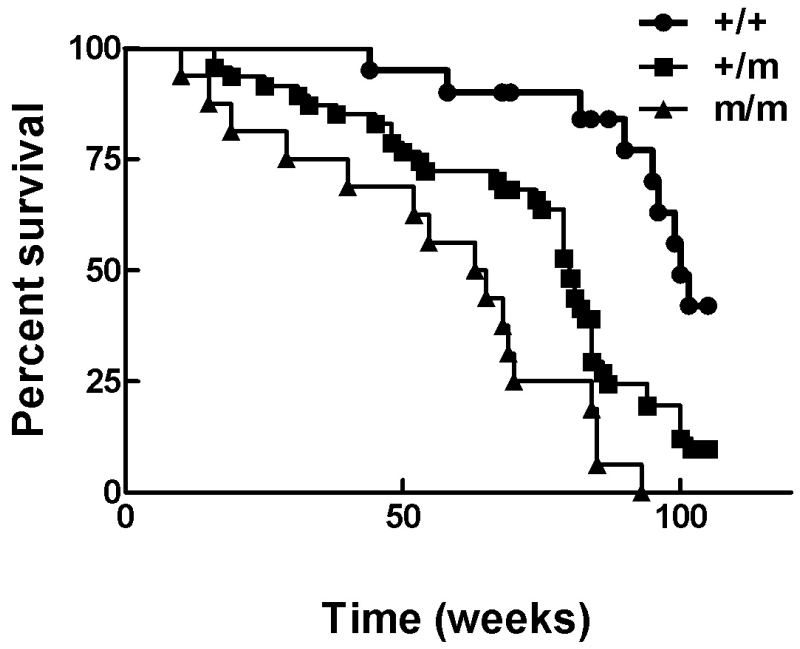

Figure 12. Survival Curve of Foxn3 mice. Kaplan–Meier survival curves of Foxn3 mutant and wild-type littermates **(*n*=20 +/+ , *n*=47 +/m , *n*=16 m/m). The percentages of survival are plotted as a function of age in weeks. Animals were monitored for tumors, morbidity or spontaneous death over a period of 105 weeks. Of the animals analyzed for each group (numbers mentioned above), 40 of the heterozygous mutants died, all 16 of the nullizygous mutants died in comparison with 9 for the wild-type controls during a period of 2 years. The differences in the lifespan of wild type and mutant animals were statistically significant (p<0.0001). All mice were of mixed inbred C57BL/6X129/Sv background.**

Additionally, we wanted to assess whether Foxn3 plays a role in DNA damage response or cell-cycle progression in mice. Therefore, we treated 4 weeks old wild type (WT), heterozygous mutant (Het) and homozygous mutant (Null) with 7 Gray X-ray radiation (1.75 Gy/week, 4 weeks) and checked for signs of lethargy, tumor formation, or death. Survival analysis showed drastic reduction in the lifespan of the *Foxn3* heterozygous (Het) and homozygous (Null) mutant mice (median lifespan of 28.28 weeks for Het, median lifespan of 28.84 weeks for Null) compared to wild-type littermates (median lifespan of 60 weeks) (Figure 13). The earliest incidence of lymphomas in the mutant

mice was at 13 weeks of age. The difference in lifespan between the mutant animals and wild-type littermates was statistically significant (P<0.01), but there was no statistical significance between heterozygous and nullizygous mutant animals (similar median lifespan).

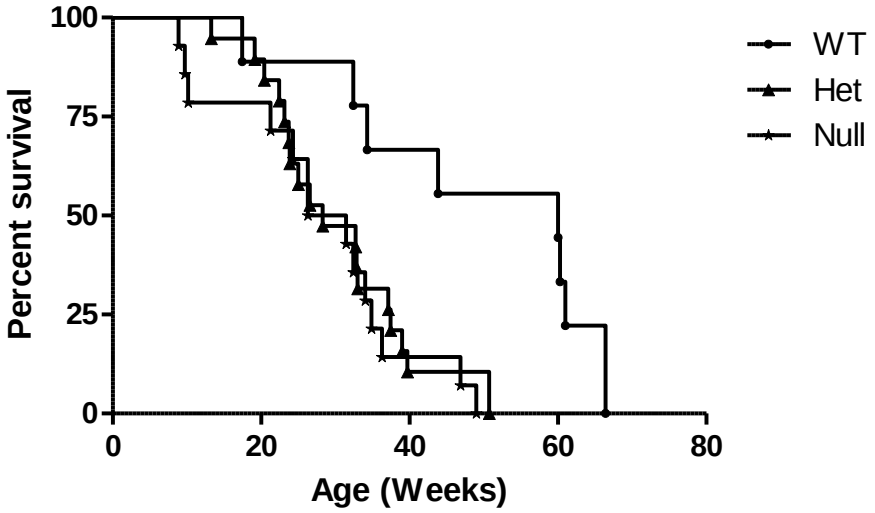

Figure 13. Survival analysis of Foxn3-deficient mice after IR treatment. **Kaplan–Meier survival curves of Foxn3 mutant and wild-type littermates (*n*=9 Wt , *n*=13 Het , *n*=14 Null). The percentages of survival are plotted as a function of age in weeks. Animals were monitored for tumors, morbidity or spontaneous death over a period of 66 weeks. Of the animals analyzed for each group (numbers mentioned above), all 13 of the heterozygous mutants died, all 14 of the nullizygous mutants died in comparison with 6 for the wild-type controls during a period of 66 weeks (all animals were terminated at 66 weeks). The differences in the lifespan of wild type and mutant animals were statistically significant (p<0.01). All mice were of mixed inbred C57BL/6X129/Sv background.**

3.2. Impaired DNA repair in Foxn3 deficient MEFs

Transformation of normal rodent cells into tumorogenic cells requires the deregulation of genes that control cell growth and differentiation. Several genes can be upregulated in combinations to produce this phenomenon (i.e. activated Ras and Myc, activated Ras and activated E1A, simian virus SV40 large T antigen) (165). We used

SV40 large T-antigen transformed *Foxn3* MEFs to perform a clonogenicity assay to test the long term survival of MEFs after DNA damage (UV and X-ray radiation). SV40 transformed *Foxn3* mutant MEFs were strikingly hypersensitive to UV and X-ray radiation treatment compared to wildtype MEFs (Figure 14). These observations signify that immortalization of *Foxn3* heterozygous and homozygous mutant MEFs makes them vulnerable to UV and X-ray radiation DNA damage, which in turn reduces their clonogenic survival.

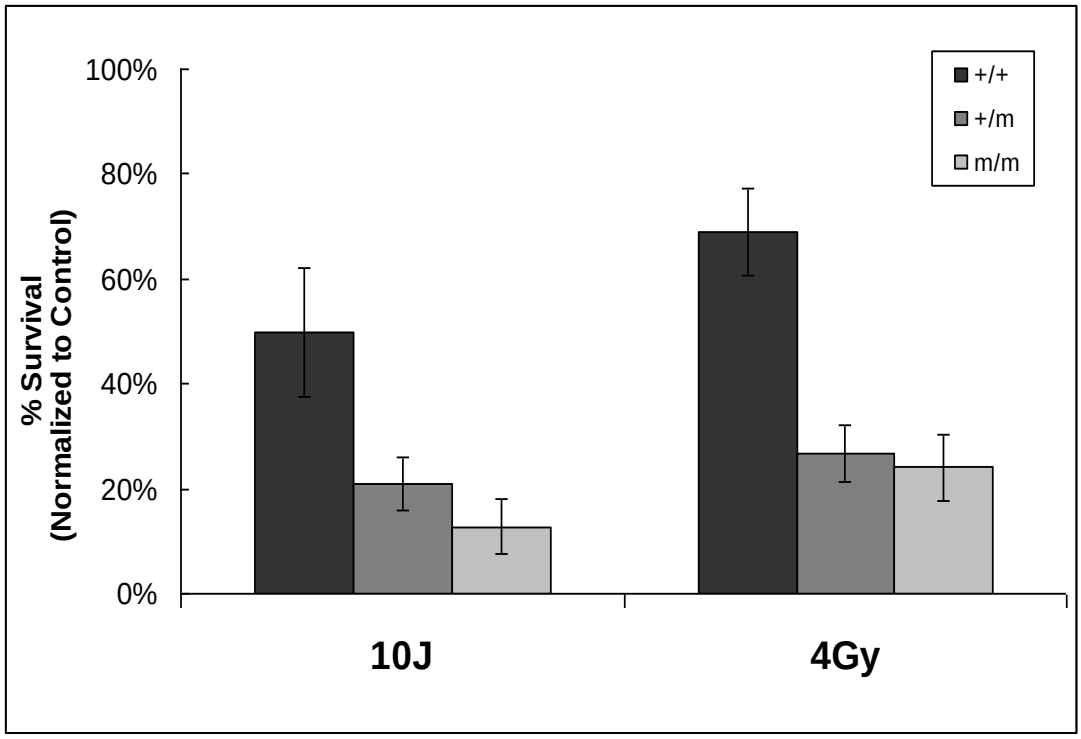

Figure 14. Clonogenicity assay of Foxn3 MEFs. **SV40 transformed wildtype (+/+), heterozygous (+/m) and homozygous (m/m) MEFs were seeded and treated with 10 J/m2 UV radiation or 4 Gy dose of IR. Colonies were counted after 8-10 days. The percentage of untreated cells was set at 100%. Error bars represent SD values for quadruplet samples. Heterozygous and homozygous mutant MEFs were statistically significant compared to wildtypes ($p<0.05$).**

3.3. FOXN3 deficiency deregulates cell-cycle regulatory proteins

RNA interference was effective in reducing FOXN3 protein levels in HCT116 cells. We checked the protein levels of cell-cycle progression proteins including CDK1 (CDC2) and CDK2, Cyclin-B1, Cyclin D1, and Retinoblastoma for any deregulation of expression due to loss of FOXN3. The levels of Cyclin-B1 and CDK2 proteins were reduced in FOXN3 deficient cells compared to F-Luc controls (Figure 15). Meanwhile, Cyclin-D1, Cdc2 (CDK1), and Rb protein levels were not changed due to FOXN3 deficiency, indicating a specific effect of FOXN3 on cell cycle progression proteins. Beta-Tubulin indicates equal loading of cell lysates.

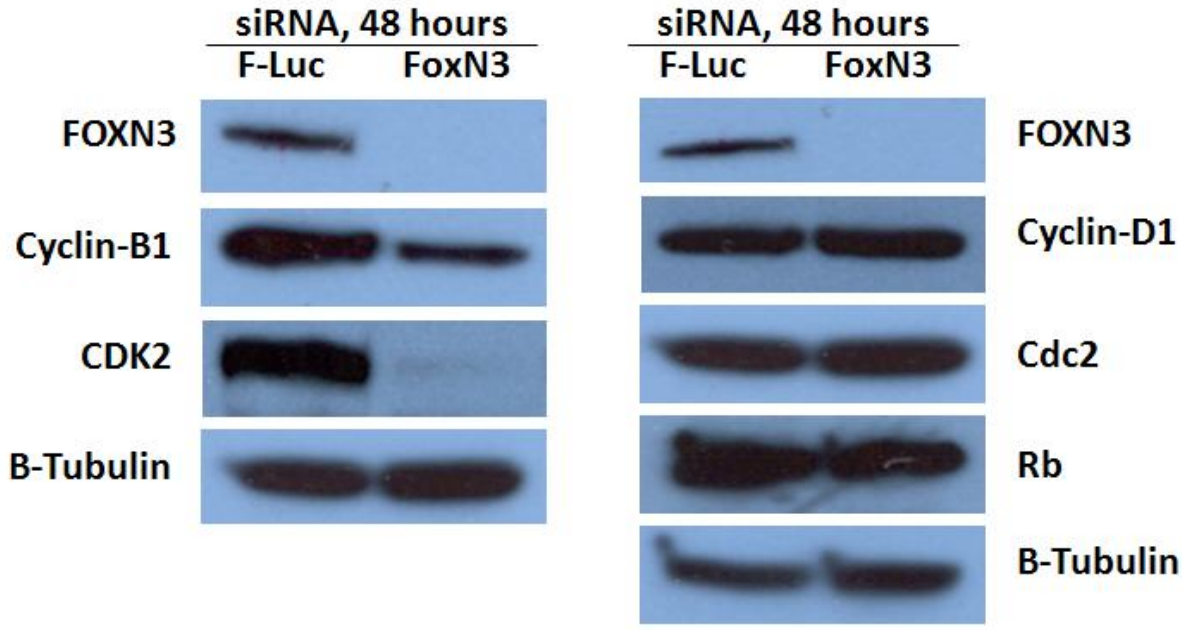

Figure 15. Western blot analysis of cell-cycle progression proteins after RNAi-mediated knockdown of FOXN3. **FOXN3-specific RNAi was efficient in reducing FOXN3 levels in HCT116 cells 48 hours after transfection. Cyclin-B1 and CDK2 proteins were reduced in FOXN3 deficient cells, meanwhile, Cyclin-D1, Cdc2 (CDK1), and Rb protein levels were not changed due to FOXN3 deficiency. Beta tubulin indicates loading control.**

Discussion

The data accumulated thus far indicates a role for Foxn3 in tumorigenesis as heterozygous and homozygous mutant animals exhibit a decreased lifespan and increased spontaneous tumor susceptibility. The majority of the mice succumbed to splenic or thymic lymphomas with some cases of adenomas and sarcomas. This is the first mouse model that studies the role of Foxn3 in initiating tumorigenesis in mammals. All previous reports have shown deregulation of FOXN3 expression in human patients' cancer samples, but never a role for Foxn3 as a true oncogene or tumor suppressor gene (161). Additionally, Foxn3 deficiency can enhance tumorigenesis in mice that have been treated with ionizing radiation (IR) to initiate tumors. Exposure of mice to IR leads to DNA double stand breaks (DSB's), which if left unresolved, can lead to rearrangements and aberrations of chromosomes, thus leading to genomic instability and initiate cancer formation (28, 166). This data suggested a possible role for FOXN3 in DNA repair or regulation of cell-cycle progression. To investigate the underlying mechanism of Foxn3-mediated tumorogenesis, we performed a clonogenicity assay to test the long term survival of Foxn3 mutant MEFS after DNA damage. The colony forming competency of Foxn3 mutant MEFs was greatly compromised after UV and X-ray radiation treatments, which may suggest a role for Foxn3 in the DNA damage repair pathway. However, more experiments are required to further prove the role of Foxn3 in DNA damage repair. Interestingly, we have performed a cell proliferation assay and results indicate a tremendous proliferation capacity for the heterozygous MEFs, while the homozygous and wild type MEFs maintain the same growth kinetics over the same time (data not shown). However, we have not been able to generate three independent MEF cell lines to cover the biological replicas to prove this data.

To determine a functional mechanism for FOXN3 in tumorogenesis and cell-cycle regulation or DNA repair pathways, we checked the levels of specific cell-cycle checkpoint proteins that have either been reported to be deregulated in FOXN3-deficient cancer samples, have a Forkhead Box binding site in their promoter, or are known interaction partners of checkpoint proteins. We were able to shown that cyclin-dependent kinase 2 (CDK2) and Cyclin-B1 levels in FOXN3-deficient HCT116 cells were lower than their Firefly-luciferase (control) transfected cells. Meanwhile, other proteins levels like Cyclin-D1, Retinoblastoma, and Cdc2 (CDK1) were not changed due to FOXN3 deficiency. This contradicted a previous report which showed con-committal reduction of FOXN3 and CDK1 in oral cancers (167). Arguably, analysis done on oral cancer samples can have different driver mutations that can perturb other pathways leading to deregulation of CDK's and other checkpoint proteins (168). Meanwhile, our system ascertains the expression of proteins due to reduction of FOXN3 only in a more controlled experimental setting. Yet, a deeper look into the redundant and cooperative mechanisms of cyclins and CDKs can shed some light on this contradiction.

Regulation of the cell cycle is an essential process in maintaining cellular homeostasis and regulation of cell proliferation. At the epicenter of cell-cycle control are heterodimeric protein kinases composed of a catalytic subunit, the Cyclin-dependent kinase (Cdk), and a regulatory subunit known as Cyclin (169). Each family of cyclins binds to a specific Cdk, forming a complex that is active at a specific phase of the cell cycle. Progression of the cell-cycle is driven by sequential activation of several Cyclin-dependent kinases (Cdk), but mainly Cdk4, Cdk6, Cdk2 and Cdk1. Cues to initiate cell division come from mitogenic factors (IGF-1, MAPK pathway) to induce expression of

cyclin D and thus stimulate the activities of its binding partners Cdk4 and Cdk6 (170). Increasing accumulation of cyclin D/Cdk4, Cyclin D/Cdk6 and cyclin E/Cdk2 complexes, along with proliferating cell nuclear antigen (PCNA) coordinate DNA replication and transition the cell into S phase entry. G2/M transition is regulated by cyclin A/Cdk2 and cyclin B/Cdk1 complexes. Additional factors termed 'cyclin-dependent kinase inhibitors' (CKIs) further control cell cycle progression; $p27^{KIP1}$ and $p21^{CIP1}$are examples of CDKI's that bind to CDKs and prevent their activation. Some of these CDKI's are controlled by tumor suppressor proteins like p53, which can arrest cell cycle progression at any stage to ensure proper proliferation (171). Recent studies in different Cdk knockout mice have shown redundancy in Cdk function and compensation by other complexes to promote cell division. Mice lacking Cdk4 and Cdk2 complete embryonic development but die shortly thereafter presumably; however, conditional ablation of Cdk2 in adult mice lacking Cdk4 does not result in obvious abnormalities (172). Cdk2 and Cdk4 are required to phosphorylate Retinoblastoma (Rb), a protein essential for G1/S cell cycle progression. Additionally, Cdk2 and Cdk4 cooperatively regulate the expression of Cdc2 (CDK1), which cannot compensate for the loss of CDK2 during embryogenesis (170). Recent work has also shown that Rb/Cdk2/Cdk4 triple mutant MEFs sustain enhanced S-phase entry and proliferation rates similar to wild type (173).

Taken together, the loss of CDK2 could have been compensated by CDK4 or RB to progress cells through the G1/S transition. Meanwhile, Cyclin-B1 interacts with CDK1 at G2/M during transition into G2/M phase of the cell cycle and brings the onset of mitosis (174). G2/M phase possesses a complex regulatory network of proteins that ensure proper entry into this phase and proper mitosis. FOXN3-mediated loss of cyclin-B1 can

potentially drive cells into G2/M without any regulatory input from CDK1, thus explaining the cell proliferation seen in Foxn3 heterozygous mutant MEFs. Additionally, a Cyclin A/Cdk2 complex regulates the activation of Cdk1 and Cdc25 (175). This provides another potential link to the role of FOXN3 in regulating the activation of CDKs and controlling cell cycle proliferation, a hallmark of cancer.

Chapter IV

Chromatin Remodeling Proteins and The Role of CHD2 in Tumorgenesis

Chromatin and Nucleosome Dynamics

DNA in the eukaryotic cell is organized into a hierarchical structure called 'chromatin'. Nucleosomes are the fundamental units of the chromatin and they consist of 146 base pairs of DNA wound around a histone octamer. The histone octamer is a positively charged protein unit that wraps the negatively charged DNA around it to condense the DNA. The histone octamer consists of two molecules of each of the histones H2A, H2B, H3, and H4 (176). This initial nucleosome organization forms a 10-nm fiber known as "beads-on-a-string". The binding of the linker histone H1 to the 10-nm fiber allows for organization of the nucleosome arrays into a 30-nm chromatin fiber (177). Higher-order compaction of DNA involves the 'structural maintenance of chromosomes' proteins (SMC) which include condensin and cohesin (178). Intrinsic functions of the cell, such as DNA replication, transcription, and DNA repair require accession of DNA sequences by DNA binding factors at proper regions and during the appropriate times, however the compaction of DNA into chromatin makes it inaccessible to such factors. To overcome this problem, a multitude of factors including histone modifying enzymes and chromatin remodeling complexes work together enhance accessibility of DNA regions to their specific factors. These chromatin modifying proteins regulate the dynamics of nucleosomes, enabling the cell to perform its functions.

Histone modifying proteins

Histones, the protein component of the nucleosome, form two H2A-H2B dimers and one H3-H4 tetramer that form the two nearly symmetrical halves of the histone octamer. Histones fold into a helix-turn-helix region that composes the core of the histone and also contain an N-terminal tail that protrudes outside the compact histone core. Histones are subjected to posttranslational modifications at their N-terminal tails that affect the affinity of DNA wrapping around the histone octamer (179). These post-translational modifications include methylation, acetylation, phosphorylation, ubiquitination, SUMOylation, ADP-ribosylation, and other modifications. These modifications regulate the DNA at a 'higher level' than the underlying DNA sequence; the term "Epigenetics" or "histone code" has been given to these modifications (180). The two best-studied modifications are methylation and acetylation.

Histone methylation of arginine and lysine residues is a major determinant for formation of transcriptionally active and inactive regions of the genome as well as the state of chromatin compaction. Arginine mono or di-methylation of histones H3 (Arg2, 17, 26) and H4 (Arg3) promotes transcriptional activation and a more loose chromatin (181-183). This methylation is mediated by the protein arginine methyltransferases (PRMTs) family that includes the co-activators PRMT1 and CARM1 (PRMT4) (184). In contrast, histone lysine methylation is a more diverse and a better studied histone modification. Lysine methylation is associated with both transcriptional activation as in histone H3 (Lys4, 36, 79) and silencing as in histone H3 (Lys9, 27) and histone H4 (Lys20) (185). Tri-methylation of histone H3 (Lys9, 27) is associated with heterochromatin, a region of highly compacted DNA with no active transcription.

Additionally, methylated lysine histone tails can recruit proteins like HP1 (utilizing their chromodomains) to propagate the methylation pattern to nearby histones, leading to heterochromatin formation. The majority of lysine methyl-transferases identified thus far contain a conserved catalytic SET domain originally identified in the Drosophila Su[var]3-9, Enhancer of zeste, and Trithorax proteins (186). Conversely, methylated histones can become demethylated if acted upon by a group of enzymes that include the amine oxidase LSD1 (lysine specific demethylase 1), hydrolase JmjC (Jumonji) and the arginine deaminase PAD4 (peptidyl arginine deiminase 4) (187). These enzymes can remove methylated lysine or arginine residues (mono or di-methylated) and thus can act as transcriptional repressors or activators, depending on the demethylated residue.

Histone acetylation of lysine residues is another major epigenetic marker typically associated with active transcription and loose compaction of DNA around the histone octamer. Acetylation neutralizes the overall positive charge of the lysine residue, thus reducing the affinity and weakening the binding between negatively charged DNA and positively charged histones. Acetylation of histone H3 (Lys9) and histone H4 (Lys5, 8, 12, 16) are known markers for euchromatin, a region of lightly compacted chromatin associated with active transcription (188). Not only does histone acetylation loosen the DNA compaction, but it also helps recruit bromodomain-possessing proteins involved in transcription. Acetylation of histone H3 (Lys9, 14) recruits the TFIID of the RNA polymerase, while acetylation of histone H4 (Lys8) recruits the SWI/SNF complex (189). Histone acetyl-transferases (HAT) are the family of enzymes responsible for histone acetylation at lysine residues. The transcriptional coactivator families MYST,

Gcn5/PCAF, and CBP/p300 compose the main three families of HAT, but are also capable of acetylating other transcription factors like p53 (64, 190). Histone deacetylases (HDACs), on the other hand, remove the acetyl group from lysine residues of histones and promote compaction of DNA. The transcriptional co-repressor complex Mi2/NuRD, mSin3a, and NCoR/SMRT recruit HDAC1, 2, and 3 respectively to de-acetylate histones and compact the genome (191, 192).

ATP-dependent chromatin remodeling factors

Nucleosomes can be actively modified through ATP-dependent protein complexes that utilize the energy from ATP hydrolysis to physically slide or remove histones and nucleosomes (193). The common feature of the ATP-dependent chromatin remodeling complexes is a highly conserved SNF2-related ATPase domain present in four main classes and variations: The switching/ sucrose non-fermenting (SWI/SNF), imitation switch (ISWI), INO80, and chromodomain helicase DNA binding (CHD), which will be discussed in greatest detail. The INO80 complex is a conserved protein complex composed of the Ino80 ATPase subunit, two AAA+ ATPases (Rvb1/2), Actin and its related proteins Arp4, 5, and 8 (194). The Ino80 ATPase subunit provides the energy for chromatin remodeling through nucleosome sliding. The INO80 complex interacts with transcription factor Ying Yang 1 (YY1), linking INO80 to both activation and repression of transcription. Additionally, INO80 is recruited to DNA double stand breaks (DSBs) in yeast through interactions of its Arp4p and Nhp10p subunits with the phosphorylated H2AX (γ-H2AX) histone variant present at the DSB site. The Drosophila ISWI is another member of the SNF2-related family of ATPases, and it is the catalytic core of three chromatin remodeling complexes: NURF, CHRAC, and ACF (195). Residues 16-19 of

histone H4 have been shown to be required for stimulation of ISWI ATPase activity, which in turn repositions nucleosomes along DNA in *cis* (196).

The Swi2/Snf2 is the third class of SNF2-related ATPases and is the main catalytic subunit of the yeast SWI/SNF complex. SWI/SNF protein complex interacts with GAL4, a sequence-specific transcription factor, and mediate transcription in yeast (197). Homologous subunits to the ATPase domain have also been identified in *S. cerevisiae* (the RSC complex), Drosophila (BAP and PBAP), and at least two complexes in mammals including the BRG1-associated factor complex (BAF), and the polybromo BRG1-associated factor complex (PBAF; homologous to *S. cerevisiae* Rsc) (198, 199). While SWI/SNF complex mediates transcriptional activation in yeast, mammalian SWI/SNF complexes contribute to both repression and activation. During mammalian T lymphocyte development, BRG1 is required to activate CD8 expression and silence CD4 simultaneously (200). The flexibility of SWI/SNF complexes lies within their ability to utilize ATP-dependent chromatin remodeling while also recruiting histone deacetylases (HDACs), thus removing the transcriptional activating acetyl marks from histone tails. ATP-dependent chromatin remodeling involving SWI/SNF complexes is initiated through binding of the complex to nucleosomal DNA, followed by disruption of histone-DNA binding, and ATPase subunit mediated translocation of DNA from histones (198). This forms a DNA loop that can then be propagated through sliding or removal of adjacent nucleosomes to generate stretches of DNA that are more accessible to DNA binding factors.

The CHD family of chromatin remodeling proteins

The chromodomain helicase DNA-binding proteins (CHDs) are an evolutionary conserved family of proteins associated with chromatin remodeling (201). Members of the CHD family of proteins are present in all eukaryotic organism; nine CHD family members have been characterized in vertebrates. The CHD family is divided into three subfamilies depending on the presence of additional motifs and the chromodomain subclass. CHD1 and CHD2 belong to Subfamily I; CHD3 and CHD4 belong to Subfamily II, while CHD5 through CHD9 belong to Subfamily III. CHD family proteins are characterized by the presence of two chromodomains (Chromatin organization modifier), DNA-dependent SNF2-related ATPase domain, DNA-binding domains (HMG-1, PHD Zn-finger) and C-terminal helicase domain. The chromodomain is a 40-50 amino acids motif that binds to methylated lysine residues in the N-terminal tails of histones, thus recognizing epigenetic marks laid out by histone modifying enzymes (202). This domain is present in many proteins involved in chromatin remodeling (Drosophila heterochromatin protein 1, HP1), heterochromatin formation (Polycomb), and regulation of gene expression (RBP1) (203). The chromodomain is formed through a hydrophobic core and hydrophobic groove required for interaction with tri-methylated lysines and dimerization. The SNF2-related ATPase domain is a ~400 amino acids motif that possesses an N-terminal ATP-hydrolysis domain and a C-terminal transduction domain. Proteins that possess SNF2-related ATPase domain can disrupt nucleosomes by either dissociating and/or repositioning (RSC and SWI/SNF), generating ordered arrays of nucleosomes (Chd1 and ISWI), or facilitating the exchange of histones with histone variants of a nucleosome (INO80 and SWR1) (204). CHD proteins mediate chromatin remodeling by first binding to either DNA using their DNA-binding domain, or

methylated histone tails using their chromodomain, and then activate their SNF2-related ATPase/helicase domain (205). Activated ATPase domain facilitates ATP-dependent translocation of the protein along DNA, thus generating force that can be directed into arranged movement or destabilization of nucleosomes.

CHD Subfamily I

CHD subfamily I consists of two proteins: CHD1 and CHD2. Murine CHD1 protein was isolated from cDNA libraries of mouse lymphoid cell mRNA and encodes a 197-kDa protein characterized by the presence of chromodomain, a SWI2/SNF2 like helicase domain, and its ability to bind immunoglobulin promoter sequence (206). CHD1 homologues were identified in most eukaryotic organisms with conserved domain across all species. CHD1 has a pair of N-terminal chromodomains followed by SNF2-related ATPase/helicase domain, a C-terminus helicase domain, and a Myb-related DNA binding domain (201, 207). The human CHD1 is 95% homologous with murine Chd1at the amino acids level and its chromodomains have been shown to bind methylated lysine 4 in the histone H3 tail (H3K4) (204). This provides a mechanism that targets CHD1 to transcriptionally active chromatin, as marked by trimethylated H3K4 (208). However, the two chromodomains of yeast Chd1 cannot bind to methylated H3K4 peptides, and in Drosophila the chromodomains are not required for the chromatin localization of CHD1. The other possible mechanism of CHD1 localization to sequence specific regions lies within its C-terminal DNA-binding domain. Studies on murine Chd1 identified a region of ~250–300 amino acids in its C-terminus that are required for proper association with DNA and chromatin (204). Further analysis of the domain revealed two main motifs similar to the SANT domain (SWI3, ADA2, NCoR, and TFIIB)

and the SLIDE (SANT-like ISWI domain) domain. The Myb DNA binding domain is a 53 amino acids domain that forms a helix-turn helix motif characterized by the presence of aromatic residues (209).

CHD1 has been shown to interact or be a component of many protein complexes in depending on the organism. CHD1 has been shown to interact at the Drosophila's polytene chromosome with the nuclear protein structure-specific recognition protein 1 (SSRP1), which plays a role in homologous recombination-mediated DNA damage response (210). The majority of Chd1 studies were conducted in *S. Cerevisiae,* where Chd1 was shown to interact with Rtf1, a component of the Paf1 complex that interacts with the RNA Polymerase II and regulates transcription elongation (211). Chd1 has been identified as a component of the yeast SAGA/SLIK (Spt, Ada, Gcn5 acetyltransferase / SAGA-like) complex (212). Murine Chd1 interacts with hormone nuclear receptor corepressor (NCoR), thus suggesting a role for mouse Chd1 as a transcriptional repressor. Moreover, yeast Chd1 antagonizes the yeast 'facilitates chromatin transcription' complex (yFACT), which contains the SSRP1 homologue (Pob3) (213). FACT complex facilitates TATA-Binding Protein (TBP) and TFIIA binding to nucleosomal binding sites in addition to its role in transcription elongation.

The other member of the CHD Subfamily I CHD2 shares ~60% similarity with CHD1 at protein level with the least degree of homology at the N and C-termini that lack conserved domains. CHD2 will be discussed in greater details in 'Preliminary data'.

CHD Subfamily II

CHD subfamily II consists of two proteins: CHD3 and CHD4, also known as Mi-2α and Mi-2β, respectively. Mi-2 proteins are present in the animal and plant kingdoms, but are absent in yeast. CHD3 and CHD4 were initially identified as autoanitgens in the connective tissue of patients with the autoimmune disorder dermatomyositis (214). Biochemical analysis of Hela cells identified Mi-2α/β as a major component of the vertebrate Mi-2/NuRD (Nucleosome remodeling deacetylase) complex, a multi-subunit protein complex containing both histone deacetylase activity (HDAC1/2) and ATPase-dependent nucleosome remodeling activity. In addition to the Mi-2 proteins and HDAC1/2 proteins, the NuRD complex contains either one of the metastasis-associated proteins (MTA1, 2, 3), methyl CpG-binding domain-containing protein 3 (MBD3), and the retinoblastoma-associated proteins 46/48 (RbAp46/48) (214). The NuRD complex possesses seemingly contradictory activities in HDAC, which deacetylate histone tails to repress transcription, and ATPase nucleosome remodeling activities, typically associated with activating transcription. However, the NuRD complex is associated with transcriptional repression and can be targeted to different genes depending on the identity of the MTA protein (215). MTA1 has been shown to repress estrogen receptor-associated transcription and interacts directly with ER in breast cancer cell lines and affect ER-mediated BRCA1 gene transcription. Meanwhile MTA3 was shown to regulate Snail transcription and interact with its promoter *in-vivo* (215).

Differently from the CHD subfamily I that possess the C-terminal HMG-1 (Myb-related) DNA-binding domain, Mi-2 proteins have a plant homeodomain (PHD) as the DNA-binding domain located in the N-terminus of the proteins. The PHD domain is

homologous to the leukemia associated protein (LAP) and contains a RING domain (Cys_3-His-Cys_4), a Zinc-finger like domain that binds DNA to activate transcription (216). However, activation of ATPase activity of Mi-2 proteins is stimulated by binding chromatin, not by free DNA or histones (214).

CHD Subfamily III

CHD subfamily III consists of five proteins: CHD5 through CHD9. CHD subfamily III proteins have the two chromodomains, ATPase/helicase domain, and the C-terminal helicase domain; moreover, they possess additional domains like a paired BRK (Brahma and Kismet) domain, a CR domain, SANT-like domains, and a DNA-binding domain (217). CHD subfamily III proteins have been identified in higher eukaryotes and studied mostly in mammals and to some extent in Drosophila. CHD5 has two PHD Zinc finger-like domains and a DEAD-box helicase domain associated with RNA-dependent helicase/ATPase activity. Human CHD5 protein is expressed in neural-derived tissues and was identified in patients with neuroblastoma (217, 218). CHD5 is located on the chromosomal locus 1p36, a region commonly deleted in neural, epithelial and hematopoietic cancers (219). CHD5 was proposed to function as a tumor suppressor protein because Chd5-compromised cells (thus termed due to rearrangement or deletions) had reduced expression of $p16^{Ink4a}$, $p19^{Arf}$, and p53. CpG mediated hypermethylation of CHD5 promoter was found in human cancers including gastric and ovarian cancer, and also in squamous cell carcinoma cell lines (220, 221).

CHD6 is a ubiquitously expressed protein and a component of the PRIC complex (PPARα-interacting cofactor) through interaction with PPAR protein (222). CHD6 also

interacts with the transcription factor Nrf, which is implicated in cellular respiration homeostasis, cell growth and heme biosynthesis (223). Additionally, CHD6 was shown to co-localize with active RNA polymerase, but not a strong physical interaction. CHD6 has a DNA binding domain that has an affinity for A+T rich DNA (AT hook).

CHD7 is also a ubiquitously expressed protein first identified as a genetic mutation in CHARGE syndrome patients (224). Additional studies using mouse models with mutant Chd7 gene phenocopied CHARGE syndrome with mice exhibiting heart defects, atresia of the choanae, and prenatal death (225). In a study to analyze CHD7 chromatin localization in mouse embryonic stem (ES) cells using a ChIP-Seq strategy, CHD7 was shown to localize with ES cell master regulators OCT4, SOX2, and NANOG (226). Additionally, CHD7 interacts with and co-localizes with P300 binding sites, a component of the P300/CBP (CREB-binding protein) which functions as a HAT and a known enhancer-binding protein complex.

CHD8 was first identified as a protein that interacts with the insulator binding protein CTCF at its target site 'differentially methylated region' (DMR) of H19. siRNA mediated CHD8 knockdown abolished CTCF-mediated insulator activity, as the expression of the H19 locus gene IGF2 was increased (227). Additionally, knockdown of CHD8 affected CpG hypermethylation, histone acetylation, and heterochromatin spreading around the CTCF binding sites of BRCA1 and c-Myc genes. Moreover, CHD8 interacts and negatively regulates β-catenin of the canonical Wnt signaling pathway, as shown by increased β-catenin target gene expression in response to loss of CHD8 (228). In a study of CHD8 function in mice, **Chd8**−/− embryos die **in utero** between embryonic day E5.5 and E7.5 due to widespread apoptosis (229). Further investigation

showed an interaction of CHD8 with the tumor suppressor p53 and histone H1. Depletion of CHD8 or histone H1 resulted in p53 activation and apoptosis, and CHD8 promoted histone H1 binding with p53 promoter to form a trimeric complex and inhibit p53-dependent transactivation of its target genes.

The last member of CHD subfamily III proteins, CHD9, was first isolated as a chromatin-related mesenchymal modulator (CReMM) from mesenchymal stromal cells (MSC) (230). Murine CHD9 expression is restricted to marrow stromal progenitor cells during the embryonic development of the mouse skeletal system by binding osteocalcin, collagen-II, myosin, and CBFA1 promoters which affect tissue specificity (217).

Chromodomain Helicase DNA Binding Protein 2 (CHD2)

CHD2 is the second member of the CHD subfamily I proteins and shares 59% homology and 70% identity with CHD1. Human CHD2 is located on chromosome 15q26.1 and encodes an 1828 amino acids protein with a proposed splice variant of 1739 AA protein (Ensembl: ENSG00000173575), while the murine CHD2 is located on chromosome 7D1 and encodes an 1827 amino acids protein (Ensembl: ENSMUST-00000169922). CHD2 has two N-terminal chromodomains, a central SNF2-related ATPase/helicase domain, a C-terminus helicase domain, a Myb-related DNA binding domain, and a C-terminal HMG-I domain (A+T hook). CHD2 is a poorly characterized nuclear protein without biochemical analysis of its chromodomains and DNA binding domains. The function of CHD2 is predicted from its homologue CHD1 and a few reports in the literature. The human CHD2 chromodomain has a 30 fold weaker interaction with the H3K4Me3 modified histone than human CHD1 (208). A microarray profile study showed that CHD2 was down-regulated in blood cells of urinary bladder

cancer patients compared to healthy controls (231). A study utilizing comparative genomic hybridization revealed a homozygous deletion at chromosomal locus 15q26.2 (encodes CHD2 and RGMA) in the Hodgkin's lymphoma cell line HDLM2 (232). Two recent studies revealed that Chd2 is important for mouse embryo development and survival, and Chd2 mutant mice have impaired kidney function (126, 233). Additionally, Chd2 mutant mice develop spontaneous lymphomas and lymphoid hyperplasias. Chd2 mutant mice also exhibit defective hematopoietic SC differentiation and defective DNA damage response (126).

Chd2 mouse models

In an effort to understand the role CHD2 in mammalian development, Chd2 deficient mice were generated in our laboratory using the Baygenomics gene trap embryonic stem cell resource. The Baygenomics insertional mutagenesis strategy involves the use of a gene-trap cassette consisting of a splice-acceptor-βgeo cassette (β-galactosidase-neomycin fusion gene) and characterized using 5' RACE (123). We obtained one of the ES cell clones that had been characterized to have a gene trap insertion within the *Chd2* gene for analysis of the gene trap insertion site. Using multiple intronic forward primers and a gene-trap specific reverse primer for PCR amplifications and sequencing, we determined the gene-trap insertion site to be in intron 1 (Figure 17). Chd2-targeted ES cells were used for blastocyst injections using the microinjection services at the University of Massachusetts Medical School, Worcester. The colonies from germline founders were further expanded for the analysis of the mutant offspring. This mouse model was termed N-terminal Chd2 mutant and differs from the C-terminal mouse model previously generated and characterized (126). The working model for the

Baygenomics gene traps encompasses the splicing of the 5' splice site of Chd2's exon 1 (encoding 20 amino acids) into the 3' splice site in the gene trap (Figure 16). This allows for the in-frame fusion of the first two exons of Chd2 to the B-galacotsidase neomycin sequence and the expression of this fusion protein. The expression of this fusion protein is mandated for the blastocysts to survive neomycin selection. Upon efficient splicing of the Chd2 gene into the gene trap, the remainder exons will not be translated and thus the wildtype protein will be haplo-insufficient in heterozygous progeny and completely lost in nullisomic progeny.

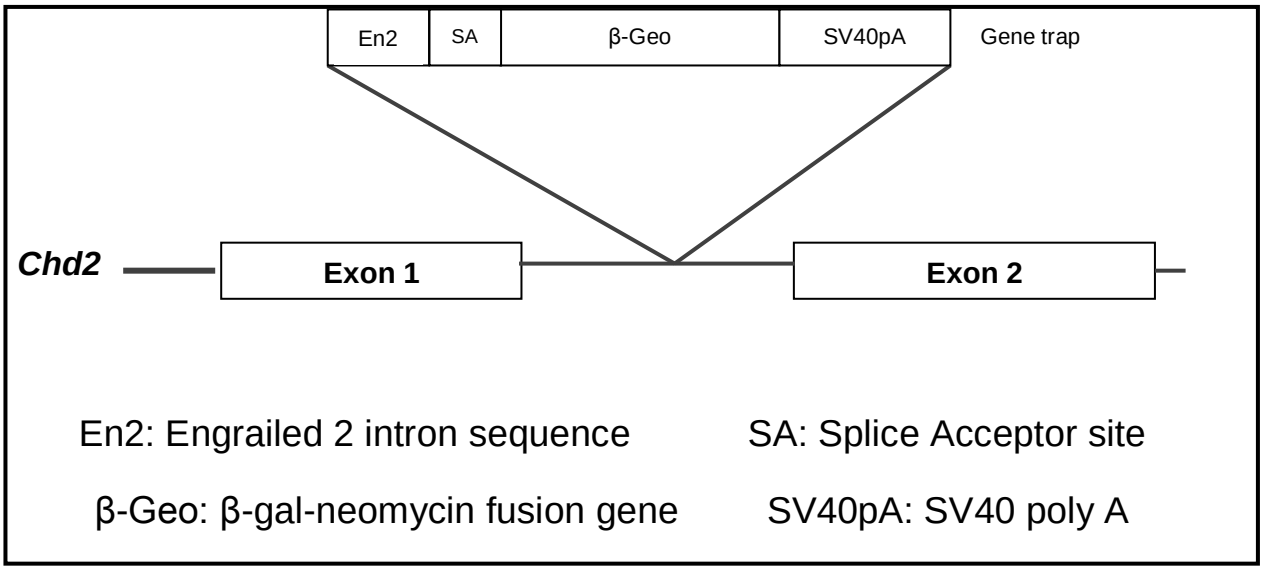

Figure 16. Schematic representation of the Baygenomics gene trap integrated into intron 1 of Chd2 gene. **The gene trap cassette contains an Engrailed 2 intronic sequence followed by a splice acceptor site, an ATG-less B-galactosidase neomycin fusion gene and a ploy-A site.**

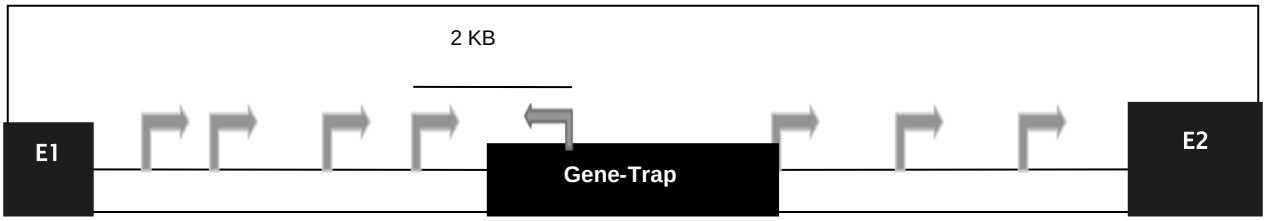

Figure 17. Schematic representation of the strategy utilized to configure the gene-trap insertion site in intron 1 of the Chd2 gene. **Chd2 Intron 1 forward primers were designed at 2kB space intervals and tested against one gene-trap specific reverse primer in WT and Heterozygous mice. Sequencing of the PCR product indicated that the gene trap was integrated within intron 1 (13.9KB from the beginning of the intron) of the *Chd2* gene.**

Materials and Methods

Reagents and Antibodies

Cell culture reagents were obtained from the following sources: fetal calf serum and Antibiotic-Antimycotic (Gibco, Invitrogen),5-Flurouracil & Doxorubicin (Sigma). Antibodies were obtained from the following sources: CHD2 (Dr. Venkatachalam's lab, University of Tennessee); p53-K382 Acetylated, H3K9 Acetlated, H4K8 Acetylated (Cell Signaling Technology, Danvers, MA, USA); Beta-Tubulin (Gentex, Zeeland, MI); Puma (Calbiochem, Gibbstown, NJ); p21, p53 (FL-393X) (Santa Cruz Biotechnology, Santa Cruz, CA).

Western Blot

HCT116 cells were scraped in phosphate-buffered saline (PBS), pellet at 1'400 RPM, and lysed in BE lysis buffer plus 1mM PMSF (40mM NaCl, 20mM Tris-HCl pH 7.8, 2mM MgCl2, 0.5% NP40, Benzonase (Novagen, 250U/mL)). Cell lysates were mixed with reducing Laemmli sample buffer, separated on 8% or 12% SDS–polyacrylamide gels and transferred onto PVDF membranes (Thermo scientific). Membranes were blocked

at room temperature for 2 hours in TBS-Tween (100mM Tris-HCl, 200mM NaCl, 0.1% Tween 20) with 5% milk, and incubated with primary antibody overnight at 4^{o}C or at room temperature for 2 hours. After washing with TBS-T, membranes were incubated with horse radish peroxidase-conjugated secondary antibody for 1 hour at room temperature in TBS-T containing 2% milk. Signal was visualized with the Super Signal chemiluminescent reagent (Thermo scientific).

RNA isolation and RT-PCR protocol

Total RNA was extracted using Trizol according to the manufacturer's instructions (Invitrogen) from adult thymus (6 weeks of age). First strand cDNA synthesis was performed with 2µg of total RNA (pretreated with RNAse free DNAse) with random hexamers and M-MLV reverse transcriptase (Promega) for 1 hour at 42^{o}C followed by inactivation of the reverse transcriptase at 70^{0}C for 15 minutes. A similar amount (2µg) of total RNA (pre-treated with RNAse free DNAse) was subjected to the above mentioned conditions in the absence of M-MLV reverse transcriptase. For PCR assays, 2µl of the reaction mixture (from a total of 40 µl) obtained from the first strand cDNA synthesis reaction was used. The PCR conditions were 94^{0}C for 2 min, followed by 30 cycles at 94^{0}C for 30 sec, 58^{0}C for 30 sec and 72^{0}C for 10 sec. PCR products were resolved by agarose gel electrophoresis and stained with ethidium bromide. For Chromatin Immunoprecipitation isolated DNA, 2.5µL of DNA eluant were subjected to primer specific amplification, and the PCR conditions were 94^{0}C for 2 min, followed by 40 cycles at 94^{0}C for 30 sec, 59^{0}C for 30 sec and 72^{0}C for 5 sec.

Quantitative PCR

The real-time PCRs were performed on a Bio-Rad iQ5 system using TaqMan® Gene Expression Assays (Applied Biosystems, Foster City, CA). Gene-specific primers were obtained from Applied Biosystems (Mm00519268_m1 for Puma, Mm01197698_m1for Beta-Glucuronidase). Puma expression values (Ct-values) were normalized to the mean expression value of the housekeeping genes Beta-Glucuronidase to obtain relative expression levels using the $2^{-\Delta\Delta Ct}$. All reactions were performed in triplicate.

Chromatin Immunoprecipitation

HCT116 cells were scraped and crosslinked with 1% formealdehyde (Fisher) in ChIP Buffer (60mM KCl, 15mM NaCl, 4mM $MgCl_2$, 15mM HEPES pH 7.6, 0.5% Triton X-100, 10mM Sodium Butyrate) for 20 minutes at room temperature with rotation. For Mouse thymocytes: 6 weeks old mice were sacrificed, thymus was isolated and cells were strained using a 0.45µM strainer, cells were then crosslinked with 1% Formealdehyde in ChIP buffer. Samples were then neutralized with 300mM Glycine for 10 minutes, washed twice with ChIP Buffer, and lysed with ChIP LB (140mM NaCl, 15mM HEPES pH7.6, 1mM EDTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 10mM Sodium Butyrate, Protease Inhibitors (Roche), 5µM Trichostatin A) for 10 minutes. Samples were sonicated to obtain 500-1000BP DNA fragments and then pre-cleared with Protein-G agarose beads overnight at 4^oC. Pre-cleared samples (1.5mg) were then incubated with primary antibody (4µg) for 6 hours at 4^oC and then fresh Protein G agarose beads were added to samples and rotated overnight at 4^oC. The Immunoprecipitates were washed successively with low salt buffer (20mM Tris-HCl [pH

7.8], 150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), high salt buffer (20mM Tris-HCl [pH 7.8], 500mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), LiCl washing buffer (10mM Tris-HCl [pH 7.8], 250mM LiCl, 1% NP40, 1% Sodium Deoxycholate, 1mM EDTA), and TE Buffer. DNA-protein complex were eluted with IP elution buffer (1% SDS, 0.1M $NaHCO_3$), reversed crosslink by adding 200mM NaCl and incubating eluant at 65^{o}C. Protein and RNA was digested by adding Proteinase K and RNase A, respectively and DNA was extracted using GeneJET kit (Fermentas, Glen Burnie, MD) and eluted with equal volume of nuclease-free water.

RNA Interference Transfection

RNAi oligonucleotides were obtained from IDT DNA (F-LUC-Si, CHD2si-2) or Dharmacon (CHD2si, used in most experiments) and resuspended in nuclease-free water to a concentration of 40µM. HCT116 cells were plated in 60mm-well plates in DMEM media containing 10% fetal bovine serum to give 30–50% confluence. Transfection of the RNAi oligonucleotides was performed using Trifectin (IDT-DNA) to result in a final RNA concentration of 40 nm. The cells were harvested at different time points and lysed in BE lysis buffer (above mentioned) for Western blot analysis. siRNA sequences were:

F-Luc si: 5'-rGrCrA rUrArU rCrArA rArGrC rArCrA rUrCrA rGrGrU rCrCA C-3'
5'-rGrUrG rGrArC rCrUrG rArUrG rUrGrC rUrUrU rGrArU rArUrG rCrCrU-3'
CHD2si: 5' rUrUrA rGrArC rArUrU rGrGrG rArUrC rUrUrA rGrGrA rUrUrC rUrUrC 3'
5' rArGrA rArUrC rCrUrA rArGrA rUrCrC rCrArA rUrGrU rCrUA A -3'

CHD2si-2: 5'-rGA UAG CUG AUG UGA AGA AGA UGT G-3'

Results

4.1. Characterization of N-terminal Chd2 deficient mice.

The gene-trap in the Chd2 C-terminal model inserts in intron 27, thus retaining a substantial portion of the Chd2 protein (translation of the first 1198/1828 amino acids). This truncated protein retains all common CHD domains fused to the β-galactosidase-neomycin gene. This fusion protein may have a dominant negative effect, a gain of function effect, or sequester the function of binding proteins necessary to mediate Chd2 activity. Therefore, we resorted to the Chd2 N-terminal mouse model (Here forth, the only model studied) as a better alternative to study Chd2's function in mammalian development, DNA damage response, and tumorogenesis. We analyzed the post-natal lethality of *Chd2* mutant mice. Analysis of Heterozygous matings at weaning (day 21) shows a 27% survival of the null offspring, which indicates lethality of the *Chd2* mutant mice. The mutant offspring that survived beyond the perinatal stage did not show any overt developmental abnormalities. However, after 5–12 months of age, the heterozygous mice began to exhibit weight loss, lordokyphosis (hunch-back spine) and loss of vitality. Survival analysis showed drastic reduction in the lifespan of the *Chd2* heterozygous mice with a median lifespan of 59 weeks (Figure 18)

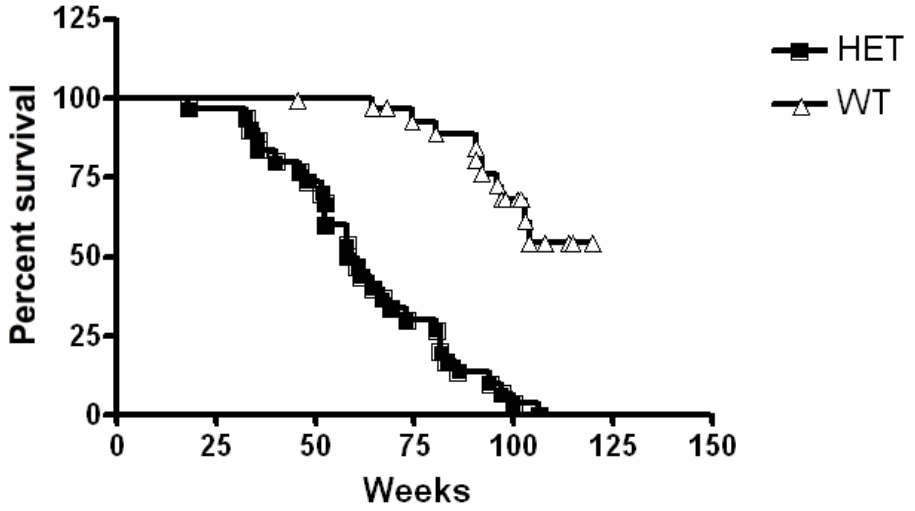

Figure 18. Chd2 deficiency leads to reduced lifespan in mice. Kaplan-Meier survival curves of Chd2 mutant and wild-type littermates **(n=30). The percentages of survival are plotted as a function of age in weeks. Animals were monitored for tumors, morbidity, or spontaneous death over a period of 120 weeks. Of the 30 animals analyzed for each group, all the heterozygous mutants have died in comparison to 10 for the wild-type controls. The median lifespan of the Chd2 +/- mice was 59.35 weeks in comparison to 91.3 weeks in the wild type littermate controls. All mice were of mixed inbred C57BL/6X129/Sv background.**

Histological examination of organs harvested from morbid mice showed that a majority of the mice were succumbing to splenic or thymic lymphomas or lymphoid hyperplasias with a fraction of adenomas and sarcomas (Table). The earliest incidence of lymphomas in the mutant mice was at 26 weeks of age. 53% of the heterozygous mutants (16/30) exhibited spontaneous tumor formation. In comparison, only 25% of wild-type mice were diagnosed with a tumor during the analysis period (4/12). Wild-type mice also develop lymphomas as a function of age, and such tumors account for about 5–20% incidence as reported for mice of various genetic backgrounds by others. Histological examination of organs also revealed Glomerulo-nephoropathies and inflammation of the heart or arteries in a minority of the animals as well as Extra Medullary Hematopoiesis (EMH) in 57% of the animals examined.

Table 3. Distribution of pathological conditions in Chd2 deficient (N-terminal mutant) mice. **Tissues from a total of 30 mice (Chd2+/-) were analyzed to determine the reasons for morbidity. Hearts were examined for 8 mice. To avoid over-estimation of lymphoid hyperplasias, animals exhibiting lymphomas as well as lymphoid hyperplasias (in other organs) were categorized under lymphomas. *15 out of 17 animals diagnosed with EMH exhibited either lymphoid hyperplasia or lymphoma. ** Both the animals exhibiting nephropathy were diagnosed with either lymphoid hyperplasia or lymphoma. ***The animal diagnosed with heart inflammation exhibited either lymphoid hyperplasia. ****Three animals were diagnosed with bronchoalveolar adenoma and two were diagnosed with sarcomas.**

Lymphoma	53% (16 of 30)
Lymphoid hyperplasia	36.6% (11 of 30)
Extra Medullary Hematopoiesis* (EMH)	56.6% (17 of 30)
Glomerulo-nephropathy**	10% (2 of 19)
Inflammation of heart/artery***	11% (1 of 9)
Other cancers****	16.6% (5 of 30)

Since histological analysis of Chd2 heterozygous mice revealed thymic lymphomas and thymic hyperplasias, we wanted to test whether Chd2 plays its greatest role specifically in the thymus of mice. For that, we crossed Chd2 heterozygous males with *Lck*-Cre transgenic founder female mice expressing Cre recombinase exclusively in thymus. This leads to excision of the splice-acceptor site of the gene trap, thus bypassing it and leading to wildtype expression of Chd2 in the thymus only (termed LY-Cre). Chd2 expression in the thymus restored normal life span of heterozygous mice to that of wild type mice. Median lifespan for LY-Cre heterozygous mice was 102 weeks

(n=8) compared to a median lifespan of 59 weeks for Chd2 heterozygous mice (Figure 19).

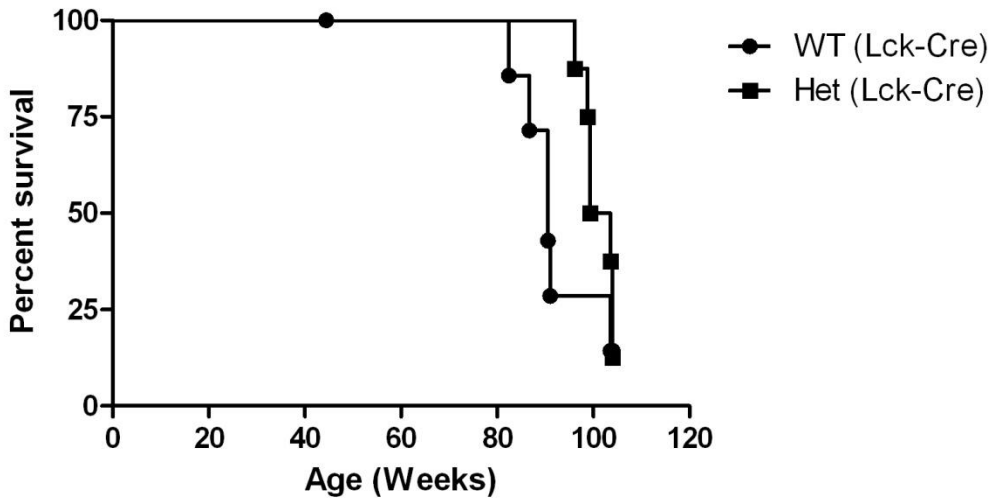

Figure 19. Restoration of CHD2 expression in thymus restores normal life span of CHD2 heterozygous mice. **Kaplan-Meier survival curves of Chd2 mutant and wild-type littermates (n=8). The percentages of survival are plotted as a function of age in weeks. Animals were monitored for tumors, morbidity, or spontaneous death over a period of 120 weeks. The median lifespan of the Chd2 +/- mice was 59.35 weeks in comparison to 102 weeks in the LY Cre Het mice. All mice were of mixed inbred C57BL/6X129/Sv background.**

To determine the mechanistic role of Chd2 in tumorogenesis, we tested the fidelity of the cell cycle progression in wildtype and Chd2 mutant MEFs. Incorporation of the thymidine analog, Bromodeoxyuridine (BrdU), indicates cell proliferation. Cells that are treated with X-ray radiation arrest at the G1/S and G2/M boundaries and therefore don't incorporate BrdU. Flow cytometry analysis indicates that Chd2+/m and Chd2m/m keep incorporating BrdU 16 hours after treatment with X-ray radiation (Figure 20). Furthermore, Chd2+/m and Chd2m/m do not arrest at the G2/M boundary either. This is indicative of a cell cycle arrest deficiency which can be caused by the deregulation of many gene including the cyclin-dependent kinase inhibitor 1 (p21/WAF1).

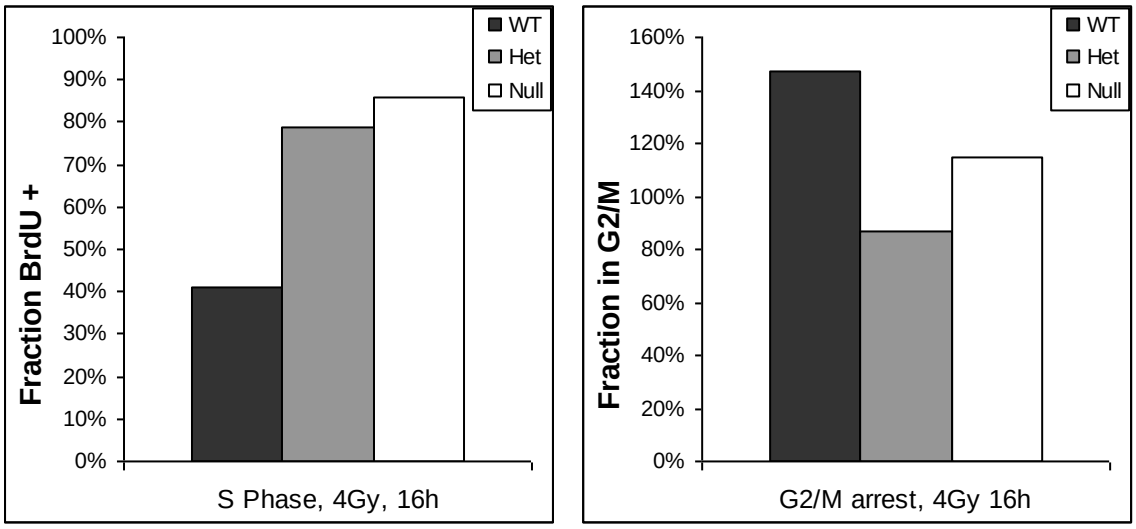

Figure 20. Effects of Chd2 on Cell-cycle Arrest. **DNA damage induces the cell-cycle inhibitor p21 causing arrest in the G1/S and G2/M phase. Wildtype cells exhibit cell-division arrest after treatment with 4Gy X-Ray. Chd2 heterozygous and nullizygous MEFs exhibit a diminished G2/M arrest and increased S phase synthesis.**

4.2. Chd2 deficient thymocytes exhibit deficient Puma response.

Since many of the Chd2 heterozygous mice were succumbing to lymphomas and lymphoid hyperplasias and hematopoietic deficiencies, we tested the ability of Chd2 mutant animals to undergo apoptosis. We picked the p53-upregulated modulator of apoptosis (Puma) as a candidate gene to test its induction in wildtype and Chd2 mutant thymocytes. RT-PCR analysis showed a severe deficiency in Chd2+/m and Chd2m/m thymocytes' ability to induce Puma after DNA damage (Figure 21). Q-PCR analysis further confirmed our initial findings where Puma was 27 folds higher in treated wildtype thymocytes over the untreated control (Figure 22). Meanwhile, Puma was only 2.5 and 1.3 folds higher in treated Chd2+/m and Chd2m/m thymocytes over the untreated control respectively.

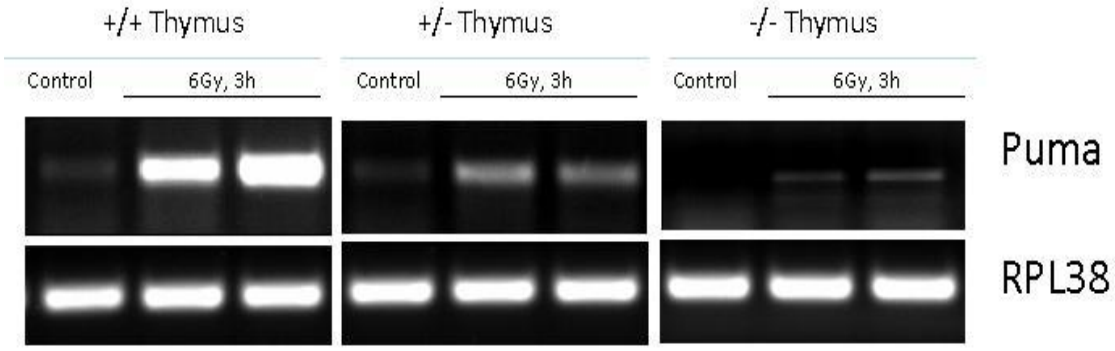

Figure 21. Expression of PUMA in Chd2 Thymocytes. **Chd2 thymus were isolated and total RNA was harvested 3H after 6Gy X-ray treatment. RT-PCR analysis show no induction of the p53-upregulated modulator of apoptosis (PUMA) in Chd2 deficient mice.**

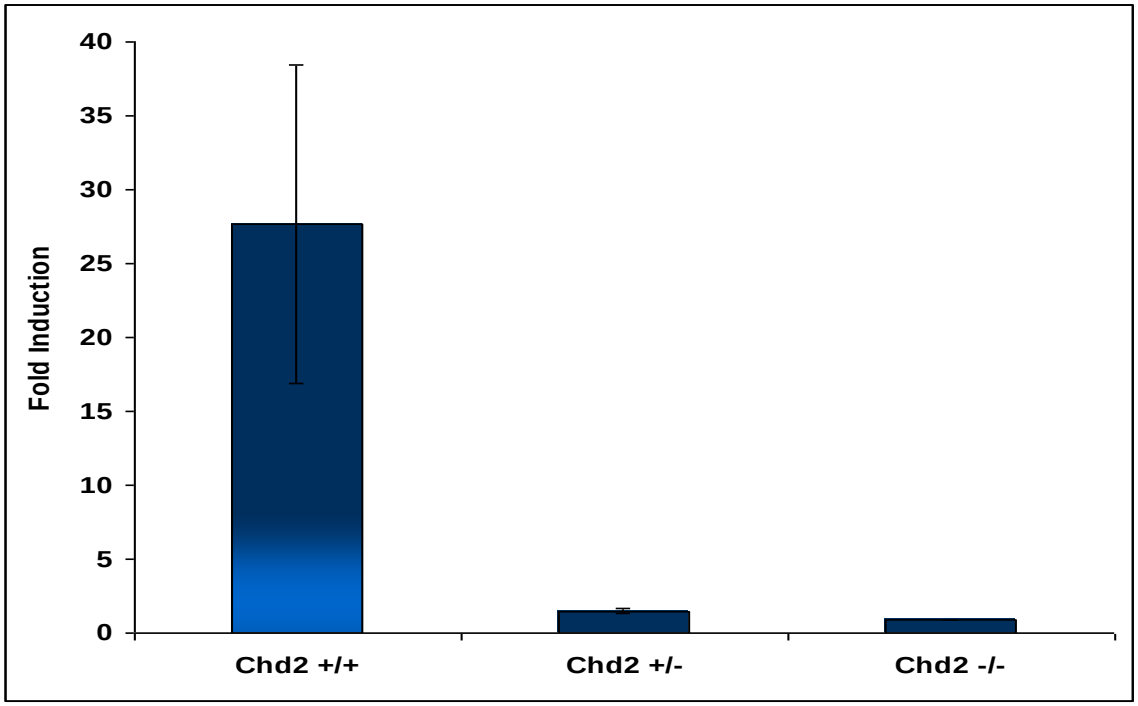

Figure 22. Real-Time analysis of PUMA levels in Chd2 thymocytes. **Wildtype Thymocytes show a 27 fold induction of Puma three hours after treatment with 6Gy X-ray. Chd2 heterozygous and nullisomic thymocytes exhibit a lack of induction after DNA damage. B-Glucuronidase (GusB) was used as an internal control.**

4.3. CHD2 deficient cells exhibit reduced PUMA induction

RNA interference was effective in reducing CHD2 protein levels in HCT116 cells. CHD2 is not induced after treatment of cells with the thymidylate synthase inhibitor 5-Flurouracil (Figure 23). Chd2 deficiency did not have an effect on the lysine 382 acetylation of p53 after 5FU treatment. However, Puma levels were greatly reduced in CHD2 deficient cells compared to controls (Figure 23). Additionally, p21 levels were also reduced in CHD2 deficient cells, indicating an insufficient induction of these p53 regulated genes. Beta-tubulin indicates equal loading control of cell lysates (Figure 23). Since CHD2 appears to play a role in transcriptional upregulation of Puma, we wanted to test whether CHD2 deficiency affects histone tail markers associated with transcriptional activation. Acetylation of lysine 9 of Histone 3 (H3K9Ac) and acetylation of lysine 8 of Histone 4 (H4K8Ac) are well known markers of active transcription. Global acetylation of these two histones was reduced in untreated CHD2-deficient cells compared to wild type cells (Figure 24). 5FU treated CHD2-deficient cells showed a greater reduction of each of these marks compared to controls, thus further indicating a deficiency in histone acetylation in CHD2-deficient cells. However, although PUMA was not being induced, apoptosis was still proficient as indicated by cleavage of poly ADP ribose polymerase (Figure 24). Same results were obtained using a second set of CHD2 siRNA, further supporting our findings.

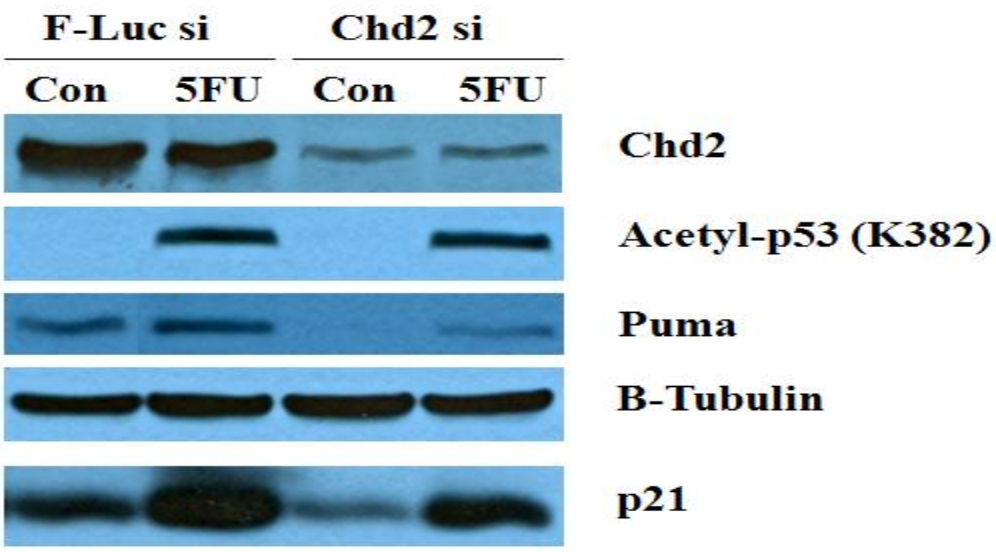

Figure 23. Western blot analysis of CHD2 deficient DNA-damage response in HCT116 cells. **RNAi was used to attenuate CHD2 protein levels. P53 acetylation at lysine 382 was not affected by CHD2 knockdown, yet levels of the p53 target gene Puma were reduced in CHD2-deficient cells compared to control (F-Luc). P21, another p53 target, was also reduced in CHD2-deficient cells, but to a lesser extent. Beta-Tubulin shows equal loading.**

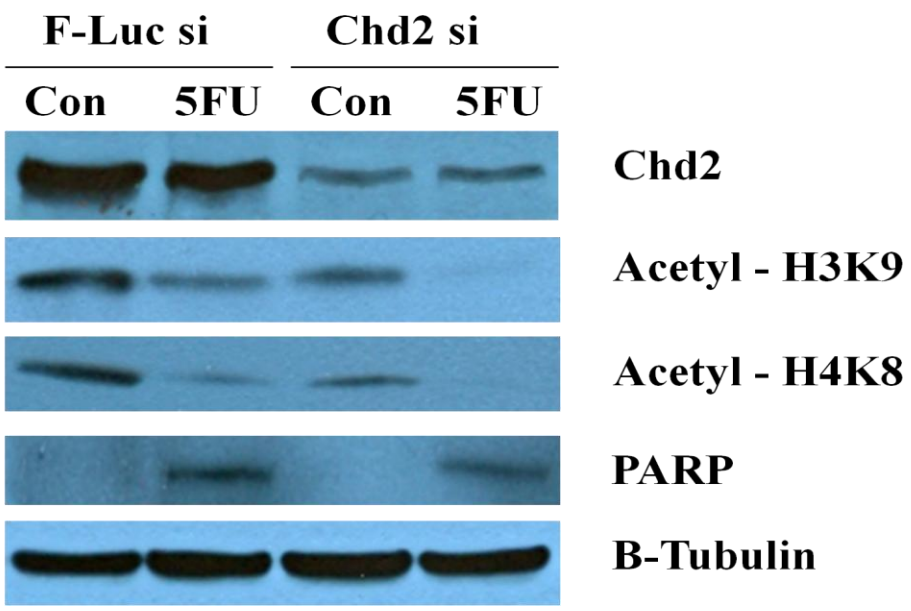

Figure 24. Western blot analysis of CHD2 deficient DNA-damage response in HCT116 cells**. RNAi was used to attenuate CHD2 protein levels. Levels of histone H3 (Lys9 acetylated) and histone H4 (Lys8 acetylated) were reduced in CHD2-deficient cells compared to control (F-Luc). The apoptosis associated protein PARP levels remained similar in 5FU treated cells in both control and CHD2-deficient cells. Beta-Tubulin shows loading control**

4.4. CHD2 localizes to *Puma* in mouse thymocytes.

Based on these findings, we wanted to determine whether Chd2 plays a role in the induction of Puma. We utilized chromatin immunoprecipitation (ChIP) to determine whether CHD2 binds to the *Puma* promoter and the p53 response element there. Our results initially indicated that Chd2 binds to the Puma promoter in wildtype thymocytes (primers designed to amplify p53 RE located in Intron 1 of the Puma promoter) specifically after DNA damage (Figure 25). Further experiments showed that Chd2 is enriched at the Puma promoter after DNA damage. Moreover, we tested a region 1.5KB upstream of the Puma promoter as an additional control and were able to amplify that region, which could indicate the spreading of Chd2 throughout the region (Figure 26). However, CHD2 didn't localize to an unrelated gene HGPRT. We then tested whether

Chd2 plays a role in the recruitment of p53 to its response element on the Puma promoter. DNA analysis indicates that p53 is recruited to its cognate region sufficiently in wildtype and Chd2 mutant thymocytes; furthermore, p53 is enriched at its Puma RE after DNA damage, which is consistent with previous studies (Figure 27).

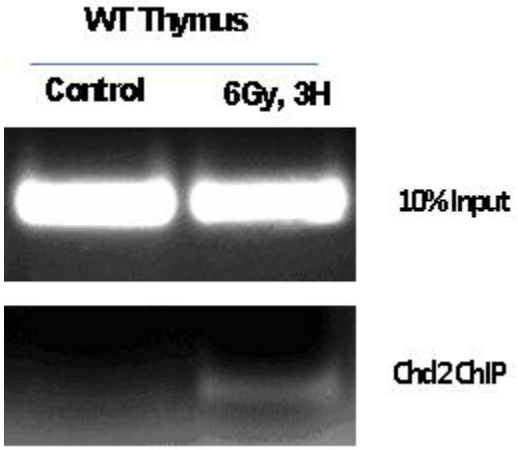

Figure 25. Chromatin immunoprecipitation (ChIP) on wildtype thymocytes. **DNA analysis on Chd2 IP shows localization to the Puma promoter after DNA damage. 10% Input shows total lysate control.**

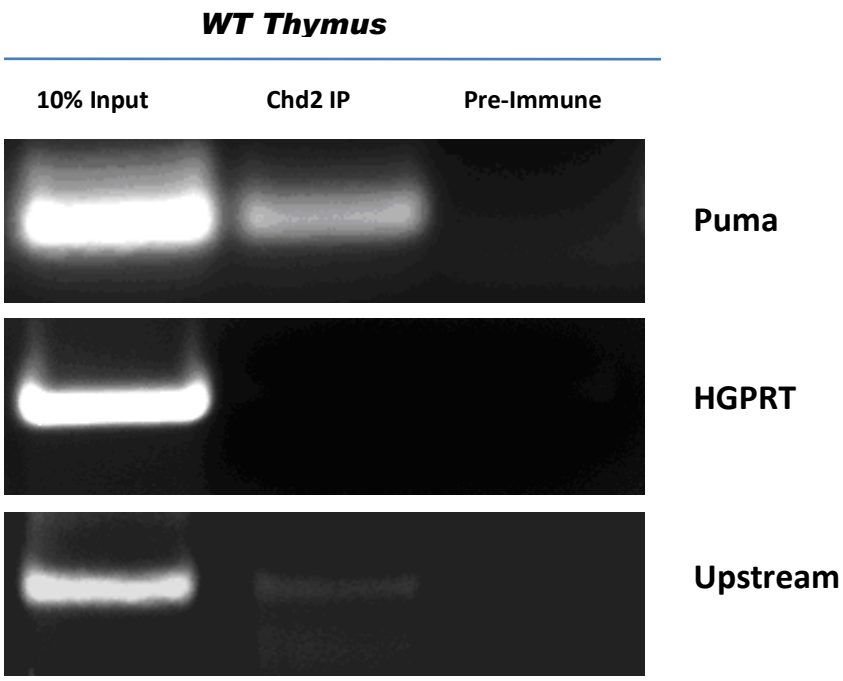

Figure 26. Chromatin immunoprecipitation (ChIP) on wildtype thymocytes. **DNA analysis on Chd2 IP shows localization to the Puma promoter. Lack of amplification of the HGPRT region indicates specificity of Chd2 IP. 10% Input shows total lysate control. Upstream indicates the 1.5KB region upstream of the Puma promoter.**

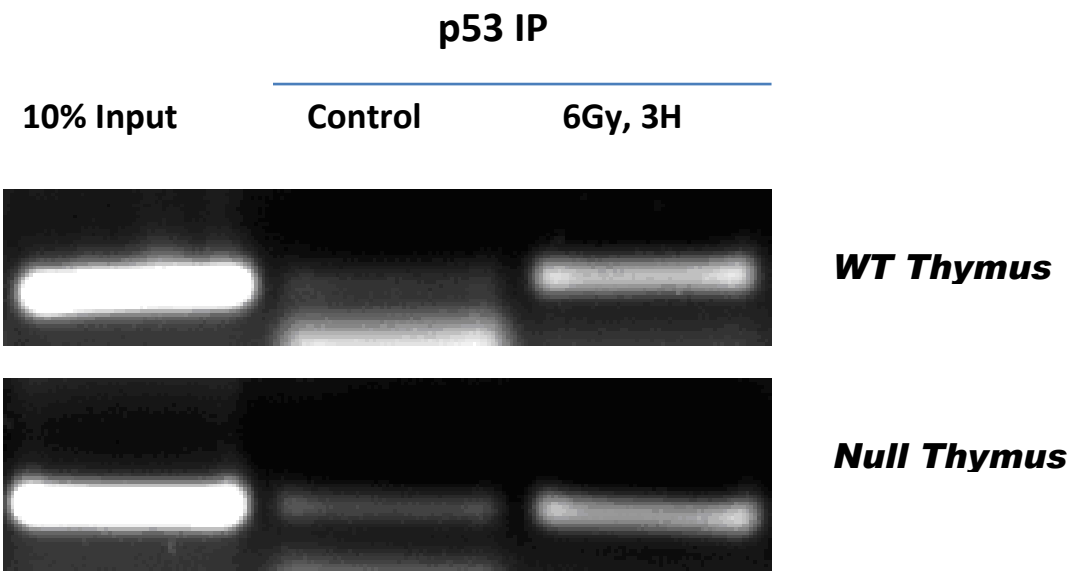

Figure 27. Chromatin Immunoprecipitation analysis on Chd2 Thymocytes. **DNA analysis on Chd2 IP shows localization to the PUMA promoter after DNA damage. p53 IP on WT and Chd2 nullisomic thymocytes shows no role for Chd2 in localizing p53 to the PUMA promoter. 10% Input shows total lysate control.**

4.5. CHD2 localizes to Puma in p53-dependent manner in HCT116 cells.

In HCT116 wild type cells, CHD2 localizes to the p53 response element of the *Puma* promoter before DNA damage, but is enriched there after DNA damage (Figure 28Figure 29). Moreover, Chd2 localizes to the p53 response element of the Puma promoter in the thymocytes of wildtype mice. CHD2 localizes to other regions within the *Puma* locus including Intron 3, which has been reported to be enriched in trimethylated H3K9 histones and occupied by the insulator protein CTCF (234). Additionally CHD2 is slightly enriched in the 5KB region upstream of the PUMA promoter after DNA damage, indicating its ability to spread beyond the promoter (Figure 30). However, CHD2 doesn't localize to the Beta-Glucuronidase promoter before or after DNA damage (Figure 31). The tumor suppressor protein p53 also is enriched at the PUMA promoter, which has

been established before. This enrichment of CHD2 in a manner similar to p53 at the PUMA promoter led us to test whether p53 and CHD2 are co-dependent on each other to be recruited to the PUMA promoter. For, we tested the localization of CHD2 to PUMA promoter in HCT116 Null cells which lack p53. The localization of CHD2 to the Puma promoter was reduced as indicated by the band intensity (Figure 29). Additionally, the localization of CHD2 to the 5KB region upstream of PUMA promoter was greatly increased after DNA damage in the p53 deficient cells. Moreover, CHD2 appears to localize to the Beta-Glucuronidase promoter after DNA damage in p53 deficient cells, indicating a possible role for p53 in guiding the localization of CHD2 (Figure 31). We then tested whether CHD2 deficiency affects p53 localization to its cognate response element in the PUMA promoter. P53 was enriched after DNA damage at the PUMA promoter in CHD2 deficient cells but at lower intensity compared to wild type cells (Figure 29). H3K9Ac markers enrichment at the *Puma* promoter appears to remain unchanged in wild-type and CHD2-deficient cells (Figure 32). Although we have shown that CHD2 deficiency reduces global acetylation, especially after DNA damage, it's possible that *Puma* acetylation is not affected in this case, or that we acetylation of Puma remains constant, while only CTCF mediates transcription of exon 3 to form the functional mRNA (234). Three biological replicates were done for each ChIP.

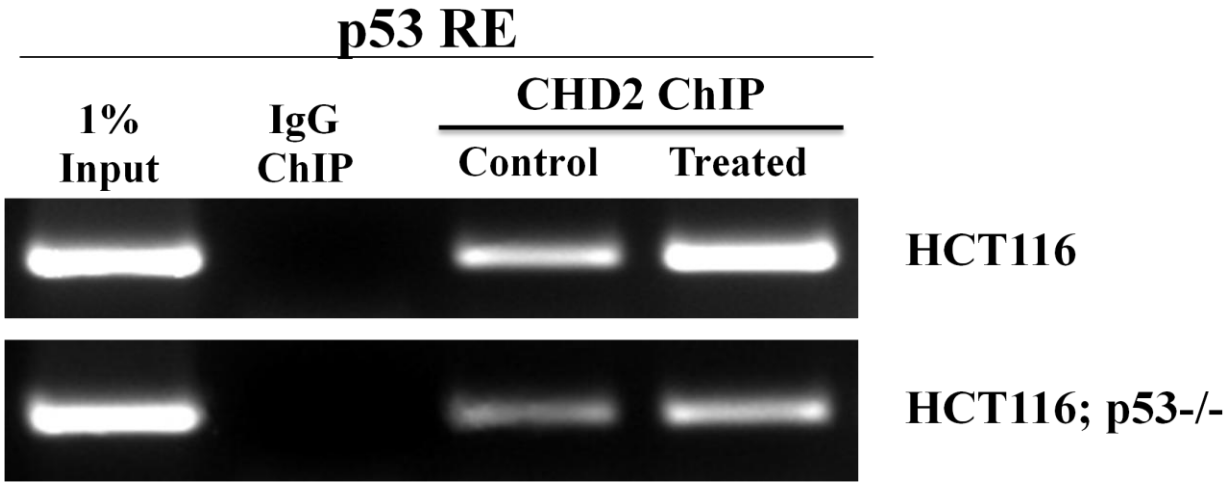

Figure 28. Chromatin Immunoprecipitation analysis of CHD2 localization to p53 response element in Puma. **DNA analysis on Chd2 IP shows localization and enrichment to the *PUMA* promoter after DNA damage. CHD2 localization is reduced in HCT116 p53-/- cells as shown by reduced intensity of bands. 1% Input shows equal total lysate control, and IgG ChIP shows antibody specificity.**

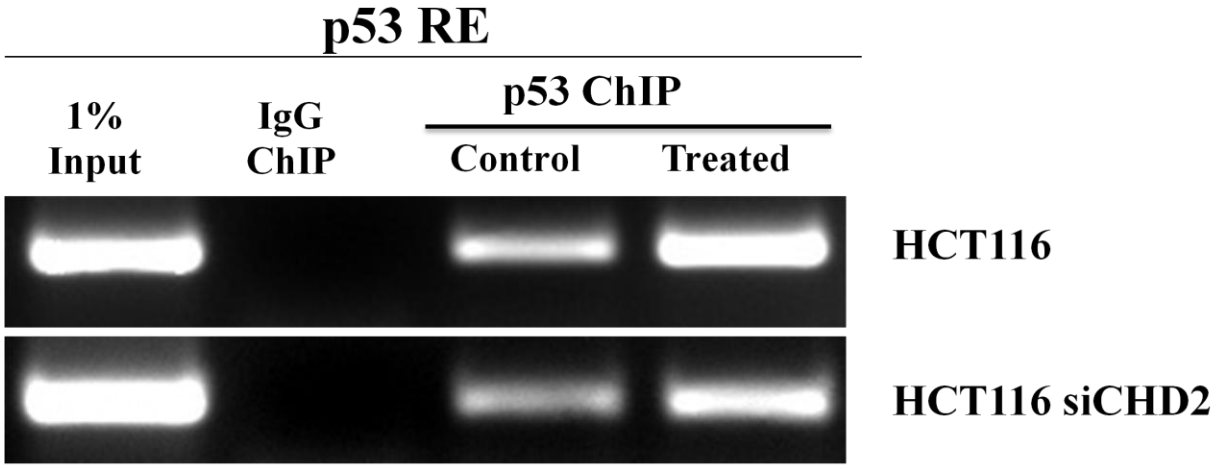

Figure 29. ChIP analysis of p53 localization to its response element in Puma. **DNA analysis on p53 IP shows localization and enrichment to the *PUMA* promoter after DNA damage. P53 localization is reduced in HCT116 treated with CHD2-specific siRNA cells as shown by reduced intensity of bands. 1% Input shows equal total lysate control, and IgG ChIP shows antibody specificity.**

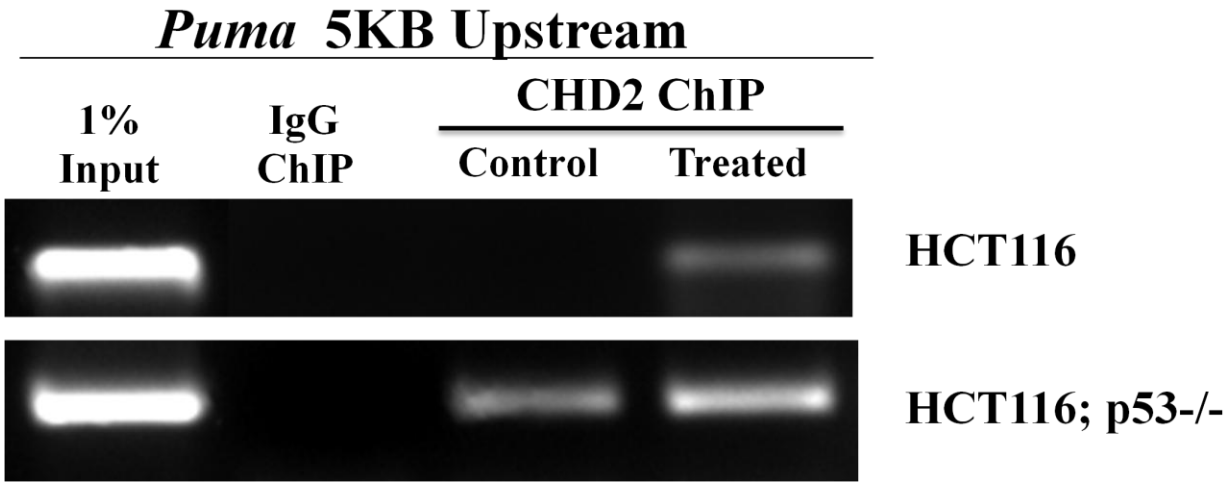

Figure 30. ChIP analysis of CHD2 localization to a 5KB region upstream of Puma promoter. **DNA analysis on Chd2 IP shows localization of CHD2 to the upstream region of the *PUMA* promoter after DNA damage. CHD2 localization is greater in HCT116 p53-/- cells as shown by increased intensity of bands, possibly indicating a role for p53 in CHD2 localization. 1% Input shows equal total lysate control, and IgG ChIP shows antibody specificity.**

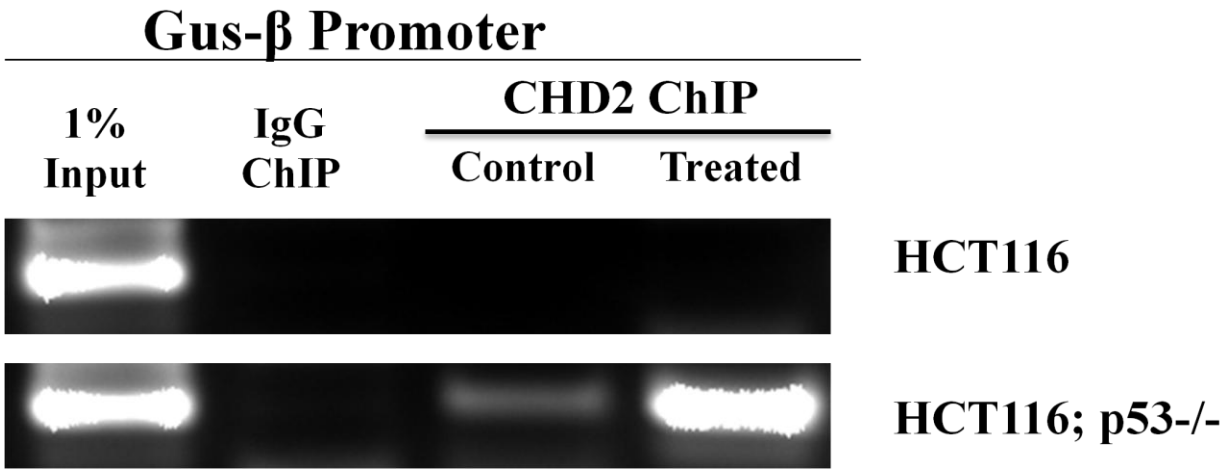

Figure 31. ChIP analysis of CHD2 localization to an unrelated Beta-Glucuronidase promoter. **DNA analysis on Chd2 IP shows no localization of CHD2 to the Gus-B promoter in HCT116 cells. CHD2 localization to this region is seen in HCT116 p53-/- cells as shown by presence of bands, possibly indicating a role for p53 in CHD2 localization. 1% Input shows equal total lysate control, and IgG ChIP shows antibody specificity.**

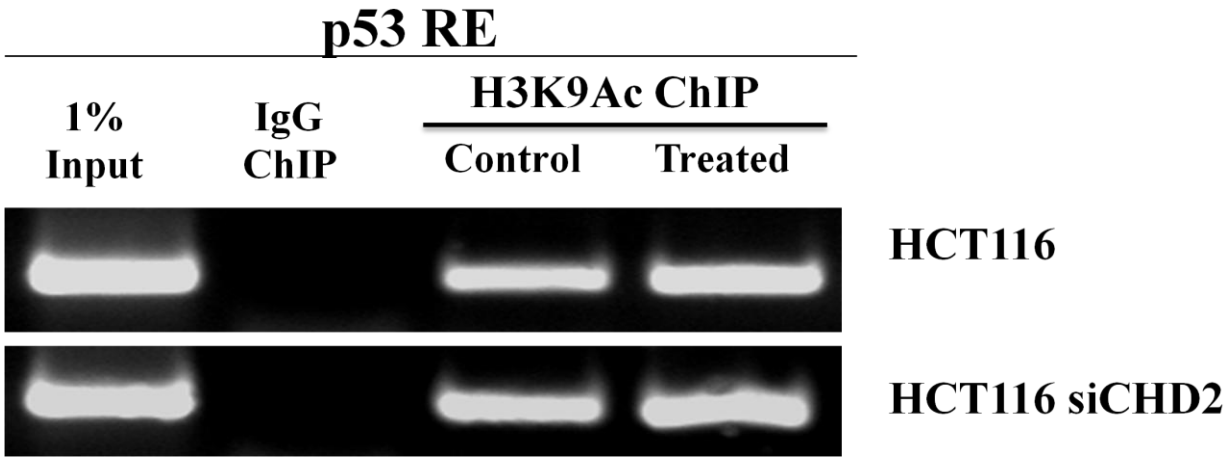

Figure 32. ChIP analysis of Acetylated histone H3 (Lys9) localization to Puma promoter. DNA analysis on H3K9Ac IP shows similar levels of localization before and after DNA damage at the Puma promoter and in wild type and CHD2-deficient cells. 1% Input shows equal total lysate control, and IgG ChIP shows antibody specificity.

Discussion

Our results suggest an indispensible role for CHD2 in preventing tumorogenesis and inducing Puma after DNA damage. Previous work in our laboratory had shown that Chd2 mutant mice succumb to lymphomas and have a deficient DNA damage response (126). However, our previous mouse model (C-terminal) expresses the first 27 exons of Chd2, which encode for the chromodomain, SNF2-related helicase domain, C-terminal helicase domain, fused to the neomycin-β-gal gene. This can potentially create a dominant negative protein that can interact with wild type Chd2. Our current model (N-terminal) possesses the gene-trap after exon 1 (20 amino acids) of *Chd2*, thus creating a superior hypomorphic model to study the role of Chd2. Similar to the C-terminal mouse model, N-terminal mutant mice also exhibit reduced lifespan and spontaneous lymphomas. However, the spectrum of tumors expands to include sarcomas and adenomas (16.6% of mice), which was absent from the C-terminal model. Additionally, restoration of Chd2 expression in the thymus of mutant mice restores their lifespan to that of wild type littermates, thus suggesting that Chd2 plays its greatest role in the thymus. We suspected Chd2 to play a role in DNA damage repair pathways and wanted to test the effect of Chd2 deficiency on the induction of the cell-cycle progression gene (p21) and apoptosis related genes (Puma, Noxa, Bax). Indeed, Puma induction both in mouse thymus was reduced after X-ray irradiation in Chd2 deficient cells, indicating a role for Chd2 as a transcriptional regulator of Puma induction. Because of the time consuming nature of animal work and lack of reliable mouse specific antibodies, we wanted to test the effect of CHD2 deficiency on PUMA induction in HCT116 cells. HCT116 cells are epithelial colon cancer cells that are capable of undergoing apoptosis when chemically stressed. Another advantage to using HCT116 cells is they are

available in both wild type and p53 null lines. Therefore, we can test the effect of p53 on a proteins function while maintaining every other pathway in the cell line. We tested multiple chemical stressors in HCT116 cells to check the best induction of PUMA, and concluded that 5-flurouracil is the best compound to induce PUMA. CHD2 deficient cells exhibited reduced PUMA induction after treatment with 5FU. This pattern is similar to the reduced Puma induction in Chd2 mutant mouse thymocytes. Interestingly, it appears that apoptosis is not affected by the lack of PUMA induction as PARP cleavage is similar in both control and CHD2-deficient cells treated with 5-FU. Since apoptosis is induced by p53-dependent and p53-independent pathways (TNF-α, Fas- pathways), an alternative pathway can possibly trigger apoptosis in the absence of PUMA induction after DNA damage.

P21, another p53 target, was induced in CHD2 deficient cells after DNA damage but a lower level than control cells, possibly indicating a role for CHD2 in the induction of p53 target genes. In previous work, CHD8 was shown to affect the acetylation of p53 and the induction of its downstream targets; therefore, we wanted to test whether CHD2 deficiency affects p53 acetylation to possibly designate a mechanism of action for CHD2 (229). Acetylation of p53 at lysine 382 of its C-terminal regulatory region has been shown to enhance p53's activity as a transcription factor, allowing p53 to bind its RE with a greater affinity. CHD2 deficiency did not affect p53 acetylation at Lysine 382 after 5FU treatment, thus indicating a different mechanism of action for CHD2 on transcription activation. We next looked at CHD2's affect on Histone acetylation because CHD1 has been shown to be a component of the SAGA histone acetyltransferase complex (212). H3K9 and H4K8 acetylation are well known

transcriptional activation markers (235). Histone acetylation at majority of genome is reduced after DNA damage to allow the cell to resolve the damaged DNA before continuing through the cell cycle. CHD2 deficient cells exhibit global reduction in H3K9 and H4K8 acetylation when comparing untreated cells (No DNA damage). After DNA damage, H3K9 and H4K8 acetylation is reduced in control cells but is almost undetectable in CHD2 deficient cells. This indicates a role for CHD2 in histone acetylation and possibly preludes a mechanism for the lack of PUMA induction after DNA damage.

We further verified a role for CHD2 in PUMA induction by checking CHD2 localization to the *Puma* locus using Chromatin Immunoprecipitation. DNA eluted from CHD2 ChIP was amplified using specific primers to test the p53 response element of the *Puma* promoter, a 5KB region upstream of this response element, a region of Intron 3 of *Puma* locus, and an unrelated Beta-Glucuronidase promoter (GusB). The 5KB region and the Intron 3 test whether CHD2 spreads beyond the promoter of Puma, while the GusB promoter works as a control of a housekeeping gene. CHD2 is enriched at the p53 response element after DNA damage, but this enrichment is slightly reduced in p53 deficient HCT116 cells. CHD2 spreads to the 5KB region upstream of Puma promoter and to Intron 3 after DNA damage, showing that it spreads all along a locus to remodel chromatin. CHD2 is present in higher levels at the 5KB upstream region and is reduced at the Puma promoter in p53 deficient HCT116 cells after DNA damage. This points to a role for p53 in guiding CHD2 binding on the genome or at least laying the groundwork for CHD2 binding through its chromodomain or its A+T hook DNA binding domain. Also, CHD2 binds to the GusB promoter in p53 deficient cells after DNA damage; this is not

seen in wild type cells and further leads to a role for p53 in guiding CHD2 binding. CHD2 deficiency also slightly reduces p53 localization to its response element after DNA damage. This could be due to reduced global acetylation in CHD2 deficient cells, which makes the chromatin more compact and less accessible to transcription factors. We also tested whether CHD2 binds to p53 RE at the *CDKN1A* locus (p21 locus) and indeed CHD2 is enriched there after DNA damage (Data not shown).

Refernces

1. Bolker, J. A. 1995. Model systems in developmental biology. *Bioessays* 17:451-455.
2. Washington, N. L., M. A. Haendel, C. J. Mungall, M. Ashburner, M. Westerfield, and S. E. Lewis. 2009. Linking human diseases to animal models using ontology-based phenotype annotation. *PLoS Biol* 7:e1000247.
3. Lee, P. S., and K. H. Lee. 2003. Escherichia coli--a model system that benefits from and contributes to the evolution of proteomics. *Biotechnol Bioeng* 84:801-814.
4. Goeddel, D. V., D. G. Kleid, F. Bolivar, H. L. Heyneker, D. G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A. D. Riggs. 1979. Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proc Natl Acad Sci U S A* 76:106-110.
5. Meyer, M., and J. Vilardell. 2009. The quest for a message: budding yeast, a model organism to study the control of pre-mRNA splicing. *Brief Funct Genomic Proteomic* 8:60-67.
6. Molina, M., V. J. Cid, and H. Martin. Fine regulation of Saccharomyces cerevisiae MAPK pathways by post-translational modifications. *Yeast* 27:503-511.
7. Scannell, D. R., G. Butler, and K. H. Wolfe. 2007. Yeast genome evolution--the origin of the species. *Yeast* 24:929-942.
8. Brenner, S. 1974. The genetics of Caenorhabditis elegans. *Genetics* 77:71-94.
9. Fischer, S. E. Small RNA-mediated gene silencing pathways in C. elegans. *Int J Biochem Cell Biol* 42:1306-1315.
10. Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* 391:806-811.
11. Sofer, W., and L. Tompkins. 1994. Drosophila genetics in the classroom. *Genetics* 136:417-422.
12. Rubin, G. M. 1988. Drosophila melanogaster as an experimental organism. *Science* 240:1453-1459.
13. Simmons, D. 2008. The Use of Animal Models in Studying Genetic Disease: Transgenesis and Induced Mutation. *Nature Education* 1.
14. Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S. E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M. R. Brent, D. G. Brown, S. D. Brown, C. Bult, J. Burton, J. Butler, R. D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A. T. Chinwalla, D. M. Church, M. Clamp, C. Clee, F. S. Collins, L. L. Cook, R. R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K. D. Delehaunty, J. Deri, E. T. Dermitzakis, C. Dewey, N. J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D. M. Dunn, S. R. Eddy, L. Elnitski, R. D. Emes, P. Eswara, E. Eyras, A. Felsenfeld, G. A. Fewell, P. Flicek, K. Foley, W. N. Frankel, L. A. Fulton, R. S. Fulton, T. S. Furey, D. Gage, R. A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T. A. Graves, E. D. Green, S. Gregory, R. Guigo, M. Guyer, R. C. Hardison, D. Haussler, Y. Hayashizaki, L. W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D. B. Jaffe, L. S. Johnson, M. Jones, T. A. Jones, A. Joy, M. Kamal, E. K. Karlsson, D. Karolchik, A. Kasprzyk, J. Kawai, E. Keibler, C. Kells, W. J. Kent, A. Kirby, D. L. Kolbe, I. Korf, R. S. Kucherlapati, E. J. Kulbokas, D. Kulp, T. Landers, J. P. Leger, S. Leonard, I. Letunic, R. Levine, J. Li, M. Li, C. Lloyd, S. Lucas, B. Ma, D. R. Maglott, E. R. Mardis, L. Matthews, E. Mauceli, J. H. Mayer, M. McCarthy, W. R. McCombie, S. McLaren, K. McLay, J. D. McPherson, J. Meldrim, B. Meredith, J. P. Mesirov, W. Miller, T. L. Miner, E. Mongin, K. T. Montgomery, M. Morgan, R. Mott, J. C. Mullikin, D. M. Muzny, W. E. Nash, J. O. Nelson, M. N. Nhan, R. Nicol, Z. Ning, C. Nusbaum, M. J. O'Connor, Y. Okazaki, K.

Oliver, E. Overton-Larty, L. Pachter, G. Parra, K. H. Pepin, J. Peterson, P. Pevzner, R. Plumb, C. S. Pohl, A. Poliakov, T. C. Ponce, C. P. Ponting, S. Potter, M. Quail, A. Reymond, B. A. Roe, K. M. Roskin, E. M. Rubin, A. G. Rust, R. Santos, V. Sapojnikov, B. Schultz, J. Schultz, M. S. Schwartz, S. Schwartz, C. Scott, S. Seaman, S. Searle, T. Sharpe, A. Sheridan, R. Shownkeen, S. Sims, J. B. Singer, G. Slater, A. Smit, D. R. Smith, B. Spencer, A. Stabenau, N. Stange-Thomann, C. Sugnet, M. Suyama, G. Tesler, J. Thompson, D. Torrents, E. Trevaskis, J. Tromp, C. Ucla, A. Ureta-Vidal, J. P. Vinson, A. C. Von Niederhausern, C. M. Wade, M. Wall, R. J. Weber, R. B. Weiss, M. C. Wendl, A. P. West, K. Wetterstrand, R. Wheeler, S. Whelan, J. Wierzbowski, D. Willey, S. Williams, R. K. Wilson, E. Winter, K. C. Worley, D. Wyman, S. Yang, S. P. Yang, E. M. Zdobnov, M. C. Zody, and E. S. Lander. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520-562.

15. Copeland, N. G., N. A. Jenkins, and D. L. Court. 2001. Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* 2:769-779.
16. Beckers, J., W. Wurst, and M. H. de Angelis. 2009. Towards better mouse models: enhanced genotypes, systemic phenotyping and envirotype modelling. *Nat Rev Genet* 10:371-380.
17. Coghill, E. L., A. Hugill, N. Parkinson, C. Davison, P. Glenister, S. Clements, J. Hunter, R. D. Cox, and S. D. Brown. 2002. A gene-driven approach to the identification of ENU mutants in the mouse. *Nat Genet* 30:255-256.
18. Morin, P. J., A. B. Sparks, V. Korinek, N. Barker, H. Clevers, B. Vogelstein, and K. W. Kinzler. 1997. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* 275:1787-1790.
19. Bedell, M. A., N. A. Jenkins, and N. G. Copeland. 1997. Mouse models of human disease. Part I: techniques and resources for genetic analysis in mice. *Genes Dev* 11:1-10.
20. McAllister, K. A., C. D. Houle, J. Malphurs, T. Ward, N. K. Collins, W. Gersch, L. Wharey, J. C. Seely, L. Betz, L. M. Bennett, R. W. Wiseman, and B. J. Davis. 2006. Spontaneous and irradiation-induced tumor susceptibility in BRCA2 germline mutant mice and cooperative effects with a p53 germline mutation. *Toxicol Pathol* 34:187-198.
21. Palmiter, R. D., and R. L. Brinster. 1985. Transgenic mice. *Cell* 41:343-345.
22. Jaenisch, R. 1977. Germ line integration of moloney leukemia virus: effect of homozygosity at the m-mulV locus. *Cell* 12:691-696.
23. Chaible, L. M., M. A. Corat, E. Abdelhay, and M. L. Dagli. Genetically modified animals for use in research and biotechnology. *Genet Mol Res* 9:1469-1482.
24. Schafer, J. M., E. S. Lee, R. M. O'Regan, K. Yao, and V. C. Jordan. 2000. Rapid development of tamoxifen-stimulated mutant p53 breast tumors (T47D) in athymic mice. *Clin Cancer Res* 6:4373-4380.
25. Manis, J. P. 2007. Knock out, knock in, knock down--genetically manipulated mice and the Nobel Prize. *N Engl J Med* 357:2426-2429.
26. de Jong, M., and T. Maina. Of mice and humans: are they the same?--Implications in cancer translational research. *J Nucl Med* 51:501-504.
27. Wolfgang, M. J., and T. G. Golos. 2002. Nonhuman primate transgenesis: progress and prospects. *Trends Biotechnol* 20:479-484.
28. Hanahan, D., and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell* 144:646-674.
29. Ashworth, A., C. J. Lord, and J. S. Reis-Filho. Genetic interactions in cancer progression and treatment. *Cell* 145:30-38.
30. Fantozzi, A., and G. Christofori. 2006. Mouse models of breast cancer metastasis. *Breast Cancer Res* 8:212.
31. Stewart, T. A., P. K. Pattengale, and P. Leder. 1984. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell* 38:627-637.

32. Sinn, E., W. Muller, P. Pattengale, I. Tepler, R. Wallace, and P. Leder. 1987. Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. *Cell* 49:465-475.
33. Kang, S., E. S. Kim, and A. Moon. 2009. Simvastatin and lovastatin inhibit breast cell invasion induced by H-Ras. *Oncol Rep* 21:1317-1322.
34. Silva, J. M., R. Gonzalez, M. Provencio, G. Dominguez, J. M. Garcia, I. Gallego, J. Palacios, P. Espana, and F. Bonilla. 1999. Loss of heterozygosity in BRCA1 and BRCA2 markers and high-grade malignancy in breast cancer. *Breast Cancer Res Treat* 53:9-17.
35. Zeller, J. L., C. Lynm, and R. M. Glass. 2006. JAMA patient page. Colon cancer. *JAMA* 296:1552.
36. Segditsas, S., and I. Tomlinson. 2006. Colorectal cancer and genetic alterations in the Wnt pathway. *Oncogene* 25:7531-7537.
37. Uronis, J. M., and D. W. Threadgill. 2009. Murine models of colorectal cancer. *Mamm Genome* 20:261-268.
38. van Boxtel, R., P. W. Toonen, H. S. van Roekel, M. Verheul, B. M. Smits, J. Korving, A. de Bruin, and E. Cuppen. 2008. Lack of DNA mismatch repair protein MSH6 in the rat results in hereditary non-polyposis colorectal cancer-like tumorigenesis. *Carcinogenesis* 29:1290-1297.
39. Nandan, M. O., A. M. Ghaleb, B. B. McConnell, N. V. Patel, S. Robine, and V. W. Yang. Kruppel-like factor 5 is a crucial mediator of intestinal tumorigenesis in mice harboring combined ApcMin and KRASV12 mutations. *Mol Cancer* 9:63.
40. Brady, C. A., and L. D. Attardi. 2010. p53 at a glance. *J Cell Sci* 123:2527-2532.
41. Zhu, F., M. E. Dolle, T. R. Berton, R. V. Kuiper, C. Capps, A. Espejo, M. J. McArthur, M. T. Bedford, H. van Steeg, A. de Vries, and D. G. Johnson. Mouse models for the p53 R72P polymorphism mimic human phenotypes. *Cancer Res* 70:5851-5859.
42. Aoki, M. N., A. C. da Silva do Amaral Herrera, M. K. Amarante, J. L. do Val Carneiro, M. H. Fungaro, and M. A. Watanabe. 2009. CCR5 and p53 codon 72 gene polymorphisms: implications in breast cancer development. *Int J Mol Med* 23:429-435.
43. Whibley, C., P. D. Pharoah, and M. Hollstein. 2009. p53 polymorphisms: cancer implications. *Nat Rev Cancer* 9:95-107.
44. Weil, P. A., D. S. Luse, J. Segall, and R. G. Roeder. 1979. Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18:469-484.
45. Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 255:11992-11996.
46. Moreau, P., R. Hen, B. Wasylyk, R. Everett, M. P. Gaub, and P. Chambon. 1981. The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res* 9:6047-6068.
47. Nettles, K. W. 2008. Insights into PPARgamma from structures with endogenous and covalently bound ligands. *Nat Struct Mol Biol* 15:893-895.
48. Wolfe, S. A., L. Nekludova, and C. O. Pabo. 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* 29:183-212.
49. Stegmaier, P., A. E. Kel, and E. Wingender. 2004. Systematic DNA-binding domain classification of transcription factors. *Genome Inform* 15:276-286.
50. Regier, J. L., F. Shen, and S. J. Triezenberg. 1993. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc Natl Acad Sci U S A* 90:883-887.
51. Piskacek, S., M. Gregor, M. Nemethova, M. Grabner, P. Kovarik, and M. Piskacek. 2007. Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* 89:756-768.

52. Valentine, J. E., E. Kalkhoven, R. White, S. Hoare, and M. G. Parker. 2000. Mutations in the estrogen receptor ligand binding domain discriminate between hormone-dependent transactivation and transrepression. *J Biol Chem* 275:25322-25329.
53. Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell* 39:499-509.
54. Cansizoglu, A. E., B. J. Lee, Z. C. Zhang, B. M. Fontoura, and Y. M. Chook. 2007. Structure-based design of a pathway-specific nuclear import inhibitor. *Nat Struct Mol Biol* 14:452-454.
55. Huxford, T., D. B. Huang, S. Malek, and G. Ghosh. 1998. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell* 95:759-770.
56. Latimer, M., M. K. Ernst, L. L. Dunn, M. Drutskaya, and N. R. Rice. 1998. The N-terminal domain of IkappaB alpha masks the nuclear localization signal(s) of p50 and c-Rel homodimers. *Mol Cell Biol* 18:2640-2649.
57. Marchenko, N. D., W. Hanel, D. Li, K. Becker, N. Reich, and U. M. Moll. Stress-mediated nuclear stabilization of p53 is regulated by ubiquitination and importin-alpha3 binding. *Cell Death Differ* 17:255-267.
58. Fodde, R., R. Smits, and H. Clevers. 2001. APC, signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer* 1:55-67.
59. Dobrzycka, K. M., S. M. Townson, S. Jiang, and S. Oesterreich. 2003. Estrogen receptor corepressors -- a role in human breast cancer? *Endocr Relat Cancer* 10:517-536.
60. Pratt, W. B., Y. Morishima, M. Murphy, and M. Harrell. 2006. Chaperoning of glucocorticoid receptors. *Handb Exp Pharmacol*:111-138.
61. Tibbetts, R. S., K. M. Brumbaugh, J. M. Williams, J. N. Sarkaria, W. A. Cliby, S. Y. Shieh, Y. Taya, C. Prives, and R. T. Abraham. 1999. A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes Dev* 13:152-157.
62. Hirao, A., Y. Y. Kong, S. Matsuoka, A. Wakeham, J. Ruland, H. Yoshida, D. Liu, S. J. Elledge, and T. W. Mak. 2000. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 287:1824-1827.
63. Stommel, J. M., N. D. Marchenko, G. S. Jimenez, U. M. Moll, T. J. Hope, and G. M. Wahl. 1999. A leucine-rich nuclear export signal in the p53 tetramerization domain: regulation of subcellular localization and p53 activity by NES masking. *EMBO J* 18:1660-1672.
64. Liu, L., D. M. Scolnick, R. C. Trievel, H. B. Zhang, R. Marmorstein, T. D. Halazonetis, and S. L. Berger. 1999. p53 sites acetylated in vitro by PCAF and p300 are acetylated in vivo in response to DNA damage. *Mol Cell Biol* 19:1202-1209.
65. Zhang, Y., Y. Xiong, and W. G. Yarbrough. 1998. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell* 92:725-734.
66. Maiti, B., J. Li, A. de Bruin, F. Gordon, C. Timmers, R. Opavsky, K. Patil, J. Tuttle, W. Cleghorn, and G. Leone. 2005. Cloning and characterization of mouse E2F8, a novel mammalian E2F family member capable of blocking cellular proliferation. *J Biol Chem* 280:18211-18220.
67. Daniels, D. L., and W. I. Weis. 2005. Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol* 12:364-371.
68. Kretzner, L., E. M. Blackwood, and R. N. Eisenman. 1992. Myc and Max proteins possess distinct transcriptional activities. *Nature* 359:426-429.
69. Ayer, D. E., L. Kretzner, and R. N. Eisenman. 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. *Cell* 72:211-222.
70. Dhillon, A. S., S. Hagan, O. Rath, and W. Kolch. 2007. MAP kinase signalling pathways in cancer. *Oncogene* 26:3279-3290.

71. Herbst, R. S. 2004. Review of epidermal growth factor receptor biology. *Int J Radiat Oncol Biol Phys* 59:21-26.
72. Anjum, R., and J. Blenis. 2008. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* 9:747-758.
73. Gilmore, T. D. 2006. Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene* 25:6680-6684.
74. Bonizzi, G., and M. Karin. 2004. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol* 25:280-288.
75. Nishikori, M. 2005. Classical and Alternative NF-kB Activation Pathways and Their Roles in Lymphoid Malignancies. *J.Clin.Exp.Hematopathol* 45:15-24.
76. Rawlings, J. S., K. M. Rosler, and D. A. Harrison. 2004. The JAK/STAT signaling pathway. *J Cell Sci* 117:1281-1283.
77. Bray, S. J. 2006. Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* 7:678-689.
78. Miele, L., T. Golde, and B. Osborne. 2006. Notch signaling in cancer. *Curr Mol Med* 6:905-918.
79. Shah, N., and S. Sukumar. The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* 10:361-371.
80. Pearson, J. C., D. Lemons, and W. McGinnis. 2005. Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* 6:893-904.
81. Bromleigh, V. C., and L. P. Freedman. 2000. p21 is a transcriptional target of HOXA10 in differentiating myelomonocytic cells. *Genes Dev* 14:2581-2586.
82. Carlsson, P., and M. Mahlapuu. 2002. Forkhead transcription factors: key players in development and metabolism. *Dev Biol* 250:1-23.
83. Katoh, M. 2004. Human FOX gene family (Review). *Int J Oncol* 25:1495-1500.
84. Hannenhalli, S., and K. H. Kaestner. 2009. The evolution of Fox genes and their role in development and disease. *Nat Rev Genet* 10:233-240.
85. Lai, E., V. R. Prezioso, E. Smith, O. Litvin, R. H. Costa, and J. E. Darnell, Jr. 1990. HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev* 4:1427-1436.
86. Weigel, D., and H. Jackle. 1990. The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? *Cell* 63:455-456.
87. Clark, K. L., E. D. Halay, E. Lai, and S. K. Burley. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* 364:412-420.
88. Kaestner, K. H., W. Knochel, and D. E. Martinez. 2000. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev* 14:142-146.
89. Kappen, C., K. Schughart, and F. H. Ruddle. 1989. Organization and expression of homeobox genes in mouse and man. *Ann N Y Acad Sci* 567:243-252.
90. Mazet, F., C. T. Amemiya, and S. M. Shimeld. 2006. An ancient Fox gene cluster in bilaterian animals. *Curr Biol* 16:R314-316.
91. Ormestad, M., J. Astorga, H. Landgren, T. Wang, B. R. Johansson, N. Miura, and P. Carlsson. 2006. Foxf1 and Foxf2 control murine gut development by limiting mesenchymal Wnt signaling and promoting extracellular matrix production. *Development* 133:833-843.
92. Wijchers, P. J., J. P. Burbach, and M. P. Smidt. 2006. In control of biology: of mice, men and Foxes. *Biochem J* 397:233-246.
93. Benayoun, B. A., S. Caburet, and R. A. Veitia. Forkhead transcription factors: key players in health and disease. *Trends Genet* 27:224-232.
94. Seo, S., and T. Kume. 2006. Forkhead transcription factors, Foxc1 and Foxc2, are required for the morphogenesis of the cardiac outflow tract. *Dev Biol* 296:421-436.

95. Lehmann, O. J., J. C. Sowden, P. Carlsson, T. Jordan, and S. S. Bhattacharya. 2003. Fox's in development and disease. *Trends Genet* 19:339-344.
96. Hulander, M., W. Wurst, P. Carlsson, and S. Enerback. 1998. The winged helix transcription factor Fkh10 is required for normal development of the inner ear. *Nat Genet* 20:374-376.
97. Johansson, C. C., M. K. Dahle, S. R. Blomqvist, L. M. Gronning, E. M. Aandahl, S. Enerback, and K. Tasken. 2003. A winged helix forkhead (FOXD2) tunes sensitivity to cAMP in T lymphocytes through regulation of cAMP-dependent protein kinase RIalpha. *J Biol Chem* 278:17573-17579.
98. Freyaldenhoven, B. S., M. P. Freyaldenhoven, J. S. Iacovoni, and P. K. Vogt. 1997. Avian winged helix proteins CWH-1, CWH-2 and CWH-3 repress transcription from Qin binding sites. *Oncogene* 15:483-488.
99. Wierstra, I., and J. Alves. 2007. FOXM1, a typical proliferation-associated transcription factor. *Biol Chem* 388:1257-1274.
100. Korver, W., J. Roose, K. Heinen, D. O. Weghuis, D. de Bruijn, A. G. van Kessel, and H. Clevers. 1997. The human TRIDENT/HFH-11/FKHL16 gene: structure, localization, and promoter characterization. *Genomics* 46:435-442.
101. Zhou, S., L. Zawel, C. Lengauer, K. W. Kinzler, and B. Vogelstein. 1998. Characterization of human FAST-1, a TGF beta and activin signal transducer. *Mol Cell* 2:121-127.
102. Seoane, J., H. V. Le, L. Shen, S. A. Anderson, and J. Massague. 2004. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* 117:211-223.
103. Hanashima, C., L. Shen, S. C. Li, and E. Lai. 2002. Brain factor-1 controls the proliferation and differentiation of neocortical progenitor cells through independent mechanisms. *J Neurosci* 22:6526-6536.
104. Li, C., A. J. Lusis, R. Sparkes, S. M. Tran, and R. Gaynor. 1992. Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein HTLF. *Genomics* 13:658-664.
105. Nehls, M., D. Pfeifer, M. Schorpp, H. Hedrich, and T. Boehm. 1994. New member of the winged-helix protein family disrupted in mouse and rat nude mutations. *Nature* 372:103-107.
106. Segre, J. A., J. L. Nemhauser, B. A. Taylor, J. H. Nadeau, and E. S. Lander. 1995. Positional cloning of the nude locus: genetic, physical, and transcription maps of the region and mutations in the mouse and rat. *Genomics* 28:549-559.
107. Balciunaite, G., M. P. Keller, E. Balciunaite, L. Piali, S. Zuklys, Y. D. Mathieu, J. Gill, R. Boyd, D. J. Sussman, and G. A. Hollander. 2002. Wnt glycoproteins regulate the expression of FoxN1, the gene defective in nude mice. *Nat Immunol* 3:1102-1108.
108. Pati, D., C. Keller, M. Groudine, and S. E. Plon. 1997. Reconstitution of a MEC1-independent checkpoint in yeast by expression of a novel human fork head cDNA. *Mol Cell Biol* 17:3037-3046.
109. Gouge, A., J. Holt, A. P. Hardy, J. C. Sowden, and H. K. Smith. 2001. Foxn4--a new member of the forkhead gene family is expressed in the retina. *Mech Dev* 107:203-206.
110. Katoh, M. 2004. Identification and characterization of human FOXN5 and rat Foxn5 genes in silico. *Int J Oncol* 24:1339-1344.
111. Katoh, M. 2004. Identification and characterization of human FOXN6, mouse Foxn6, and rat Foxn6 genes in silico. *Int J Oncol* 25:219-223.
112. Schuff, M., A. Rossner, C. Donow, and W. Knochel. 2006. Temporal and spatial expression patterns of FoxN genes in Xenopus laevis embryos. *Int J Dev Biol* 50:429-434.
113. Katoh, M. 2007. WNT signaling pathway and stem cell signaling network. *Clin Cancer Res* 13:4042-4045.
114. Scott, K. L., and S. E. Plon. 2003. Loss of Sin3/Rpd3 histone deacetylase restores the DNA damage response in checkpoint-deficient strains of Saccharomyces cerevisiae. *Mol Cell Biol* 23:4522-4531.

115. Kadosh, D., and K. Struhl. 1998. Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Mol Cell Biol* 18:5121-5127.
116. Scott, K. L., and S. E. Plon. 2005. CHES1/FOXN3 interacts with Ski-interacting protein and acts as a transcriptional repressor. *Gene* 359:119-126.
117. Folk, P., F. Puta, and M. Skruzny. 2004. Transcriptional coregulator SNW/SKIP: the concealed tie of dissimilar pathways. *Cell Mol Life Sci* 61:629-640.
118. Dahl, R., B. Wani, and M. J. Hayman. 1998. The Ski oncoprotein interacts with Skip, the human homolog of Drosophila Bx42. *Oncogene* 16:1579-1586.
119. Zhou, S., M. Fujimuro, J. J. Hsieh, L. Chen, A. Miyamoto, G. Weinmaster, and S. D. Hayward. 2000. SKIP, a CBF1-associated protein, interacts with the ankyrin repeat domain of NotchIC To facilitate NotchIC function. *Mol Cell Biol* 20:2400-2410.
120. Knijnenburg, J., A. van Haeringen, K. B. Hansson, A. Lankester, M. J. Smit, R. D. Belfroid, E. Bakker, C. Rosenberg, H. J. Tanke, and K. Szuhai. 2007. Ring chromosome formation as a novel escape mechanism in patients with inverted duplication and terminal deletion. *Eur J Hum Genet* 15:548-555.
121. Schlade-Bartusiak, K., G. Macintyre, J. Zunich, and D. W. Cox. 2008. A child with deletion (14)(q24.3q32.13) and auditory neuropathy. *Am J Med Genet A* 146A:117-123.
122. Schuff, M., A. Rossner, S. A. Wacker, C. Donow, S. Gessert, and W. Knochel. 2007. FoxN3 is required for craniofacial and eye development of Xenopus laevis. *Dev Dyn* 236:226-239.
123. Stryke, D., M. Kawamoto, C. C. Huang, S. J. Johns, L. A. King, C. A. Harper, E. C. Meng, R. E. Lee, A. Yee, L. L'Italien, P. T. Chuang, S. G. Young, W. C. Skarnes, P. C. Babbitt, and T. E. Ferrin. 2003. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. *Nucleic Acids Res* 31:278-281.
124. Samaan, G., D. Yugo, S. Rajagopalan, J. Wall, R. Donnell, D. Goldowitz, R. Gopalakrishnan, and S. Venkatachalam. Foxn3 is essential for craniofacial development in mice and a putative candidate involved in human congenital craniofacial defects. *Biochem Biophys Res Commun* 400:60-65.
125. Cecconi, F., and B. I. Meyer. 2000. Gene trap: a way to identify novel genes and unravel their biological function. *FEBS Lett* 480:63-71.
126. Nagarajan, P., T. M. Onami, S. Rajagopalan, S. Kania, R. Donnell, and S. Venkatachalam. 2009. Role of chromodomain helicase DNA-binding protein 2 in DNA damage response signaling and tumorigenesis. *Oncogene* 28:1053-1062.
127. Schliekelman, M., D. O. Cowley, R. O'Quinn, T. G. Oliver, L. Lu, E. D. Salmon, and T. Van Dyke. 2009. Impaired Bub1 function in vivo compromises tension-dependent checkpoint function leading to aneuploidy and tumorigenesis. *Cancer Res* 69:45-54.
128. Gonzalez-Billault, C., E. Demandt, F. Wandosell, M. Torres, P. Bonaldo, A. Stoykova, K. Chowdhury, P. Gruss, J. Avila, and M. P. Sanchez. 2000. Perinatal lethality of microtubule-associated protein 1B-deficient mice expressing alternative isoforms of the protein at low levels. *Mol Cell Neurosci* 16:408-421.
129. Leland, S., P. Nagarajan, A. Polyzos, S. Thomas, G. Samaan, R. Donnell, F. Marchetti, and S. Venkatachalam. 2009. Heterozygosity for a Bub1 mutation causes female-specific germ cell aneuploidy in mice. *Proc Natl Acad Sci U S A* 106:12776-12781.
130. Strachan, T. 1999. Genes in Pedigree. In *Human Molecular Genetics 2*. John Wiley and Sons, Inc,.
131. Zelzer, E., and B. R. Olsen. 2003. The genetic basis for skeletal diseases. *Nature* 423:343-348.
132. Jiang, X., S. Iseki, R. E. Maxson, H. M. Sucov, and G. M. Morriss-Kay. 2002. Tissue origins and interactions in the mammalian skull vault. *Dev Biol* 241:106-116.

133. Karsenty, G. 2008. Transcriptional control of skeletogenesis. *Annu Rev Genomics Hum Genet* 9:183-196.
134. Gans, C., and R. G. Northcutt. 1983. Neural crest and the origin of vertebrates: a new head. *Science* 220:268-273.
135. Opperman, L. A. 2000. Cranial sutures as intramembranous bone growth sites. *Dev Dyn* 219:472-485.
136. Nie, X., K. Luukko, and P. Kettunen. 2006. FGF signalling in craniofacial development and developmental disorders. *Oral Dis* 12:102-111.
137. Richman, J. M., and S. H. Lee. 2003. About face: signals and genes controlling jaw patterning and identity in vertebrates. *Bioessays* 25:554-568.
138. Deckelbaum, R. A., A. Majithia, T. Booker, J. E. Henderson, and C. A. Loomis. 2006. The homeoprotein engrailed 1 has pleiotropic functions in calvarial intramembranous bone formation and remodeling. *Development* 133:63-74.
139. Kraus, P., and T. Lufkin. 2006. Dlx homeobox gene control of mammalian limb and craniofacial development. *Am J Med Genet A* 140:1366-1374.
140. Tsuji, K., K. Cox, A. Bandyopadhyay, B. D. Harfe, C. J. Tabin, and V. Rosen. 2008. BMP4 is dispensable for skeletogenesis and fracture-healing in the limb. *J Bone Joint Surg Am* 90 Suppl 1:14-18.
141. Wright, V., H. Peng, A. Usas, B. Young, B. Gearhart, J. Cummins, and J. Huard. 2002. BMP4-expressing muscle-derived stem cells differentiate into osteogenic lineage and improve bone healing in immunocompetent mice. *Mol Ther* 6:169-178.
142. Myatt, S. S., and E. W. Lam. 2007. The emerging roles of forkhead box (Fox) proteins in cancer. *Nat Rev Cancer* 7:847-859.
143. Tuteja, G., and K. H. Kaestner. 2007. SnapShot: forkhead transcription factors I. *Cell* 130:1160.
144. Jackson, B. C., C. Carpenter, D. W. Nebert, and V. Vasiliou. Update of human and mouse forkhead box (FOX) gene families. *Hum Genomics* 4:345-352.
145. Krupczak-Hollis, K., X. Wang, V. V. Kalinichenko, G. A. Gusarova, I. C. Wang, M. B. Dennewitz, H. M. Yoder, H. Kiyokawa, K. H. Kaestner, and R. H. Costa. 2004. The mouse Forkhead Box m1 transcription factor is essential for hepatoblast mitosis and development of intrahepatic bile ducts and vessels during liver morphogenesis. *Dev Biol* 276:74-88.
146. Ye, H., T. F. Kelly, U. Samadani, L. Lim, S. Rubio, D. G. Overdier, K. A. Roebuck, and R. H. Costa. 1997. Hepatocyte nuclear factor 3/fork head homolog 11 is expressed in proliferating epithelial and mesenchymal cells of embryonic and adult tissues. *Mol Cell Biol* 17:1626-1641.
147. Biggs, W. H., 3rd, J. Meisenhelder, T. Hunter, W. K. Cavenee, and K. C. Arden. 1999. Protein kinase B/Akt-mediated phosphorylation promotes nuclear exclusion of the winged helix transcription factor FKHR1. *Proc Natl Acad Sci U S A* 96:7421-7426.
148. Ramaswamy, S., N. Nakamura, I. Sansal, L. Bergeron, and W. R. Sellers. 2002. A novel mechanism of gene regulation and tumor suppression by the transcription factor FKHR. *Cancer Cell* 2:81-91.
149. Paik, J. H., R. Kollipara, G. Chu, H. Ji, Y. Xiao, Z. Ding, L. Miao, Z. Tothova, J. W. Horner, D. R. Carrasco, S. Jiang, D. G. Gilliland, L. Chin, W. H. Wong, D. H. Castrillon, and R. A. DePinho. 2007. FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* 128:309-323.
150. Greer, E. L., and A. Brunet. 2005. FOXO transcription factors at the interface between longevity and tumor suppression. *Oncogene* 24:7410-7425.
151. Banham, A. H., N. Beasley, E. Campo, P. L. Fernandez, C. Fidler, K. Gatter, M. Jones, D. Y. Mason, J. E. Prime, P. Trougouboff, K. Wood, and J. L. Cordell. 2001. The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p. *Cancer Res* 61:8820-8829.

152. Ferland, R. J., T. J. Cherry, P. O. Preware, E. E. Morrisey, and C. A. Walsh. 2003. Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *J Comp Neurol* 460:266-279.
153. Wlodarska, I., E. Veyt, P. De Paepe, P. Vandenberghe, P. Nooijen, I. Theate, L. Michaux, X. Sagaert, P. Marynen, A. Hagemeijer, and C. De Wolf-Peeters. 2005. FOXP1, a gene highly expressed in a subset of diffuse large B-cell lymphoma, is recurrently targeted by genomic aberrations. *Leukemia* 19:1299-1305.
154. Lai, C. S., S. E. Fisher, J. A. Hurst, F. Vargha-Khadem, and A. P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519-523.
155. Brown, P. J., S. L. Ashe, E. Leich, C. Burek, S. Barrans, J. A. Fenton, A. S. Jack, K. Pulford, A. Rosenwald, and A. H. Banham. 2008. Potentially oncogenic B-cell activation-induced smaller isoforms of FOXP1 are highly expressed in the activated B cell-like subtype of DLBCL. *Blood* 111:2816-2824.
156. Fox, S. B., P. Brown, C. Han, S. Ashe, R. D. Leek, A. L. Harris, and A. H. Banham. 2004. Expression of the forkhead transcription factor FOXP1 is associated with estrogen receptor alpha and improved survival in primary human breast carcinomas. *Clin Cancer Res* 10:3521-3527.
157. Shu, W., H. Yang, L. Zhang, M. M. Lu, and E. E. Morrisey. 2001. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. *J Biol Chem* 276:27488-27497.
158. Bennett, C. L., J. Christie, F. Ramsdell, M. E. Brunkow, P. J. Ferguson, L. Whitesell, T. E. Kelly, F. T. Saulsbury, P. F. Chance, and H. D. Ochs. 2001. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet* 27:20-21.
159. Schubert, L. A., E. Jeffery, Y. Zhang, F. Ramsdell, and S. F. Ziegler. 2001. Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation. *J Biol Chem* 276:37672-37679.
160. Wang, L., R. Liu, W. Li, C. Chen, H. Katoh, G. Y. Chen, B. McNally, L. Lin, P. Zhou, T. Zuo, K. A. Cooney, Y. Liu, and P. Zheng. 2009. Somatic single hits inactivate the X-linked tumor suppressor FOXP3 in the prostate. *Cancer Cell* 16:336-346.
161. Chang, J. T., H. M. Wang, K. W. Chang, W. H. Chen, M. C. Wen, Y. M. Hsu, B. Y. Yung, I. H. Chen, C. T. Liao, L. L. Hsieh, and A. J. Cheng. 2005. Identification of differentially expressed genes in oral squamous cell carcinoma (OSCC): overexpression of NPM, CDK1 and NDRG1 and underexpression of CHES1. *Int J Cancer* 114:942-949.
162. Struckmann, K., P. Schraml, R. Simon, K. Elmenhorst, M. Mirlacher, J. Kononen, and H. Moch. 2004. Impaired expression of the cell cycle regulator BTG2 is common in clear cell renal cell carcinoma. *Cancer Res* 64:1632-1638.
163. Markowski, J., T. Tyszkiewicz, M. Jarzab, M. Oczko-Wojciechowska, T. Gierek, M. Witkowska, J. Paluch, M. Kowalska, Z. Wygoda, D. Lange, and B. Jarzab. 2009. Metal-proteinase ADAM12, kinesin 14 and checkpoint suppressor 1 as new molecular markers of laryngeal carcinoma. *Eur Arch Otorhinolaryngol* 266:1501-1507.
164. Venkatachalam, S., Y. P. Shi, S. N. Jones, H. Vogel, A. Bradley, D. Pinkel, and L. A. Donehower. 1998. Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. *EMBO J* 17:4657-4667.
165. Akagi, T., K. Sasai, and H. Hanafusa. 2003. Refractory nature of normal human diploid fibroblasts with respect to oncogene-mediated transformation. *Proc Natl Acad Sci U S A* 100:13567-13572.
166. Bristow, R. G., Q. Hu, A. Jang, S. Chung, J. Peacock, S. Benchimol, and R. Hill. 1998. Radioresistant MTp53-expressing rat embryo cell transformants exhibit increased DNA-dsb rejoining during exposure to ionizing radiation. *Oncogene* 16:1789-1802.
167. Arora, S., A. Matta, N. K. Shukla, S. V. Deo, and R. Ralhan. 2005. Identification of differentially expressed genes in oral squamous cell carcinoma. *Mol Carcinog* 42:97-108.

168. Haber, D. A., and J. Settleman. 2007. Cancer: drivers and passengers. *Nature* 446:145-146.
169. Malumbres, M., and M. Barbacid. 2005. Mammalian cyclin-dependent kinases. *Trends Biochem Sci* 30:630-641.
170. Berthet, C., and P. Kaldis. 2006. Cdk2 and Cdk4 cooperatively control the expression of Cdc2. *Cell Div* 1:10.
171. Garcia-Tunon, I., M. Ricote, A. Ruiz, B. Fraile, R. Paniagua, and M. Royuela. 2006. Cell cycle control related proteins (p53, p21, and Rb) and transforming growth factor beta (TGFbeta) in benign and carcinomatous (in situ and infiltrating) human breast: implications in malignant transformations. *Cancer Invest* 24:119-125.
172. Barriere, C., D. Santamaria, A. Cerqueira, J. Galan, A. Martin, S. Ortega, M. Malumbres, P. Dubus, and M. Barbacid. 2007. Mice thrive without Cdk4 and Cdk2. *Mol Oncol* 1:72-83.
173. Li, W., S. Kotoshiba, C. Berthet, M. B. Hilton, and P. Kaldis. 2009. Rb/Cdk2/Cdk4 triple mutant mice elicit an alternative mechanism for regulation of the G1/S transition. *Proc Natl Acad Sci U S A* 106:486-491.
174. Lindqvist, A., W. van Zon, C. Karlsson Rosenthal, and R. M. Wolthuis. 2007. Cyclin B1-Cdk1 activation continues after centrosome separation to control mitotic progression. *PLoS Biol* 5:e123.
175. Mitra, J., and G. H. Enders. 2004. Cyclin A/Cdk2 complexes regulate activation of Cdk1 and Cdc25 phosphatases in human cells. *Oncogene* 23:3361-3367.
176. Zhou, V. W., A. Goren, and B. E. Bernstein. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet* 12:7-18.
177. Li, G., and D. Reinberg. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev* 21:175-186.
178. Hirano, T. 2005. SMC proteins and chromosome mechanics: from bacteria to humans. *Philos Trans R Soc Lond B Biol Sci* 360:507-514.
179. Sawan, C., and Z. Herceg. Histone modifications and cancer. *Adv Genet* 70:57-85.
180. Bird, A. 2007. Perceptions of epigenetics. *Nature* 447:396-398.
181. Wang, H., Z. Q. Huang, L. Xia, Q. Feng, H. Erdjument-Bromage, B. D. Strahl, S. D. Briggs, C. D. Allis, J. Wong, P. Tempst, and Y. Zhang. 2001. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* 293:853-857.
182. Kirmizis, A., H. Santos-Rosa, C. J. Penkett, M. A. Singer, R. D. Green, and T. Kouzarides. 2009. Distinct transcriptional outputs associated with mono- and dimethylated histone H3 arginine 2. *Nat Struct Mol Biol* 16:449-451.
183. Feng, Q., B. He, S. Y. Jung, Y. Song, J. Qin, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley. 2009. Biochemical control of CARM1 enzymatic activity by phosphorylation. *J Biol Chem* 284:36167-36174.
184. Kleinschmidt, M. A., G. Streubel, B. Samans, M. Krause, and U. M. Bauer. 2008. The protein arginine methyltransferases CARM1 and PRMT1 cooperate in gene regulation. *Nucleic Acids Res* 36:3202-3213.
185. Peters, A. H., S. Kubicek, K. Mechtler, R. J. O'Sullivan, A. A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J. H. Martens, and T. Jenuwein. 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12:1577-1589.
186. Qian, C., and M. M. Zhou. 2006. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci* 63:2755-2763.
187. Bannister, A. J., and T. Kouzarides. 2005. Reversing histone methylation. *Nature* 436:1103-1106.
188. Kimura, A., K. Matsubara, and M. Horikoshi. 2005. A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J Biochem* 138:647-662.

189. Agalioti, T., G. Chen, and D. Thanos. 2002. Deciphering the transcriptional histone acetylation code for a human gene. *Cell* 111:381-392.
190. Marmorstein, R., and S. Y. Roth. 2001. Histone acetyltransferases: function, structure, and catalysis. *Curr Opin Genet Dev* 11:155-161.
191. Pal, S., R. Yun, A. Datta, L. Lacomis, H. Erdjument-Bromage, J. Kumar, P. Tempst, and S. Sif. 2003. mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. *Mol Cell Biol* 23:7475-7487.
192. Zhang, J., M. Kalkum, B. T. Chait, and R. G. Roeder. 2002. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol Cell* 9:611-623.
193. Tsukiyama, T. 2002. The in vivo functions of ATP-dependent chromatin-remodelling factors. *Nat Rev Mol Cell Biol* 3:422-429.
194. Conaway, R. C., and J. W. Conaway. 2009. The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends Biochem Sci* 34:71-77.
195. Deuring, R., L. Fanti, J. A. Armstrong, M. Sarte, O. Papoulas, M. Prestel, G. Daubresse, M. Verardo, S. L. Moseley, M. Berloco, T. Tsukiyama, C. Wu, S. Pimpinelli, and J. W. Tamkun. 2000. The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol Cell* 5:355-365.
196. Whitehouse, I., C. Stockdale, A. Flaus, M. D. Szczelkun, and T. Owen-Hughes. 2003. Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Mol Cell Biol* 23:1935-1945.
197. Burns, L. G., and C. L. Peterson. 1997. The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. *Mol Cell Biol* 17:4811-4819.
198. Wilson, B. G., and C. W. Roberts. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* 11:481-492.
199. Xue, Y., J. C. Canman, C. S. Lee, Z. Nie, D. Yang, G. T. Moreno, M. K. Young, E. D. Salmon, and W. Wang. 2000. The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proc Natl Acad Sci U S A* 97:13015-13020.
200. Chi, T. H., M. Wan, K. Zhao, I. Taniuchi, L. Chen, D. R. Littman, and G. R. Crabtree. 2002. Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. *Nature* 418:195-199.
201. Woodage, T., M. A. Basrai, A. D. Baxevanis, P. Hieter, and F. S. Collins. 1997. Characterization of the CHD family of proteins. *Proc Natl Acad Sci U S A* 94:11472-11477.
202. Bottomley, M. J. 2004. Structures of protein domains that create or recognize histone modifications. *EMBO Rep* 5:464-469.
203. Jones, D. O., I. G. Cowell, and P. B. Singh. 2000. Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays* 22:124-137.
204. Ryan, D. P., R. Sundaramoorthy, D. Martin, V. Singh, and T. Owen-Hughes. The DNA-binding domain of the Chd1 chromatin-remodelling enzyme contains SANT and SLIDE domains. *EMBO J* 30:2596-2609.
205. Flanagan, J. F., L. Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, and S. Khorasanizadeh. 2005. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 438:1181-1185.
206. Delmas, V., D. G. Stokes, and R. P. Perry. 1993. A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proc Natl Acad Sci U S A* 90:2414-2418.
207. Stokes, D. G., and R. P. Perry. 1995. DNA-binding and chromatin localization properties of CHD1. *Mol Cell Biol* 15:2745-2753.

208. Flanagan, J. F., B. J. Blus, D. Kim, K. L. Clines, F. Rastinejad, and S. Khorasanizadeh. 2007. Molecular implications of evolutionary differences in CHD double chromodomains. *J Mol Biol* 369:334-342.
209. Ogata, K., C. Kanei-Ishii, M. Sasaki, H. Hatanaka, A. Nagadoi, M. Enari, H. Nakamura, Y. Nishimura, S. Ishii, and A. Sarai. 1996. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat Struct Biol* 3:178-187.
210. Kumari, A., O. M. Mazina, U. Shinde, A. V. Mazin, and H. Lu. 2009. A role for SSRP1 in recombination-mediated DNA damage response. *J Cell Biochem* 108:508-518.
211. Simic, R., D. L. Lindstrom, H. G. Tran, K. L. Roinick, P. J. Costa, A. D. Johnson, G. A. Hartzog, and K. M. Arndt. 2003. Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *EMBO J* 22:1846-1856.
212. Pray-Grant, M. G., J. A. Daniel, D. Schieltz, J. R. Yates, 3rd, and P. A. Grant. 2005. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433:434-438.
213. Biswas, D., R. Dutta-Biswas, D. Mitra, Y. Shibata, B. D. Strahl, T. Formosa, and D. J. Stillman. 2006. Opposing roles for Set2 and yFACT in regulating TBP binding at promoters. *EMBO J* 25:4479-4489.
214. Denslow, S. A., and P. A. Wade. 2007. The human Mi-2/NuRD complex and gene regulation. *Oncogene* 26:5433-5438.
215. Bowen, N. J., N. Fujita, M. Kajita, and P. A. Wade. 2004. Mi-2/NuRD: multiple complexes for many purposes. *Biochim Biophys Acta* 1677:52-57.
216. Pascual, J., M. Martinez-Yamout, H. J. Dyson, and P. E. Wright. 2000. Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. *J Mol Biol* 304:723-729.
217. Marfella, C. G., and A. N. Imbalzano. 2007. The Chd family of chromatin remodelers. *Mutat Res* 618:30-40.
218. Thompson, P. M., T. Gotoh, M. Kok, P. S. White, and G. M. Brodeur. 2003. CHD5, a new member of the chromodomain gene family, is preferentially expressed in the nervous system. *Oncogene* 22:1002-1011.
219. Bagchi, A., and A. A. Mills. 2008. The quest for the 1p36 tumor suppressor. *Cancer Res* 68:2551-2556.
220. Wang, X., K. K. Lau, L. K. So, and Y. W. Lam. 2009. CHD5 is down-regulated through promoter hypermethylation in gastric cancer. *J Biomed Sci* 16:95.
221. Gorringe, K. L., D. Y. Choong, L. H. Williams, M. Ramakrishna, A. Sridhar, W. Qiu, J. L. Bearfoot, and I. G. Campbell. 2008. Mutation and methylation analysis of the chromodomain-helicase-DNA binding 5 gene in ovarian cancer. *Neoplasia* 10:1253-1258.
222. Lutz, T., R. Stoger, and A. Nieto. 2006. CHD6 is a DNA-dependent ATPase and localizes at nuclear sites of mRNA synthesis. *FEBS Lett* 580:5851-5857.
223. Biswas, M., and J. Y. Chan. Role of Nrf1 in antioxidant response element-mediated gene expression and beyond. *Toxicol Appl Pharmacol* 244:16-20.
224. Zentner, G. E., W. S. Layman, D. M. Martin, and P. C. Scacheri. Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *Am J Med Genet A* 152A:674-686.
225. Bosman, E. A., A. C. Penn, J. C. Ambrose, R. Kettleborough, D. L. Stemple, and K. P. Steel. 2005. Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. *Hum Mol Genet* 14:3463-3476.
226. Schnetz, M. P., L. Handoko, B. Akhtar-Zaidi, C. F. Bartels, C. F. Pereira, A. G. Fisher, D. J. Adams, P. Flicek, G. E. Crawford, T. Laframboise, P. Tesar, C. L. Wei, and P. C. Scacheri. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet* 6:e1001023.

227. Ishihara, K., M. Oshimura, and M. Nakao. 2006. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 23:733-742.
228. Thompson, B. A., V. Tremblay, G. Lin, and D. A. Bochar. 2008. CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. *Mol Cell Biol* 28:3894-3904.
229. Nishiyama, M., K. Oshikawa, Y. Tsukada, T. Nakagawa, S. Iemura, T. Natsume, Y. Fan, A. Kikuchi, A. I. Skoultchi, and K. I. Nakayama. 2009. CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat Cell Biol* 11:172-182.
230. Shur, I., and D. Benayahu. 2005. Characterization and functional analysis of CReMM, a novel chromodomain helicase DNA-binding protein. *J Mol Biol* 352:646-655.
231. Osman, I., D. F. Bajorin, T. T. Sun, H. Zhong, D. Douglas, J. Scattergood, R. Zheng, M. Han, K. W. Marshall, and C. C. Liew. 2006. Novel blood biomarkers of human urinary bladder cancer. *Clin Cancer Res* 12:3374-3380.
232. Feys, T., B. Poppe, K. De Preter, N. Van Roy, B. Verhasselt, P. De Paepe, A. De Paepe, and F. Speleman. 2007. A detailed inventory of DNA copy number alterations in four commonly used Hodgkin's lymphoma cell lines. *Haematologica* 92:913-920.
233. Marfella, C. G., N. Henninger, S. E. LeBlanc, N. Krishnan, D. S. Garlick, L. B. Holzman, and A. N. Imbalzano. 2008. A mutation in the mouse Chd2 chromatin remodeling enzyme results in a complex renal phenotype. *Kidney Blood Press Res* 31:421-432.
234. Gomes, N. P., and J. M. Espinosa. Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev* 24:1022-1034.
235. Martinato, F., M. Cesaroni, B. Amati, and E. Guccione. 2008. Analysis of Myc-induced histone modifications on target chromatin. *PLoS One* 3:e3650.

Vita

George Samaan was born on July 25th, 1985 in Hims, Syria Arab Republic. He migrated to the United States in June of 1997 and lived in Jacksonville, FL for eight months before moving to Surfside, FL. He attended and graduated from Miami Beach Senior High School in June of 2002 after which he attended Miami-Dade College and obtained Associates in Arts. He graduated from the University of Florida in 2007 with a Bachelor's in Arts with a major of Interdisciplinary Studies (Biology track) and finished the requirements for a double major in Microbiology. In 2007, he joined the PhD program in Biochemistry, Cellular and Molecular Biology at the University of Tennessee, Knoxville. After graduation, he plans on attending Medical School and pursue a career in Medicine.