Variants in *Hdac9* Intronic Enhancer As Candidates for *Skin Tumor Susceptibility 5* (*Skts5*) Locus

Undergraduate Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Research Distinction in Biomedical Science in the School of Health and Rehabilitation Sciences at The Ohio State University

Tyler E. Siekmann

Undergraduate Biomedical Science Major
School of Health and Rehabilitation Sciences
The Ohio State University

2014

Thesis Committee:
Dr. Amanda E. Toland, Ph.D, Advisor
Dr. Paul Goodfellow, Ph.D
Dr. Amanda M. Agnew, Ph.D

Abstract:

Non-melanoma skin cancers (NMSC) are the most common forms of cancer in the world accounting for nearly half of all cancer diagnoses. Rates of NMSC are on the rise with an over 300% increase in diagnosis of these cancers in the last 20 years. While environmental risk factors for skin cancer such as ultraviolet light (UV) exposure are well understood, little is known about inherited genetic risk factors for these cancers. Mouse linkage studies have identified several loci housing skin cancer susceptibility genes (susceptibility to skin cancer or *Skts*). Human tumors show evidence of preferential allelic imbalance for polymorphisms in *Hdac9*, a gene mapping to one of the linkage regions, *Skts5*. An intron in the *Hdac9* gene between exons 8 and 9 was shown by others to contain an enhancer for *Twist1* which affects early limb development and phenotypes in the skin. *Twist1* is a known regulator of skin differentiation and has a documented role in cancer including metastasis and cell growth. The hypothesis of this study is that this enhancer locus plays a role in the differential risk for NMSC between the cancer susceptible *NIH/Ola* and cancer resistant *Spretus/EiJ* mice. To test this hypothesis, we first looked for sequence differences between the strains. Sequence analyses led to the identification of several polymorphisms in this intron which are predicted by multiple *in silico* methods to disrupt known transcription factor binding. To investigate the *in vitro* effects of these variants, intron fragments from both *NIH/Ola* and *Spret/EiJ* murine DNA were cloned into an enhancer pGL3 reporter vector and transfected into both normal keratinocyte C5N and squamous cell carcinoma A5 cells. Luciferase assay and real-time PCR data suggests these variants are responsible for changes in gene expression, specifically in the *Twist1* gene. Chromatin immunoprecipitation studies found that two transcription factors, Gata3 and Oct1, preferentially bind to *NIH* or *Spretus* DNA respectively at the enhancer locus, suggesting a possible mechanism for enhancer

activity. In summary, variants at the *Hdac9* intron 8 enhancer appear to affect *Twist1* expression and may explain the skin cancer susceptibility locus *Skts5*.

Acknowledgements:

First and foremost, I would like to thank my research advisor Dr. Amanda Toland for her continued support and scientific guidance throughout the past three years. Her patience, encouragement and positive attitude were instrumental to both this project and my growth as a student scientist. Dr. Toland took me in as a novice scientist, providing all the tools necessary to give me an outstanding education in the scientific process and inspiring me to question the world around me. Without her guidance as a scientific mentor I would be neither the researcher nor person I am today.

I would also like to thank members of the Toland Lab who have played various roles in my development as a student scientist. I owe a great deal of gratitude to my original lab mentor Laura Skeeles, Jessica Fleming, Madelyn Gerber, Jessica Gillespie, Kelsey Gray, Allison Nosotti, Luke Cheng and Matt Cianciolo.

In addition, I would like to give a special thank you to Dr. John Gunn, Lori Martensen and Steven Mousetes who have served as my advisors and mentors through the Biomedical Science Undergraduate Program. They have supported me throughout my undergraduate years and have inspired me to pursue my lifelong goal of a career in medicine.

This work would not be possible if it were not for the generous support of *Pelotonia* and the Pelotonia Undergraduate Research Fellowship. Any opinions, findings, and conclusions expressed in this material are those of the author and do not necessarily reflect those of the Pelotonia Fellowship Program.

Finally, I would like to thank my loving parents Cyndi and Rick Siekmann and my brothers Drew and Brad Siekmann for their support and guidance in all that I do.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Non-Melanoma Skin Cancer (NMSC):

Non-melanoma skin cancer is a broad category of metastatic disease that accounts for nearly 3.5 million cancer diagnosis in the United States each year (Skin Cancer Foundation, SCF).The two most common forms of NMSC include basal cell carcinoma (BCC), a cancer originating in the basal cells of the epidermis, and squamous cell carcinoma (SCC), which initiates in the keratinocytes of the epidermis. These two types of NMSC account for roughly 96% of all NMSC diagnoses each year, with 80% of cases classified as BCC. Other rare non-melanoma skin cancers include Kaposi's sarcoma, normally seen only in immunocompromised patients, and Merkel cell carcinoma, a neuroendocrine related cancer usually developing in the hair follicles (UCSF Dermatology). These cancers do not include melanomas, which alternatively result from malignancies originating in pigment producing melanocytes.

While instances of mortality resulting from NMSC are low, rates of recurrence are high and NMSC patients are at increased risk for developing other cancers. The cost of NMSC treatments are a significant financial burden, totaling over 1.4 billion dollars in 2004 (SCF). Alarmingly, there has been a 300% increase in the incidence of skin cancers since 1994 (SCF). This increase can be partially attributed to a rise in environmental risk factors, most notably increased ultraviolet light (UV) exposure resulting from a depleted ozone layer and a rise in tanning bed usage. Other NMSC risk factors include: light skin and eye color, family and personal history, smoking, radiation or chemical carcinogen exposure and immunosuppressive medical conditions (Cancer.gov). While it is understood there is a genetic component of risk, the exact genetic risk factors and mechanism of SCC development is poorly understood to date.

Skin Tumor Susceptibility Locus 5 *(Skts5)*:

Linkage analysis using F1 backcrosses between skin cancer resistant *Mus spretus* and susceptible *Mus musculus* mice has led to the identification of a susceptibility locus for skin cancer on mouse chromosome 12, called *Skts5 (Mahler et al. 2008)*. In humans, the orthologous locus mapping to 7p21 and 7q31 has also been identified as a susceptibility locus through allelic imbalance studies (*Fleming et al. 2014*). The human studies revealed that roughly 20% of cutaneous SCC tumors exhibit copy neutral loss of heterozygocity and copy number increases are present in 10% of tumors. Additionally, microsatellite markers within *Skts5* show preferential allelic imbalance in SCC tumors. In humans, nine single nucleotide polymorphisms (SNPs) mapping to *Hdac9* within *Skts5* demonstrated significant preferential allelic imbalance in SCC tumors, suggesting that *Hdac9* or variants mapping near *Hdac9* may be important in human SCC.

Genetic Enhancers:

Enhancers are short DNA sequences that act as cis or trans-regulatory genetic elements, increasing the expression of a target gene. Enhancers exhibit their effects through binding specific transcription factors that favorably interact with the promoter and transcription initiation site of the target gene, increasing levels of transcription and thus gene expression. Enhancers can be found either upstream or downstream of the target gene, as long as their spatial organization in the nucleus allows interaction with the target gene. They can be located hundreds or thousands of base pairs away from this gene of interest, and can be found in either coding regions of the genome or in non-coding elements such as introns. A unique feature of enhancers is their ability to exhibit regulatory effects independent of their orientation, functional in both forward and reverse sequential order (Tuan et al. 1997). Enhancers have been implicated in many pathologic conditions including several cancers. Pancreatic cancer and skin cancer risk specifically

have been associated with enhancers in non-coding regions of the genome (*Hurst et al. 2002* and *Takahara et al. 2010*).

Histone Deacytelase 9 (*Hdac9*) and *Twist1*:

Histone deacetylase 9 (*Hdac9*) is a gene located on mouse chromosome 12 and human chromosome 7 that codes for a catalyzing enzyme that removes acetyl groups of histones. *Hdac9* is a member of the histone deacetylase family of genes that all code for enzymes involved in acetyl group removal and transcription regulation (*Kuilenburg et al. 2003*). In skin cancer, polymorphisms in *Hdac9* have been associated with genetic risk through linkage analysis and allelic imbalance studies (*Fleming et al. 2014*). A nearby gene to *Hdac9* in both mice and humans is *Twist1*, an oncogene known to play several roles in tumurogenesis including but not limited to metastasis, apoptosis inhibition and epithelial mesenchymal transition or EMT (*Geng et al. 2009*). Overexpression of *Twist1* has been implicated in several forms of cancer including various carcinomas, and the gene is a potential target for cancer therapeutics (*Fu et al. 2013*). Both *Twist1* and *Hdac9* map to the peak region of linkage for *Skts5 (Mahler et al. 2008).*

A known cis-regulatory association between *Hdac9* and *Twist1* has been shown in mice in the form of an enhancer that has been demonstrated to affect limb development and skin phenotypes *in vivo* (*Ahituv et al. 2012*). In this study, human *Hdac9* DNA mapping to the orthologous 8-9 intron of murine *Hdac9* was cloned into an Hsp68-LacZ enhancer reporting vector. Enhancer assays revealed the *Hdac9* locus was acting as an enhancer for *Twist1,* increasing expression levels in the skin and limbs. Mice lacking this enhancer demonstrated phenotypes of polydactyly and abnormal skin development with partial penetrance. Similar results were observed in zebrafish. The genetic locus of interest showing the proximity of these genes in mice and humans is shown in Figure 1.

FIGURE 1

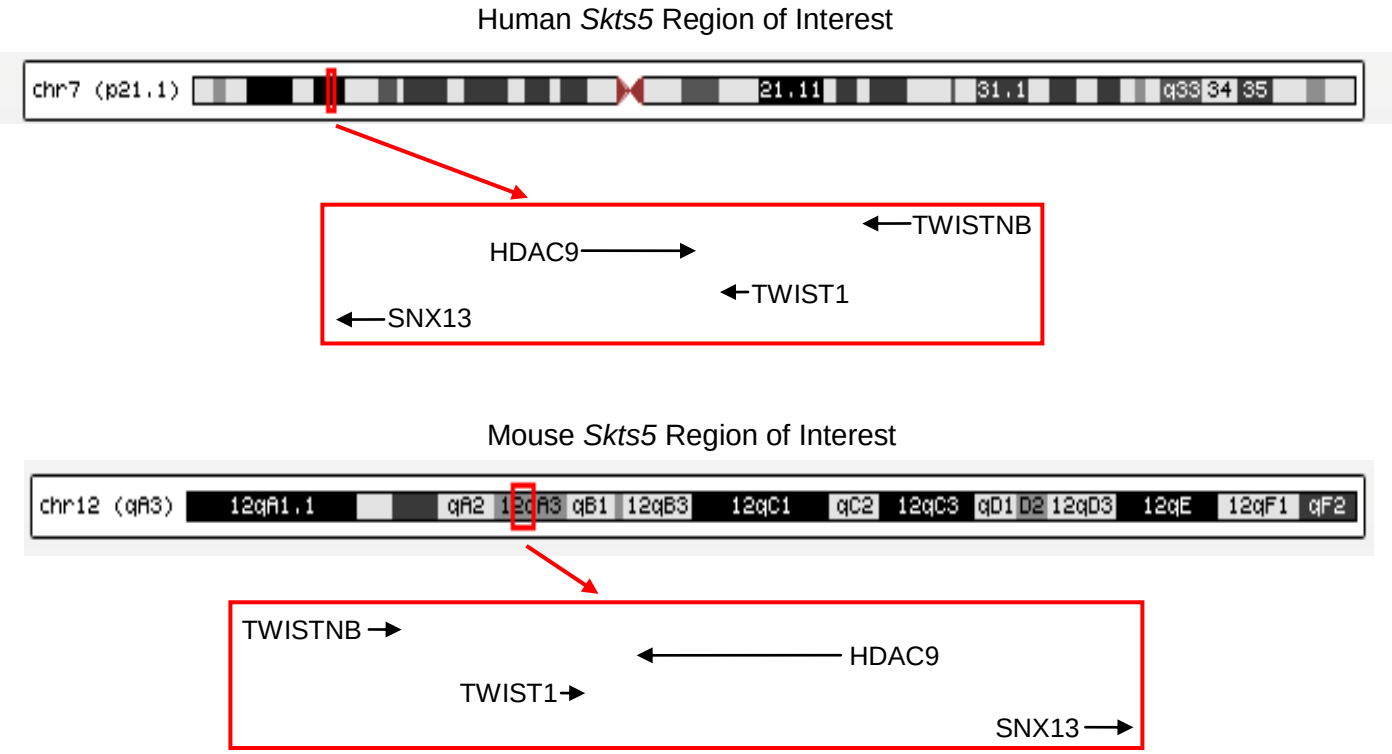

Figure 1: Map of *Skts5* region of interest in humans and mice showing the spatial proximity of *Hdac9* and *Twist1*. Human region displayed is estimated to be 2.2 Mb in length. Mouse region is estimated to be 1.8 Mb in length.

Hypothesis and Specific Aims:

The central hypothesis of my research is that variants in the proposed enhancer region of *Hdac9* between skin cancer resistant *Spretus* and susceptible *NIH/Ola* murine DNA are driving differential gene expression of enhancer target genes, contributing to the differential risk for skin cancer seen between these two mice strains. The specific aims of this project are outlined below:

1. To determine if variants exist between *NIH/Ola* and *Spretus* murine DNA at the proposed enhancer locus in *Hdac9* between exons 8 and 9.
2. To identify and characterize any enhancer sites in the murine *Hdac9* sequence between exons 8 and 9. This will be accomplished by cloning *Hdac9* intron fragments into an enhancer reporting vector, transfecting these constructs into keratinocyte and squamous cell carcinoma cell lines, and using these cells for Luciferase assays.
3. To determine which gene(s) are targets of the enhancer(s) and demonstrate this regulation through quantitative PCR.
4. To identify specific transcription factor/DNA binding interactions potentially associated with enhancer activity. First, *in silico* prediction tools will be utilized to identify potential factors. Then chromatin immunoprecipitation (ChIP) studies will be performed to demonstrate these interactions *in vitro* .

CHAPTER 2: MATERIALS AND METHODS

2.1 DNA Sample Source:

Tail snips for *NIH/Ola* mice were provided by Dr. Allan Balmain and tissues for *SPRET/Outbred* mice were provided by Hiroki Nagase from mice used for other studies as approved by the University of California, San Francisco Institutional Animal Care and Use Committee. DNA was isolated from tails using standard methods (Laird et al. 1991).

2.2 Sequencing:

Intronic and exonic sequences of the *Hdac9* gene corresponding to the published enhancer region were identified using the Ensembl database. We designed PCR primers using Integrated DNA Technology's SciTools PrimerQuest web-based program. PCR products were treated with Exo/SAP-IT to remove single stranded DNA (USB). Automated sequencing of PCR products was conducted on an ABI 3700 by standard methods at the Ohio State Nucleic Acid Shared Resource. Sequences were analyzed using DNAstar 3.0 software and polymorphisms were visually inspected when a nucleotide substitution was indicated. Sequencing primers are available for reference in Table 1.1

2.3 Cloning

Cloning was accomplished using TA Cloning (Graham et al. 1990). *Hdac9* intron 8 DNA was amplified using Polymerase Chain Reaction. Due to the large intron size of over 5000 base pairs, nine primer sets were utilized to break up the intron into smaller fragments of 500-800 base pairs. A model displaying these fragments, categorized as Inserts 1-9, can be seen in Figure 2 below. The vector chosen for the cloning process was the pGL3 enhancer vector complete with an SV40 minimal promoter, a luciferin reporter gene and a blunt restriction enzyme digest site for enhancer insertion (Promega). Sma1 blunt-end

restriction enzyme digestion was utilized to linearize the vector and 2mM dTTP's were added in a ligation to create a thymine base pair overhang. *Hdac9* Insert PCR product was ligated to this modified vector to re-circularize the cloning product. Cloning products were transformed into Stellar Competent Cells according to Protocol PT5055-2 (Clontech). Clones were picked and DNA was isolated using a GeneJET Plasmid Miniprep Kit (Thermal Scientific) and stored at -20°C. All clones were verified using both restriction enzyme digest and sequencing techniques.

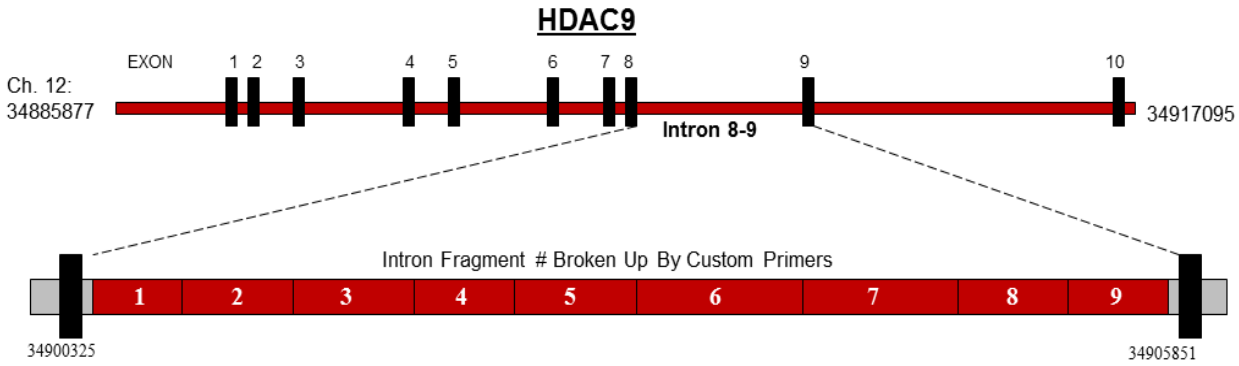

FIGURE 2: Magnified map of the *Hdac9* Intron 8 locus. The intron was divided into 500-800 bp pieces by customer primer PCR into the self-titled "Inserts 1-9" spanning the length of the intron. Primers for these inserts can be found in Table 1.1

2.4 Cell Lines

Two cell lines were used for the experiments in this project. A normal murine keratinocyte cell line, C5N, and a murine cutaneous spindle cell line, A5, were both obtained from Allan Balmain (Zoumpourlis et al. 2003). All cells were grown in Dulbecco's medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen) and 1% penicillin/streptomycin (Invitrogen).

2.5 Transfection

Transient transfections were performed using Lipofectamine reagent according to the manufacturer's protocol. Cells were plated in triplicate and were transfected at 60-80% confluency. Mock transfection and pGL3-control empty vector only transfections were carried out as controls. Cells were also co-transfected with a pRL-TK Renilla reporter gene vector as a normalizing control.

2.6 Luciferase Assays

Cells were allowed to incubate at 37°C in supplemented DMEM (10% FBS, 1% P/S) for 24 hours post-transfection. At this time-point cells were harvested and protein lysate was prepared using M-PER (Pierce Biotechnology, Rockford, IL). 30 µg of each sample was used for luciferase assay analysis, which was performed as described in Skeeles et al. 2013. Illumination measurements were made using the Veritas Microplate Illuminometer. Student t-tests were used to calculate p-values.

2.7 Quantitative Real-Time PCR

At 24 and 48 hours post-transfection, cells were harvested and RNA was isolated using a modified Trizol protocol (Life Sciences Technologies). One microgram of RNA from each sample was reversed transcribed using iScript cDNA synthesis kit (Bio-Rad, Hercules, CA). To assess mRNA expression of *Hdac9* and *Twist1*, Taqman probes were purchased from Applied Biosystems (Life Sciences Technologies). Each sample was measured in triplicate. To measure expression of *Hdac9* and *Twist 1*, mRNA expression was measured at 24 and 48 hours post-transfection of the pGL3-enhancer construct in A5 cells. Taqman probes were from Applied Biosystems/Life Technologies and are as follows *Hdac9* (Mm00458454), *Twist1* (Mm00442036) and *Hprt* (Mm00446968). Control gene *Hprt* was

used to calculate the relative expression of each test gene. Each sample was assessed in triplicate. Student t-tests were used to calculate the p-values for comparisons.

2.8 *In Silico* Predictions for Transcription Factor Binding Sites

Software tools were utilized to identify potential transcription factor binding differences at the *Hdac9* intron 8 Locus between *NIH/Ola* and *Spretus* murine DNA. TFSearch was the primary tool utilized, with a threshold score of 85.0 used as a cutoff for significance. Other databases utilized include PROMO (Transfac), TFSitescan (no longer available) and DBD. Factors given the most importance were those with high threshold scores and agreement between multiple databases.

2.9 Chromatin Immunoprecipitation (ChIP)

Chromatin Immunoprecipitation studies were completed using an altered Millipore ChIP protocol. *Hdac9* transfected cells were grown to a count of one million cells in a 10cm dish 24 hours post-transfection. Histones were cross-linked with a 1% formaldehyde solution in supplemented DMEM (10% FBS, 1% P/S) for 10 min at 37°C. Harvested cells were suspended in 500 µL of SDS lysis buffer and sonicated using the Fisher Scientific Sonic Dismembrator Model 500 at the following conditions: 12 total cycles of 30 second sonication, 60 second rest at 40% amplitude and 4°C. These conditions consistently produced PCR products between 200-400 base pairs. Immunoprecipitation antibodies were ordered for transcription factors Oct1 (Bethyl Laboraties No. A301-716A) and Gata3 (Santa Cruz Biotech HG3-31: sc-268). A no antibody negative control, as well as a non-specific binding negative control antibody for GFP (Santa Cruz Biotech B-2: sc9996) were utilized. Results were displayed both qualitatively and quantitatively, with all quantification being performed using the Alpha Imager software. Significance for all data was determined using a student t-test.

2.10 Primers

TABLE 1.1: Primers for Sequencing and TA Cloning

Amplified DNA	Forward	Reverse
NIH Insert 1	GCACTTGCAAACAACTGAAGGGTG	GGGCAGATCTCCAGCTCATTATTC
Spretus Insert 1	GCACTTGCAAACAACTGAAGGGTG	AAATCACCATTGGCTTTCCCTCGC
NIH Insert 2	CAGATGACACTTGGCAGCTTGTCT	ACTTTGGGCTTCTATTGCTAGCCT
Spretus Insert 2	CAGATGACACTTGGCAGCTTGTCT	ACTTTGGGCTTCTATTGCTAGCCT
NIH Insert 3	AGGCTGTGTTGAAGGCTAGAGA	ATGGGTGTTTCAGCATCTTGTGGC
Spretus Insert 3	AGGCTGTGTTGAAGGCTAGAGA	ATGGGTGTTTCAGCATCTTGTGGC
NIH Insert 4	ATGGGTGTTTCAGCATCTTGTGGC	ACATTCCCTAGCACTGGGCAGAAA
Spretus Insert 4	ATGGGTGTTTCAGCATCTTGTGGC	CTGAGCTTGTTTCTCCATGACAAC
NIH Insert 5	TATTTCCACTTCCCAGAGCACAGC	AGCAAGAGTCAGGGTTGGTTGGAT
Spretus Insert 5	CTTCCCAGTGCTAGGGAATGTA	AGCAAGAGTCAGGGTTGGTTGGAT
NIH Insert 6	CTAGCATTGCTTCAGGATTGGGATG	TTTCCTTTGCATAGCACACACGCC
Spretus Insert 6	TCCATGAACACTCATCTCAAGCA	AAGCCTCATTCCAGTGGATGCTGA
NIH Insert 7	TTCTTACCTGCCTGAGCAGCATGA	GGCTTTCCCACACTTCCACAACTT
Spretus Insert 7	ATACCCTAAAGGTGACGAGCAAGC	GCGCAGGGATTTATGCAATCCAGT
NIH Insert 8	CCTTGTTACCCTCTTTATTGTTTCCA	AACTAGCCTCATTGGCCTCTGGAA
Spretus Insert 8	AGTGTGGGAAAGCCAGAAACCTGA	GAGTGAAATGTACAGGATGCTGGG
NIH Insert 9	TCTGTGAGCCATAGATCCTCCTGA	CCATGGTGAACATCCTACAGGGAA

Note: Primers for NIH and Spretus were made identical when possible. All primers listed in a 5'⟶ 3' direction

Table 1.2: Chromatin Immunoprecipitation (ChIP) Primers

Amplified DNA	Forward	Reverse
Insert 1: Region 1 (base pairs 2-242)	GATTGTAGATCTGGTATGTATGCCT	GATCTGCGCCTTCTTCTCATTA
Insert 1: Region 2 (base pairs 273-472)	TGATACTTCCACAGTTAAATAGCAC	TCACAGTGACATGTTGACTTGTTT

Note: All primers are listed in a 5'⟶3' direction. The "Insert 1" portion of *Hdac9* intron 8 was targeted for ChIP studies for its promise as an enhancer locus from Luciferase Assay and qPCR results. Since multiple transcription factor binding site differences between *NIH* and *Spretus* DNA were predicted in this portion of the intron, the insert was divided into smaller "regions" of 200-250 base pairs for more specified analysis.

CHAPTER 3: RESULTS

Identification of Sequence Variants between *NIH/Ola* and *Spretus* Murine DNA

To investigate potential enhancer activity and its relation to skin cancer risk, all genetic differences within *Hdac9* intron 8 between cancer susceptible *NIH/Ola* and cancer resistant *Spretus* murine DNA were characterized. To accomplish this, the *Hdac9* intron 8 DNA for both mice strains were amplified and sequenced in small portions (Figure 2). Primers for this PCR amplification can be found in Table 1.1. In total, 44 polymorphisms were found between the two mice strains.

TABLE 2

SNP #	Locus of SNP *base pair # in Hdac9 Intron 8*	*NIH* Allele	*Spretus* Allele	Polymorphism
1	158	T	C	T→C
2	334	G	C	G→C
3	414	C	T	C→T
4	491	A	G	A→G
5	534	A	G	A→G
6	594	G	A	G→A
7	743	G	T	G→T
8	744	C	X	C→DELETION
9	797	A	X	A→DELETION
10	811	A	X	A→DELETION
11	1,521	A	G	A→G
12	1,553	A	G	A→G
13	1,622	C	T	C→T
14	1,657	A	C	A→C
15	1,768	T	C	T→C
16	1,918	A	C	A→C
17	2,045	A	G	A→G

TABLE 2 CONTINUED

SNP #	Locus of SNP *base pair # in Hdac9 Intron 8*	*NIH* Allele	*Spretus* Allele	Polymorphism
18	2,048	G	A	G→A
19	2,055	G	A	G→A
20	2,064	G	A	G→A
21	2,072	A	G	A→G
22	2,110	T	C	T→C
23	2,245	T	C	T→C
24	2,344	G	T	G→T
25	2,363	G	A	G→A
26	2,367	X	C	DELETION→C
27	2.451	T	A	T→A
28	2,465	G	A	G→A
29	2,466	G	A	G→A
30	2,559	T	C	T→C
31	2,626	T	G	T→G
32	2,696	A	G	A→G
33	2,792	C	T	C→T
34	2,880	A	T	A→T
35	3,288	T	C	T→C
36	3,290	T	A	T→A
37	3,301	T	C	T→C
38	3,433	G	A	G→A
39	3,438	T	C	T→C
40	3,440	C	G	C→G
41	3,482	CACACA	XXXXXX	CACACA→DELETION
42	3,495	GT	XX	GT→DELETION
43	3,499	AT	XX	AT→DELETION
44	3,502	TATA	XXXX	TATA→DELETION

Hdac9 Intron Demonstrates Enhancer Activity

To determine whether or not enhancer activity was present in *Hdac9* intron 8, fragments of intron DNA were cloned into a pGL3 enhancer reporting vector using TA cloning. The fragments of the intron were broken up into 9 DNA segments using the primers found in Table 1.1 and a visual of these 9 inserts can be found in Figure 2. Of the 9 total inserts, 3 have been successfully cloned into the pGL3 vector and confirmed with sequencing. These 3 inserts include Insert 1, Insert 2 and Insert 7. The vectors containing each of these inserts were transfected into C5N and A5 cell lines and luciferase assays were performed on cell harvests 24 hours post-transfection.

Luciferase assay data for Insert 1 suggests enhancer activity is present in both A5 and C5N cells transfected with the *NIH* Insert 1 construct. An estimated 1.6 fold increase in Luciferin expression in these cells compared to cells transfected with pGL3 control vector is a significant increase in both cell lines (C5N: $p<0.01$, A5: $p<0.01$; Figure 3A,B). Additionally, enhancer activity seems to be found selectively in the *NIH* Insert 1 transfected cells in comparison to the *Spretus* Insert 1 transfected cells. Differences in Luciferin expression were statistically significant in both C5N cells and A5 cells (C5N: $p<0.05$, A5: $p<0.01$; Figure 3A,B).

Luciferase Assay data for Insert 2 suggests there is also enhancer activity present in cells transfected with the *NIH* Insert 2 construct. Activity was slightly less than Insert 1, with an estimated 1.3 fold increase in Luciferin expression in *NIH* Insert 2 cells in comparison to pGL3 control vector transfected cells. Still, these differences were statistically significant according to a student t-test (C5N: $p<0.05$, A5: $p<0.05$; Figure 3A,B). Similar to Insert 1, cells transfected with Insert 2 selectively show enhancer activity in *NIH* Insert 2 cells in

comparison to *Spretus* Insert 2 transfected cells. These differences also hold statistical significance with differences in A5 cells reaching a level of high significance (C5N: $p<0.05$, A5: $p<0.01$; Figure 3A,B).

Finally, data from Insert 7 Luciferase Assays suggest that enhancer activity was not present in this region of the intron. There is no apparent difference in enhancer activity (Figure 3A,B).

Figure 3

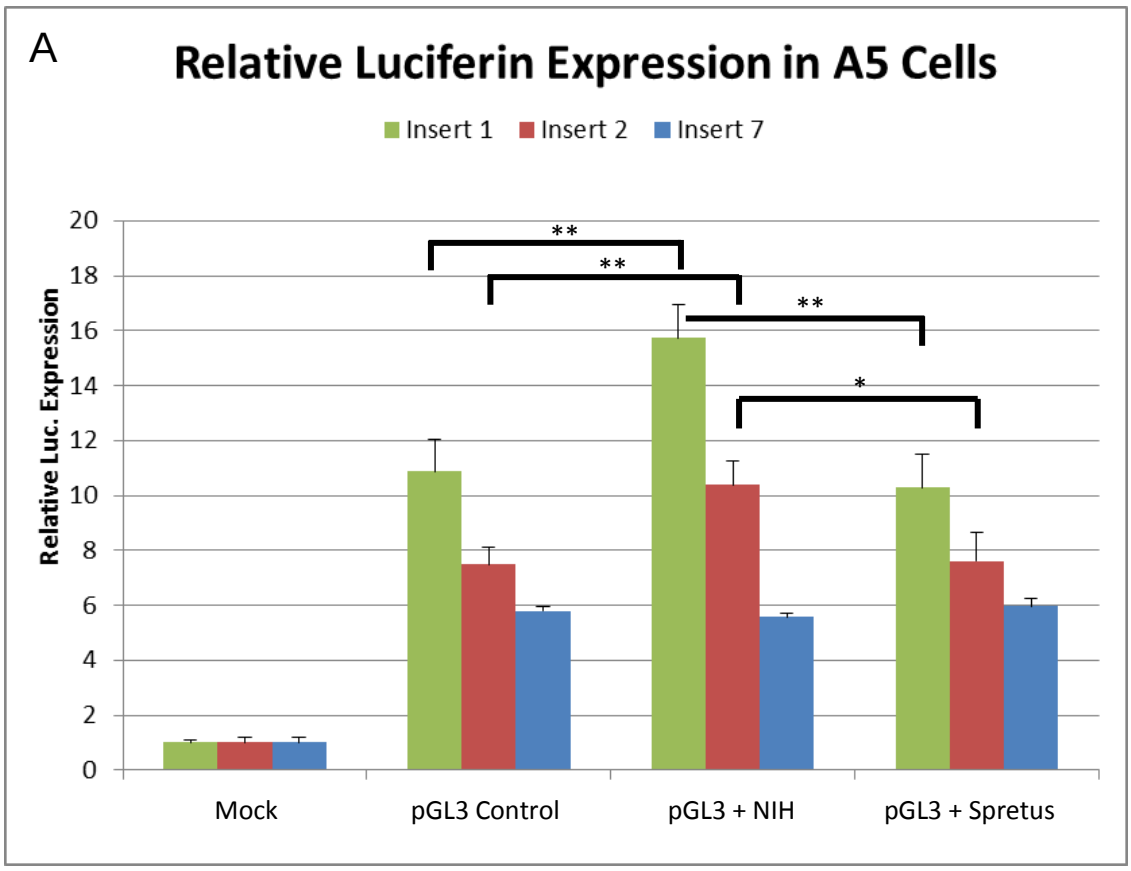

Figure 3A) Luciferase assay data for cells transfected into A5 cells (SCC). Both Inserts 1 and 2 show highly significant statistical differences between both control and *NIH* results (**$p<0.01$). Insert 1 shows highly significant differences between *NIH* and *Spretus* results (**p values <0.01). Insert 2 shows significant differences (*$p<0.05$). There are no significant changes in expression for Insert 7.

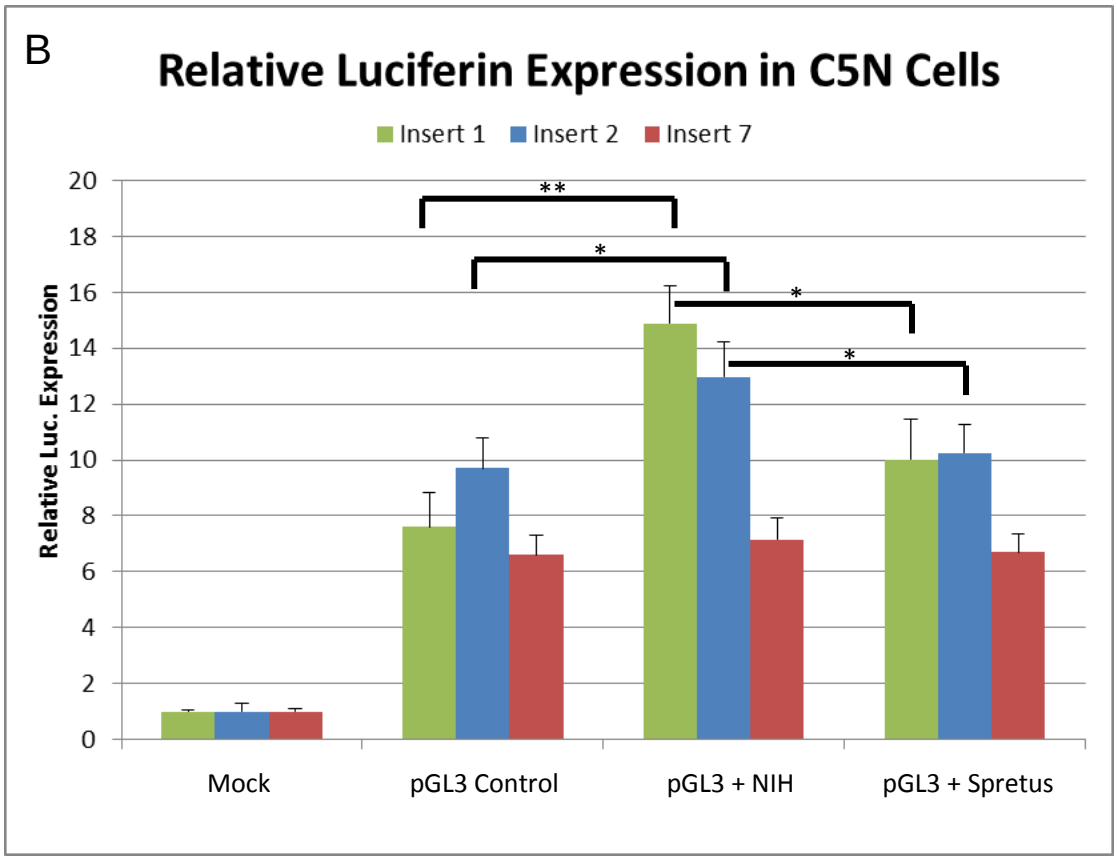

Figure 3B) Luciferase assay data for cells transfected into C5N (normal keratinocyte) cells. Insert 1 shows highly significant differences in expression between control and *NIH* results (**p values < 0.01). Insert 2 shows significant differences between control and *NIH* (*p<0.05). Inserts 1 and 2 show significant difference between *NIH* and *Spretus (*p<0.05).* Insert 7 does not show any significant changes in expression.

Hdac9 Intronic Enhancer Upregulates *Twist1* Expression

To identify target genes of the enhancer activity present in *Hdac9* intron 8, Real time quantitative PCR studies (qPCR) were performed to look for changes in gene expression in the presence of intron DNA constructs. Logical starting genes were *Hdac9* and *Twist1*, due to the increased likelihood that the enhancer acts autonomously or on a nearby gene. As *Twist1* has already been implicated to be a target of *Hdac9* enhancer activity, it was important to include (Ahituv et al. 2012). No differences in expression for *Hdac9* were present between either contruct or controls. It should be noted that expression levels relative to control gene *Hprt* were quite low, possibly indicative of errors in the experimentation process.

Results for *Twist1* expression assays indicate that the *Hdac9* Insert 1 enhancer construct is upregulating expression of *Twist1*. An estimated 2.2 fold increase in expression between pGL3 control vector transfected A5 cells and *NIH* Insert 1 transfected cells is apparent with highly significant differences ($p<0.01$). There also seems to be an estimated 2 fold increase in expression between *NIH* Insert 1 transfected cells and *Spretus* Insert 1 transfected cells, with significant differences in expression ($p<0.05$). Results for qPCR studies analyzing *Twist1* expression relative to *Hprt* control gene can be found in Figure 4.

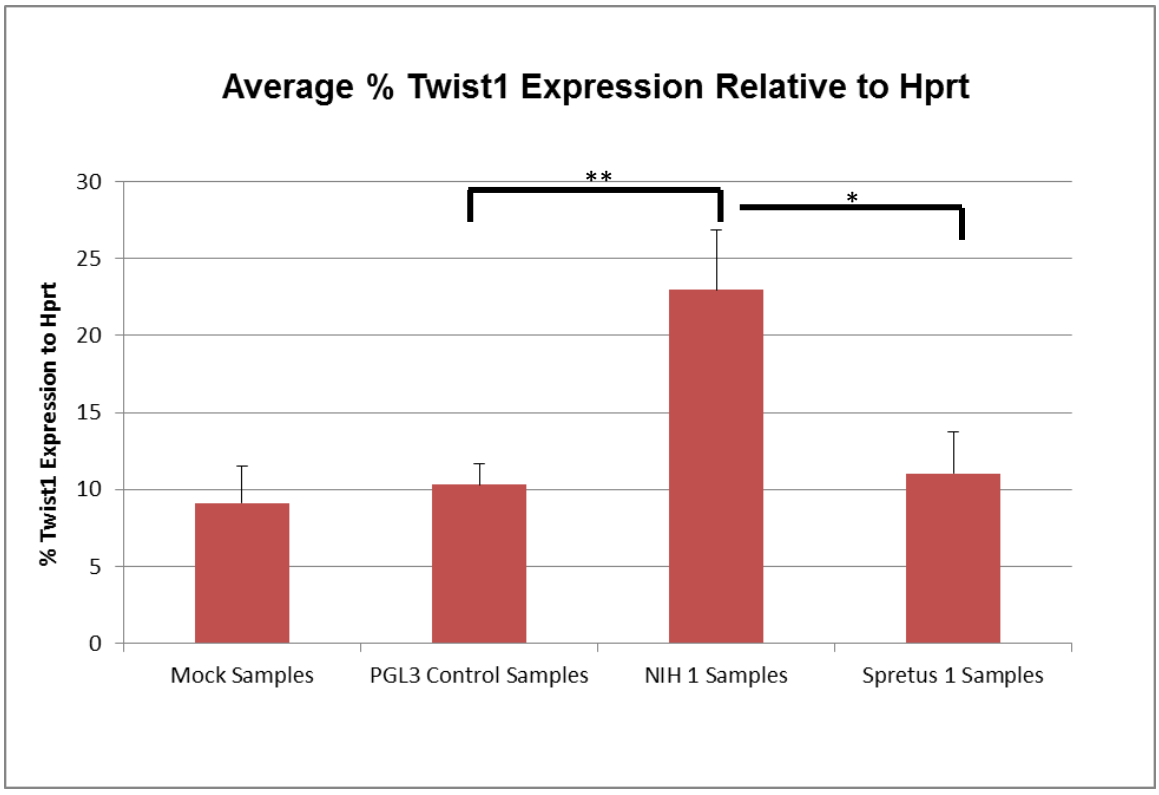

Figure 4) qPCR results for *Twist1* expression relative to *Hprt* control gene. A5 cells transfected with *NIH* Insert 1 contructs show highly significant differences in *Twist1* expression in comparison to pGL3 control (**p<0.01). Significant differences in *Twist1* expression present between *NIH* Insert 1 and *Spretus* Insert 1 transfected A5 cells (*p<0.05). Results indicate *Twist1* is a target for *Hdac9* enhancer activity and may play a role in the differential risk for skin cancer noted between *NIH/Ola* and *Spretus* mice.

In Silico Transcription Factor Binding Site Predictions

To determine which DNA/protein interactions are contributing to enhancer activity, we first utilized bioinformatics tools to predict binding differences between *NIH* and *Spretus* DNA. The primary software tool utilized for this analysis was TFSearch, powered by the Transfac 3.0 database. Other secondary analysis tools utilized include PROMO, DBD, and TFSiteScan (no longer in existence). Of the hundreds of factors initially available as candidates, results were narrowed down through cross-checking across multiple databases. From this list, factors were further narrowed according to their significance in the field of cancer research according to published literature. The top seven transcription factors remaining can be found in Table 3 below along with the polymorphisms that lead each factor to differentially bind. A cutoff score of 85.0 (as determined by TFSearch database) was utilized for significance.

Table 3

Transcription Factor	DNA Binding Site	Polymorphism	Binding Strain	Significance Score
Oct1	ATATACACT	G/C	Spretus	89.7
Gata3	AATCACG	C/T	NIH	85.9
CdxA	ATATAG	G/C	NIH	85.0
Lyf-1	TGGGAT	A/G	Spretus	85.7
Gfi-1	ATAGTTGTGAT	A/G	Spretus	85.9
Nkx-2	TCAAGTG	C/A	Spretus	89.9
AP-1	GTGATTAA	A/G	NIH	86.4

Table 3) *In Silico* predictions of transcription factors that will differentially bind *NIH* and *Spretus* DNA at the enhancer locus in *Hdac9*. Cutoff score of 85.0 used for significance. These factors will be ideal candidates for ChIP studies to confirm predictions *in vitro*.

Gata3 and Oct1 Transcription Factors Bind *Hdac9* Enhancer Differentially *in vitro*

To test whether or not *in silico* transcription factor binding predictions actually occurred in *in vitro* cell models, chromatin immunoprecipitation (ChIP) studies were performed utilizing antibodies against the Gata3 and Oct1 transcription factors. These factors were selected from Table 3 for their binding site loci occurring within Insert 1 of *Hdac9* insert 8, which displayed the most enhancer activity in luciferase assay results. ChIP results for the Gata3 factor indicate that Gata3 selectively binds to the *NIH/Ola* strain of murine DNA in comparison to *Spretus* DNA and controls. An observable increase in DNA binding the Gata3 transcription factor in Region 2 of Insert 1 can be seen in C5N cells (Figure 5A). An observable difference in Gata3 binding between the two mice strains was not noticeable in Region 1 of Insert 1 (Figure 5A).

ChIP results for the Oct1 factor indicate that Oct1 selectively binds the *Spretus* strain of murine DNA in comparison to *NIH/Ola* controls. An observable increase in DNA binding the Oct1 transcription factor can be seen in Regions 1 and 2 of Insert 1 in C5N cells (Figure 5B). ChIP experiments in C5N cells confirm the predicted transcription factor binding results from the bioinformatics tools utilized in this project.

Figure 5

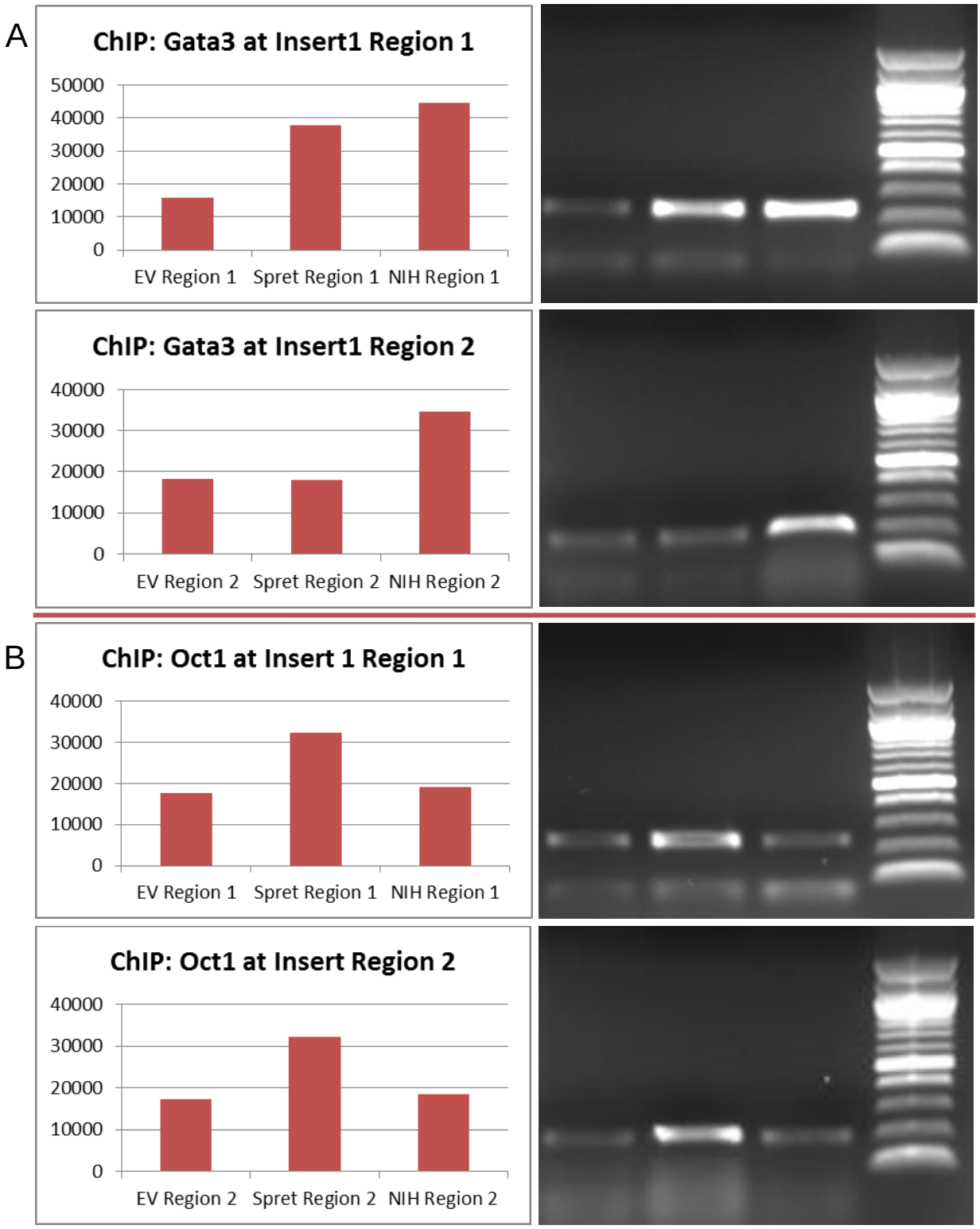

Figure 5: ChIP results for Gata3 and Oct1 transcription factors. Gata3 shows significant preferential binding to *NIH* DNA in Insert 1 Region 2. Oct1 shows significant preferential binding to *Spretus* DNA in Insert 1 Regions 1 and 2

CHAPTER 4: DISCUSSION

We hypothesized that variants in a proposed enhancer locus at *Hdac9* intron 8 between cancer susceptible *NIH/Ola* and cancer resistant *Spretus* murine DNA are playing a role in the differential risk for NMSC seen between these two mice strains. To support this hypothesis, experiments were conducted to both identify genetic differences at this locus and determine whether these differences had any effect on enhancer activity in normal keratinocyte and SCC cells. Furthermore, we hoped to elucidate the mechanism for this difference in risk by identifying genetic targets of this enhancer and determining what protein/DNA interactions in the cell nucleus were involved in any differential enhancer activity between the mice strains. In summary, this study had four major aims:

1. To determine if variants exist between *NIH/Ola* and *Spretus* murine DNA at the proposed enhancer locus in *Hdac9* between exons 8 and 9.
2. To identify and characterize any enhancer sites in the murine *Hdac9* sequence between exons 8 and 9. This will be accomplished by cloning *Hdac9* intron fragments into an enhancer reporting vector, transfecting these constructs into keratinocyte and squamous cell carcinoma cell lines, and using these cells for Luciferase assays.
3. To determine which gene(s) are targets of the enhancer(s) and demonstrate this regulation through quantitative PCR.
4. To identify specific transcription factor/DNA binding interactions potentially associated with enhancer activity. First, *in silico* prediction tools will be utilized to identify potential factors. Then chromatin immunoprecipitation (ChIP) studies will be performed to demonstrate these interactions *in vitro* .

Hdac9 Variants Alter Enhancer Activity between *NIH* and *Spretus* DNA

Based on our results, clear genetic variation at *Hdac9* intron 8 exists between cancer susceptible *NIH* and cancer resistant *Spretus* murine DNA. In total, 44 genetic variants were discovered over the span of a genetic sequence roughly 5,000 base pairs in length. Due to the nature of this sequence being intronic, none of these variants are found in known protein-coding regions of the genome. However, the potential still exists for these variants to be involved in cis-regulatory elements such as enhancers.

Previous studies have identified enhancer activity in *Hdac9*. This finding was demonstrated using enhancer assays, in which vectors containing human *Hdac9* sequence were cloned into mouse and zebrafish models. In the presence of the human *Hdac9* sequence mapping to the orthologous mouse *Hdac9* intron 8, both mice and zebrafish clearly showed enhancer activity using LacZ and GFP reporter genes, respectively. Further experimentation showed this *Hdac9* enhancer upregulated neighboring gene *Twist1,* with significant potential phenotype implications. In the presence of the *Hdac9* enhancer, both mice and zebrafish exhibited limb malformations (such as polydactyly) and defects in both anterior and posterior limbs (Ahituv et al. 2012). Interestingly, although Hdac9 is not expressed highly in the limbs, *Hdac9* knockout mice show a similar phenotype to *Twist1* knockout mice, indicating enhancer activity in *Hdac9* may be is essential for sufficient *Twist1* expression during development.

In addition to its functionality as an enhancer for *Twist1*, *Hdac9* has also been implicated to be part of a skin tumor susceptibility locus known as *Skts5*. Linkage analysis studies in mice have identified variations resulting in amino acid differences in *Hdac9* associated with differential risk for skin cancer between *NIH/Ola* and *Spretus* mice. However, no mechanism has been suggested for how these variants affect NMSC genetic risk.

A correlation between *Hdac9* enhancer activity and the gene's role in differential risk for NMSC has never been investigated. This study attempts to find a link between enhancer activity in *Hdac9* and the role *Hdac9* plays in the differential risk for skin cancer with the hope of elucidating a mechanism of *Skts5*'s relation to NMSC risk. To establish this correlation, first enhancer activity had to be demonstrated in our genetic locus within mouse *Hdac9,* as was found in the orthologous genetic locus of human *Hdac9*. Luciferase assay data clearly suggests enhancer activity is indeed present in the Insert 1 and Insert 2 regions of *Hdac9* intron 8. This data establishes that the mice strains utilized in this study display similar enhancer activity to previous studies utilizing human *Hdac9* in mice and zebrafish (Ahituv et al. 2012). While this connection is important, the more notable takeaway from these assays is that enhancer activity significantly varies between *NIH/Ola* and *Spretus* transfected cells. While not proof of causation, this finding is instrumental to the hypothesis that enhancer activity is associated with NMSC differential risk.

To further investigate the potential existence of a link between *Hdac9* enhancer activity and this gene's role in NMSC risk, Insert 1, the most promising locus for enhancer activity studied, was analyzed further to identify gene targets of the enhancer. While other inserts were not ruled out as potential enhancer loci, Insert 1's proven role in enhancer activity made it ideal for further experiments.

Hdac9 **Intronic Enhancer Upregulates** ***Twist1*** **Expression**

Based on qPCR results, *Hdac9* enhancer does not autonomously regulate *Hdac9* expression. However, the neighboring gene to *Hdac9, Twist1*, is clearly upregulated in the presence of *NIH/Ola Hdac9* Insert 1. Interestingly, this upregulation of *Twist1* is not found in the presence of *Spretus* Insert 1 or related controls. The discovery that *NIH*

Hdac9 selectively increases expression of *Twist1* has significant implications for this study. *Twist1* overexpression has been linked to several cancer phenotypes including apoptosis inhibition, increased cell proliferation and increased epithelial mesenchymal transition (EMT) leading to increased metastasis (Geng et al. 2009). If cancer susceptible *NIH/Ola* mice are overexpressing a known oncogene that cancer resistant *Spretus* mice are not overexpressing, this suggests a possible mechanism for the reduced rates of NMSC development in *Spretus* mice. Additionally, this provides strong support for our hypothesis that enhancer activity in *Hdac9* is directly correlated to *Hdac9*'s role in differential NMSC risk.

While this discovery is important, it does not elucidate how genetic variants are affecting the expression of *Twist1*. To understand this relationship, specific protein and DNA binding interaction differences between *NIH* and *Spretus* would have to be demonstrated.

Gata3 and Oct1 Transcription Factors Differentially Bind *Hdac9* at Insert 1

Gata3 and Oct1 were selected from Table 3 as the two top candidate transcription factors for ChIP studies. These factors have both been implicated previously in risk for several cancers including skin cancer and breast cancer (Nielsen et al. 2008, Sturm et al. 1993). Additionally, the binding sites for both of these factors occur in Insert 1 of *Hdac9* intron 8, the insert showing most enhancer activity in luciferase assays (Figure 3).

According to the ChIP results presented here, Gata3 transcription factor (TF) selectively binds *NIH* murine DNA at the Insert 1 Region 2 predicted binding site *in vitro.* The role of Gata3 in skin cancer is poorly understood, however it is possible Gata3 binding to *NIH* DNA is somehow essential for Hdac9 enhancer activity. This hypothesis would suggest

the lack of Gata3 in *Spretus* mice inhibits enhancer functionality. More work must be done in this area to fully understand the role of Gata3 in this mechanism of NMSC risk. ChIP results also indicate the Oct1 TF selectively binds *Spretus* murine DNA at the Insert 1 Regions 1 and 2 predicted binding sites *in vitro*. The potential mechanism for Oct1's role in differential NMSC risk is much more concrete, with evidence supported by previously published studies.

One such study has found that several Oct1 binding sites exist in the promoter region of *Twist1* (Lee et al. 2013). This study found Oct1 to be part of a repressive binding complex with PER2 and other factors including EZH2 and HDAC2. The Oct1-mediated complex binds the *Twist1* promoter, decreasing expression of *Twist1*. This mechanism would suggest that the selective binding of Oct1 in *Spretus* mice prevents upregulation of *Twist1*, resulting in reduced expression of *Twist1* in comparison to wild-type *NIH/Ola* mice. This trend is exactly what was observed in this project's qPCR data (Figure 4). This repression of *Twist1* in cancer resistant *Spretus* mice would have significant implications for NMSC risk, considering *Twist1* overexpression is associated with cancer phenotypes such as increased cell proliferation, inhibition of apoptosis and increased EMT (Geng et al. 2009).

Future Directions and Study Limits

The majority of experiments in this project focus on enhancer activity demonstrated exclusively in Insert 1 of *Hdac9* intron 8. A major reason for this narrowed focus has been the difficulty of the cloning process, for which only 3 of 9 inserts have successfully been cloned. In the future, if more inserts could be successfully cloned, qPCR and ChIP studies could be done on additional inserts, limited only by the cost of future Taqman

probes and ChIP antibodies. Additionally, it could be beneficial to design primers for later portions of *Hdac9* exon 8 to see if enhancer activity extends into the nearby exon.

Only two transcription factors were analyzed in these ChIP studies. In future experiments, it could be enlightening to investigate other TFs predicted to differentially bind *in silico* such as AP-1 and Lyf-1 (Table 3).

Two cell lines were used in this study, A5 and C5N cells, however in some experiments, errors in experimental procedure prevented experiment completion in one of the two cell lines. Future experiments could repeat these studies in the other cell line to confirm results.

Finally, this study only investigates the activity of this enhancer on genes nearby to *Hdac9*. Enhancers have the capability of affecting gene expression of genes quite far from the enhancer site, so other genes could be investigated. C4 studies could be conducted to evaluate this potentiality. If more genes are identified as possible targets, further qPCR studies could be performed.

A fundamental limitation of this study is that it does not investigate any other genes or regulatory regions in *Skts5* that may also play a role in skin cancer susceptibility.

CONCLUSION

In summary, I found that variants in *Hdac9* intron 8 play a role in an intronic enhancer found to upregulate the expression of *Twist1.* This differential expression of *Twist1* between *NIH/Ola* and *Spretus* murine DNA can potentially be explained by the differential binding of certain transcription factors including but not limited to Oct1 and Gata3. These newly discovered interactions may help explain the role of *Hdac9* and *Skts5* in NMSC risk.

WORKS CITED

1. The Skin Cancer Foundation. Skincancer.org/skin-cancer-information/skin-cancer-facts. 2014.
2. University of California San Francisco Department of Dermatology. Dermatology.ucsf.edu/skincancer. 2007.
3. Tekawa IS, Friedman GD. Association of basal cell skin cancers with other cancers (United States). Cancer Causes Control 2000; 11(10):891-7.
4. National Cancer Institute at the National Institute of Health. Cancer.gov/cancertopics/types/skin. 2014.
5. Mahler et al. Sequence divergence of Mus spretus and Mus musculus across a skin cancer susceptibility locus. BMC Genomics 2008; 9: 626.
6. Fleming et al. Allele-specific imbalance mapping identifies HDAC9 as a candidate gene for cutaneous squamous cell carcinoma. Int J Cancer 2014; 134(1)244-8.
7. Tuan D, Kong S, Bohl D. Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation position, and distance of the enhancer relative to the gene. Mol. Cell. Biol. 1997; 17(7):3955-3965.
8. Hurst HC, Shah RN, Ibbitt JC, Alitalo K. FGFR4 overexpression in pancreatic cancer is mediated by an intronic enhancer activated by HNF1α. Oncogene. 21(54)8251-61.
9. Takahara et al. An intronic enhancer driven by NF-kB contributes to transcriptional regulation of peptidylarginine deiminase type 1 gene in human keratinocytes. J Inv. Derm 2010; 130:2543-2552.
10. Kuilenburg AB, Kemp S, Caron HN, van Gennip AH, de Ruijter AJ. Histone deacetylases (HDAC's): characterization of the classic HDAC family. Biochem J. 2003; 370(3):737-49.
11. Geng JS, Yu XW, Shi QT, Wang K, Feng MY. Gene expression profiling in Twist1-depleted gastric cancer cells. Anat Rec. 2009; 292(2):262-270.

12. Fu J, Khan MA, Chen HC, Zheng D. Twist1: A molecular target in cancer therapeutics. Tumour Biol 2013; 34(5):2497-2506.
13. Ahituv et al. Coding exons function as tissue-specific enhancers of nearby genes. Genome Res 2012; 22(6):1059-68.
14. Laird PW, Zijderveld A, Linders K, Rudnicki MA, Berns A. Simplified mammalian DNA isolation procedure. Nucleic Acids Res 1991; 19(15)4293.
15. Graham MW, Holton TA. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. Nucleic Acids Res 1990; 19(5):1156.
16. Skeeles et al. The impact of 3' UTR Variants on differential expression of candidate cancer susceptibility genes. PLOS Onc 2013; 8(3).
17. Zompourlis et al. Alteration in signal transduction pathways implicated in tumour progression during multistage mouse skin carcinogenesis. Carcinogenesis 2003; 24(7): 1159-65.
18. Nielsen T, Cheang M, Voduc D. Gata3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer Epi Biomarkers Prev 2008; 17:365.
19. Sturm RA, Thomson JA, Parsons PG. In vivo and in vitro expression of octamer binding proteins in human melanoma metastasis, brain tissue and fibroblasts. Pigment Cell Res 1993; 6(1): 13-22.
20. Lee et al. Loss of corepressor PER2 under hypoxia up-regulates OCT1-mediated EMT gene expression and enhances tumor malignancy. PNAS 2013; 110(30): 12331-12336.