Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

CHARACTERIZATION OF THE CANINE *DLC1* GENE AND ASSAY FOR SP1 BINDING

by

Mohamed Jabbes

A thesis

Submitted in partial fulfilment

Of the requirement for the degree of

Master of Science in the Department of Medical Laboratory Sciences

Idaho State University

Spring 2014

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Mohamed Jabbes find it satisfactory and recommend that it be accepted.

_____

Kathleen M Spiegel

Major Advisor

_____

Susan Galindo

Committee Member

_____

Gene Scalarone

Graduate Faculty Representative

Dedication

This thesis is dedicated to my wife, my son, and my daughter for their sacrifice and support.

ACHNOWLEGMENTS

I thank Dr. Jeffrey N. Bryan for funding and the opportunity to work on this project at the Comparative Oncology Laboratory, Veterinary Clinical Sciences, Washington State University. I Thank my committee members, Dr. Kathleen Spiegel, Susan Galindo, and Dr. Gene Scalarone for their help, support, and reviewing this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Deleted in liver cancer one (*DLC1*) is a tumor suppressor gene that is frequently inactivated by aberrant CpG methylation and mutations in multiple tumor types (Yuan et al. 1998, Shikhar et al. 2009, Esteller 2002, Bryan et al. 2009). We isolated for the first time the full length cDNA of canine *DLC1* (GenBank accession number FJ602870), and compared it to human *DLC1* and mouse *DLC1*. The transcript of the full-length cDNA of canine *DLC1* was 3724 bases, organized into 14 exons encoding a 1091-amino acid (aa) polypeptide with a molecular weight of about 123 KDa. The canine *DLC1* is located on the long arm of chromosome 16 (16:39,491,300-39,535,640). The canine *DLC1* nucleotide and amino acid sequences have high homology to the human *DLC1*, with 88% and 95% identity values, respectively.

The amino acid sequences of canine *DLC1*contains four major domains that were previously identified in human and mouse (Yuan et al. 1998, Durkin et al. 2002): 1) An N-terminal sterile alpha motif (SAM) domain; 2) a RhoGAP domain, a conserved region responsible for the catalytic activity of RhoGAPs; 3) a C-terminal steroidogenic acute regulatory protein (STAR)-related lipid-transfer (START) domain; and 4) a serine-rich (SR) domain located between the SAM domain and RhoGap domain.

To determine that the transcription factor Sp1 binds to the promoter region of canine *DLC1*, we carried out western blotting using anti-human-Sp1 antibody. Western blot analysis confirmed that the antibody used for the ChIP reacts with the canine protein specificity protein 1 (Sp1) as expected. In addition, *DLC1* promoter analysis using MatInspector program included multiple Sp1-binding sites suggesting that Sp1 might be

involved in *DLC1* expression. Canine Sp1 binding sites are conserved and show similarity of transcription control to that of the human counterpart (Bryan et al. 2009).

Introduction

Deleted in liver cancer one (*DLC1*) is a tumor suppressor gene found to be deleted in a variety of cancers particularly liver, breast, prostate, and colorectal cancers (Yuan et al. 1998, Liao and Lo 2007, Goodison et al. 2005, Thangapazham et al. 2014). The gene was initially identified by Yuan and coauthors in a primary hepatocellular carcinoma (HCC) and HCC cell lines using representational difference analysis (RDA), a PCR-based method (Yuan et al. 1998). *DLC1* was mapped to the human short chromosome 8p22 region and named as deleted in liver cancer (*DLC1*) because it was found to be often deleted in primary hepatocellular carcinoma (HCC) and HCC cell lines (Yuan et al. 1998). Because the human *DLC1* on the short arm of chromosome 8 was found to be deleted in many cases of cancer including liver, breast, lung, and prostate cancers (Brotherman 1997, Wong et. al. 2003, Isola 1995, Marchio et al. 1997) but present in normal cells or tissues, it has been often speculated that *DLC1* is a tumor suppressor gene (Yuan et al. 1998). In addition, *DLC1* expression has been found to be decreased or undetectable in many prostate and colon cancers (Liao et al. 2008).

A tumor suppressor gene is generally defined as a gene whose product normally inhibits cell proliferation and tumor initiation and progression (Rivlin et al. 2011, Sherr 2004). According to Xue et al. (2008) "Tumor suppressor genes act in signaling systems that protect against tumor initiation and progression, and can be inactivated by deletions, point mutations, or promoter hypermethylation". In normal cells, tumor suppressor genes help stabilize cellular DNA and help prevent uncontrolled cell growth and proliferation. However, harmful mutations, such as in BRCA1 and BRCA 2 tumor-suppressor genes, can produce a hereditary predisposition for breast or ovarian cancer in women (Elledge

and Amon 2002). The tumor suppressor function of *DLC1* has been confirmed by demonstrating that the loss of *DLC1* expression resulted in hepatocellular carcinoma formation, based on a knock-down mouse model (Xue et al. 2008).

Epigenetic changes are also as important in the process of cancer development (Shikhar et al. 2009, Esteller 2002, Bryan et al. 2009). In many types of cancer, tumor suppressor genes are frequently down-regulated or silenced by promoter hypermethylation or histone deacetylation (Jones and Baylin 2007, Esteller 2002). Epigenetic modifications of DNA include the covalent addition of methyl groups to nucleotides and the removal of acetyl groups from histones that play regulatory roles in gene expression (Esteller 2002, Shikhar et al. 2009). One important form of DNA methylation is a chemical modification that occurs when a methyl group is added at the C-5 position of cytosine in a CpG dinucleotide motif following DNA replication to produce 5-methylcytocine (5-mC). These modifications can become heritable and sustained in subsequent generations (Hackenberg et al. 2006). As an example of this process, DNA methyltransferase 1 (DNMT1) recognizes newly replicated DNA (hemi-methylated DNA) containing an old methylated strand and a new unmethylated strand and places methylation on the newly replicated CpGs (Goll and Bestor 2005).

In mammals, clusters of CpG dinucleotides in GC rich regions are called CpG islands (Novik et al. 2002). These CpG islands are rare and unevenly distributed throughout mammalian genomes (Bartolomei and Ferguson-Smith 2011, Bird 2002). CpG rich areas or islands are located at the 5' end of the gene in the promoter region (upstream from the transcription start site) and represent an important feature in mammalian genomes (Gardiner-garden and Fomme, 1987, Hackenberg et al. 2006).

The number of CpG islands in human (27,000) has been found to be much higher than in mouse (15,500) and rat (15,975) genomes, even though all of these genomes code for a similar number of genes (Gibbs et al. 2004, Venter et al. 2001). In contrast, the number of CpG islands in the dog (58,000) has been found to be much higher than those in human and mouse (Han and Zhao 2009). In mammals, these CpG islands are typically unmethylated at all stages of development so that gene transcription can occur; however, heavy methylation of cytosines within a promoter CpG island often results in little or no expression of the associated gene (Greger et al. 1989, Yoder et al. 1997, Shikhar et al. 2009). In human cancers, genes critical to slowing cell growth can be silenced by hypermethylation.

Published data suggest that DNA-binding sites for the transcription factor specificity protein 1 (Sp1) are maintained in an unmethylated state within regulatory regions of genes (Mancini et al., 1999). Mancini and coauthors (1999) evaluated the methylation status of CpG dinucleotides within five promoter subregions in the human and mouse homologues of the neurofibromatosis (*NF1*) genes. They found that three 5' subregions were consistently methylated in all tissue analyzed, but DNA methylation was absent in the vicinity of the transcription start site bounded by Sp1 recognition sequences.

DLC1 is an important tumor suppressor gene that is under-expressed in many types of cancer in both rats and humans (Mancini et al., 1999, Greger et al. 1989, Yoder et al. 1997, Shikhar et al. 2009). Data generated by our laboratory demonstrate hypermethylation of *DLC1* in canine lymphoma; however, expression was not silenced (Bryan et al. 2009).

Objectives

The objectives of these experiments were to:

1) Characterize the canine DLC1 gene,

2) Compare it to the human and mouse, and

3) Assay of SP1 binding within the promoter region.

Materials & Methods

Sample Collection

Canine spleen samples (~200 mg) were obtained from five dogs meeting ethical standards and following Institutional Animal Care and Use Committees (IACUC) requirements, and immediately placed in 5 volumes of RNA stabilization solution (RNAlater, Ambion, Austin, TX, USA), and stored at -20°C.

Total RNA Extractions

Because RNA extraction is a critical step in the highly sensitive quantitative RT-PCR analysis, four RNA extraction kits were used for preliminary tests, including Ambion® RiboPure (Life Technologies, Carlsbad, CA, USA), Ambion® RNAqueous (Life Technologies, USA), RNeasy mini kit (Qiagen, USA), and TRIzol (Invitrogen, USA). Ambion® RiboPure™ was selected based on the quality and quantity of extracted RNA (results not shown) and ease of use.

Spleen samples were removed from RNAlater®, ground in liquid Nitrogen using a precooled mortar and pestle, and RNA was extracted using Ambion® RiboPure™ extraction Kit following the manufacturer specifications. The frozen powder was mixed immediately with 1 mL TRI buffer (to make 5% W/V homogenate) containing guanidinium thiocyanate. The homogenate was transferred into RNase-free, 1.5-mL microcentrifuge tubes (supplied in the kit) and quickly homogenized by vortexing at maximum speed. The homogenate is then mixed well by vortexing at maximum speed with 200 µL chloroform solution, incubated for 5 minutes at room temperature, and centrifuged at 12,000 g for 30 seconds. The aqueous phase (contains the RNA) was

recovered and mixed with 200 μL 100% ethanol, then it was applied to a glass-fiber filter. The RNA that bound to the glass-fiber filter was washed twice with 500 μL of wash solution to remove residual contaminants. Purified RNA was eluted from the filter in 100 μL Elution Buffer. An rDNase I treatment was performed to remove contaminating genomic DNA from the final preparation.

DNase Treatment

Genomic DNA contaminants were eliminated by treatment with rDNase I using DNA- free™ kit (Life Technologies). Total RNA (10 μg) was treated with rDNAse I consisting of 1 μL (2 units) rDNase I and 5 μL of 10X DNase I buffer. The reaction mixture was incubated in a thermocycler at 37 ^{o}C for 30 minutes, then 5 μL of inactivation reagent was added to terminate the reaction and precipitate the rDNase. The mixture was centrifuged at 10,000 g for 1.5 minutes and the RNA was transferred to a new RNase- and DNase-free microcentrifuge tube. The RNA was stored at -80 °C until use.

RNA Quality

RNA preparation and handling proceeded carefully to assure high-quality mRNA that was not degraded by ribonucleases. The combination of RNA extraction using RiboPure™ kit and rDNase I treatment with DNA-free™ kit produced very high quality RNA that was free of DNA contaminants. RNA quality was assessed by the $OD_{260/280}$ nm absorbance ratio and Experion™ RNA High Sense analysis (Bio-Rad, USA). Total

RNA isolated had an $OD_{260/280}$ ranging from 1.9-2.0, which indicated that the RNA preparation was largely free of proteins and other contaminants that adversely impact subsequent applications. In addition, the quality of total RNA was tested by performing RT-PCR using gene- specific primers followed by gel electrophoresis of the PCR products to determine the amplifiable sizes of the purified RNA. RNA yields were determined using spectrophotometer at OD 260 nm (NanoDrop 2000, Thermo Scientific, USA) following the manufacturer protocol.

Reverse Transcription:

Three cDNA synthesis kits were used in a preliminary test: Iscript™ cDNA synthesis (Bio-Rad), SuperScript® III RT (Invitrogen, USA), and High Capacity cDNA Synthesis Archive Kit (Applied Biosystems, USA). The Superscript® III RT (Invitrogen, USA) was selected for further use based on quality of cDNA products (results not shown). Total RNA was reverse transcribed by both Oligo (dT) and Random priming in a reaction volume of 50 μL using SuperScript® III RT following the manufacturer's protocol. The reaction mixture included 10 μL 2X RT reaction mix, 2 μL enzyme mix , and 1 μg of RNA. The mixture was incubated at 25 ^{o}C for 10 minutes, then at 5 ^{o}C for 40 minutes, and at 85 ^{o}C for 5 minutes. Any RNA contaminant (cDNA:RNA hybrid molecules produced after first strand synthesis) was removed using 1 μL of RNase H and samples were incubated at 37°C for 20 minutes. The cDNA was stored at -80 ^{o}C until use.

Primer Design and Validation

Primer design parameters were selected to enhance organism specificity, gene specificity, amplicon efficiency, uniformity of primer Tm values, elimination of primer-dimer artifacts, detection of single bands of expected size via gel electrophoresis, and the generation of single peak dissociation curve by qRT-PCR. Gene-specific primers were derived from the canine sequences obtained from boxer dog predicted sequences and human DLC1 sequences and used in a previous study in our laboratory (http://genome.ucsc.edu/cgi-bin/hgGateway) and human DLC1 (AF026219). Gene-specific primers were designed using Primer Express software version 2.0 (Applied Biosystems, Foster City, CA, USA) following the manufacturer general guidelines. Candidate primer sets were further analyzed using OligoAnalyzer 3.0 (Integrated DNA Technologies, USA) and primers sets were selected that had low probability to form inter and intra-molecular dimers.

RNA Ligase-mediated Rapid Amplification of cDNA Ends (RLM-RACE) and Sequencing of *DLC1* cDNA.

The FirstChoice® RLM-RACE (rapid amplification of cDNA ends) (Invitrogen, Inc. USA) kit was used to isolate the full-length mRNA of the canine homolog of *DLC1*. Polymerase chain reaction (PCR) was used to amplify (from mixtures of mRNA) regions between known sequences of a target gene and tags appended to the 5' and 3' ends of the cDNA (Schaefer, 1995).

3' RLM RACE

First Strand cDNA from canine total RNA was synthesized using a 3' adapter supplied by the manufacturer (Invitrogen, Inc. USA). A total of 20 μL reaction mix composed of 1μg total RNA, 4 μL dNTP mix, 2 μL 3' RACE adapter (table 2), 2 μL 10X RT buffer, 1 μL RNase H inhibitor, 1 μL M-MLV Reverse Transcriptase was incubated at 42^{o}C in an MJ thermocycler (MJ Research, USA) for 1 hour. The cDNA product was stored at -20 ^{o}C until use.

PCR

5' gene-specific primers (GSP) for *DLC1* were designed from highly conserved regions of dog boxer predicted sequences or human DLC1 sequences and used for amplifying canine cDNA. GSPs were selected in the range of 400-500 b from the 3' end of the RNA. Two nested PCRs were performed using canine *DLC1* GSP on the 5' end with a 3' RLM-RACE outer primer (Table 1). The reaction mixture was subjected to the following cycles: a denaturation step at 96 ^{o}C for 2 minutes, followed by 40 cycles of 96^{o}C for 30 seconds, 60 ^{o}C for 30 seconds (annealing), 72 ^{o}C for 30 seconds (extension), and a final extension step of 72 ^{o}C for 7 minutes, and kept at 4 ^{o}C. PCR products were separated by gel electrophoresis and the correct size fragment was purified and sequenced with automated fluorescent dideoxy method (see sequencing). The resulting sequence data were used to design canine *DLC1* 3' gene-specific primers for the 5' RACE.

5' RLM RACE

The 5' RACE was optimized by first selecting full-length, capped mRNA from

total RNA (Maruyama and Sugano, 1994). Total RNA was treated with Calf Intestine Alkaline Phosphatase (CIP) to remove free 5' phosphates from rRNA, fragmented mRNA, tRNA, and DNA fragments. A total of 20 μL reaction mixture consisting of 10 μg total RNA, 2 uL CIP reaction buffer, and 2 μL CIP was incubated at 37 ^{o}C for 1 hour. The CIP reaction was terminated and cDNA was extracted by adding 15 μL ammonium acetate, 115 μL nuclease-free water, and 150 μL acid phenol:chloroform. The 5' cap of full- length mRNA was removed with tobacco acid pyrophosphatase (TAP) by mixing 5 μL CIP-treated RNA (from the previous step), 1 μL 10X TAP buffer, 2 μL TAP, and 2 μL nuclease-free water, mixed well, and incubated at 37 ^{o}C in a thermal cycler for 1 hour. A 45-base RNA adapter of known sequence was ligated to the resulting 5' monophosphate using T4 RNA ligase. The reaction mixture includes 2 μL CIP-/TAP-treated RNA, 1 μL 5' RACE Adapter, 1 μL 10X RNA ligase buffer, 2 μL T4 RNA Ligase, and 4 μL nuclease-free water and incubated at 37 ^{o}C for 1 hour.

Reverse transcription

The full-length mRNA from the previous step was reverse transcribed using random primers in a 20-μL reaction mix comprised of 2 μL ligated RNA, 4 μL dNTP mix, 2 μL random decamers, 2 μL 10X RT buffer, 1 μL RNAse inhibitor, 1 μL M-MLV Reverse Transcriptase, and 8 μL nuclease-free water.

PCR

A similar procedure was used to synthesize gene-specific primers to the 3' end of the RNA. The distance of the GSP3 and GSP4 were selected to be ca. 400-500 b from the

3’ end of the 5’ adapter. Two nested PCR’s were performed, each using a primer complementary to the RACE adapter (Table 1) on the 5’ end and a canine GSP on the 3’ end. For the first nested PCR, we used canine RLM-RACE outer primer on the 5’ end and *DLC1* GSP3 (outer primer) on the 3’end, and for the second nested PCR, a RLM-RACE inner primer on the 5’ end and GSP4 outer primer on the 3’end were used. PCR mixtures were prepared in 0.2-mL tubes as follows: 2 μL RT reaction product from previous step, 5 μL 10X PCR buffer, 4 μL dNTP mixture, 2 μL (10 μM),5’ RACE GSP, 2 μL (10 μM) 5‘ RACE Outer primer or inner primer 0.25 μL DNA Polymerase, and 34.75 μL nuclease-free water. The reaction mixture was incubated in an MJ thermal cycler using the following cycling conditions: a denaturation step at 96 ^{o}C for 2 minutes, followed by 40 cycles of 96 ^{o}C for 30 seconds (denaturation), 60 ^{o}C for 30 seconds (annealing), 72 ^{o}C for 30 seconds (extension), and a final extension step of 72 ^{o}C for 7 minutes, and kept at 4 ^{o}C.

Sequencing

The PCR product was enzymatically purified with Exo-SAP-IT® (Amersham, USA). PCR products were incubated for 30 minutes at 37°C and the reaction was terminated by incubation at 80°C for 15 minutes. PCR products were further purified with DNA Clean

& Concentrator™ kit (Zymo Research Corporation, Orange, CA) and eluted in 10 μL of HyClone® HyPure water (Thermo Scientific). Clean PCR products were sequenced by cycle sequencing using BigDye® Terminator Sequencing Kits (Life Technologies). The reaction was carried out in a 10-μL reaction containing 4 μL Big-Dye®, 4 μL PCR

product, and 2 μL (3.2 pmol) primer. The mixture was cycled as follows: initial denaturation at 96°C for 2 minutes, and 36 cycles of 96°C for 10 seconds, 50°C for 10 seconds, and 60°C for 4 minutes.

Western Blot Analysis

Because of no anti-canine-Sp1 antibody is commercially available, we also demonstrated affinity of an anti-human-Sp1 antibody by Western Blot. This was performed in a canine lymphoma cell line (OSW) a canine lymph node, and canine liver, with a human lymphoma cell line (MEC1) as a control.

Chromatin Immunoprecipitation (ChIP)

Binding of Sp1 to the promoter was evaluated using chromatin Immunoprecipitation (ChIP). This was performed using the Chip-IT kit from Active Motif to identify binding of Sp1 to *DLC1*, with complementary primers designed for qPCR to analyze the relative binding of Sp1 to *DLC1*, compared to *GAPDH*, a gene not dependent on Sp1 for expression. Results were expressed as a proportion of total amplicons.

Analysis

The amino acid sequences were obtained from ncbi.nml.nih.gov. These sequences were also translated (confirmed) using an open reading frame finder (http://www.ncbi.nml.nih.gov/gof/gorf.html.) of nucleotide sequences (Bryan et al., 2009). The nucleotide and amino acid sequences of the canine *DLC1* (accession number

FJ602870) were analyzed and compared to the human *DLC1* sequences (accession numbers AF026219 and NM_006094), and the mouse sequences (accession number NM_015802) using CLC genomics (CLC bio/Qiagen, Cambridge, MA, USA) and Clustal Omega (www.clustal.org) software. The potential Sp1 binding sites in *DLC1* promoter regions were identified using MatInspector program (Cartharius et al., 2005).

Research Project

The research project was performed at the Comparative Oncology Laboratory, Veterinary Clinical Sciences, Washington State University. Dr. Jeffrey N. Bryan funded and supervised the research work.

Results

DLC1 Sequencing and Analysis

In this study, deleted in Liver cancer (*DLC1*) was successfully amplified starting with primers designed from predicted canine genome sequences that are available online and cDNA of human sequences (AF026219, and NM_006094). The cDNA sequences were determined using an automated DNA sequencer, and were submitted to GenBank database (FJ602870; Bryan et al. 2009).

The nucleotide sequences of the canine *DLC1* gene and its deduced amino acids are shown in Figures 1 and 2, respectively. The translated region of the full-length cDNA of the canine *DLC1* gene contained a transcript of 3724 bases, organized into 14 exons and 1091 amino acid (aa) polypeptide with a molecular weight of about 123 KDa. In contrast, the human *DLC1* Gene (AF026219) has a transcript size of 3766 bases, organized into 14 exons, and 1083 amino acid polypeptide with a molecular weight of about 123 KDa (Table 2). The longer human transcript 7365 bases (AB051510) encodes 1554 amino acids polypeptide with a size of 170 KD (Nagaze et al. 2000); whereas mouse *DLC1* (NM_015802) includes a transcript size of 6184 and 1092 amino acid polypeptide with a molecular weight of 123 KDa (Table 2). The canine *DLC1* gene is located on the long arm of chromosome 16, the human *DLC1* gene is located on chromosome 8, and the mouse *DLC1* is on chromosome 8 (Yuan et al., 1998, Bryan et al., 2009).

The nucleotide and amino acid sequences of the canine *DLC1* (accession number FJ602870) were highly similar to the human *DLC1* sequences (accession numbers AF026219) (Table 3). The canine cDNA sequences shares 85 to 88 % identity with the human sequences AF026219 and AB051510, respectively, and 81 % identity with the mouse sequences NM_015802. Comparison of the amino acid sequences of the canine *DLC1* with amino acid sequences of both human accession numbers AAB81637.1 and BAB21814, and the mouse amino acid sequences accession numbers NP_056617 are shown in table 4. The canine DLC1 peptide sequences share 94 % identity with human sequences (accession numbers AAB81637 and BAB21814), and 93 % with mouse sequences accession number NP_056617.

Analysis of the conserved domains of the canine amino acid sequence of DLC1 is shown in Figure 3. The DLC1 canine amino acid sequences contain four major domains that were previously identified in human and mouse (Yuan et al., 1998, Durkin et al., 2002): 1) An N-terminal sterile alpha motif (SAM) domain; 2) a RhoGAP domain, a conserved region responsible for the catalytic activity of RhoGAPs; and 3) a C-terminal steroidogenic acute regulatory protein (STAR)-related lipid-transfer (START) domain; and 4) a serine-rich (SR) domain located between the SAM domain and RhoGap domain. The SAM domain is 60 amino acids long and located at amino acid positions 17 to 76 (nucleotide position 252-453), a RhoGAP domain is 219 amino acids long and located at amino acid positions 635 to 853 (nucleotide position 2142-2760), and finally START domain is 204 amino acids long and located at amino acid positions 878 to 1081 (nucleotide position 2877-3474).

DLC1 Serine rich region

The FoldIndex analysis indicated 16 disordered regions spanning 538 residues, mostly located in the region between amino acids 80 and 630, 16 disordered regions spanning 565 amino acid residues predominantly located between amino acid 78 and 638, and 19 disordered regions spanning 602 amino acid residues predominantly located between amino acids 78 and 639 of canine, human, and mouse protein sequences, respectively. The secondary structure analysis using FoldIndex predicted that regions where SAM, RhoGAP and START domains will form globular structures and a large region between the SAM and RhoGAP domain is unlikely to fold or form a globular conformation (Figure 3). This region between the SAM and RHOGAP domain is encoded primarily by exon 5, and similar to human *DLC1*. It is an area rich in serine residues (SR) and is known to be the least conserved *DLC1* domain in both human (Durkin et. Al. 2002) and dogs (Bryan et al. 2009). This domain has many characteristics in common with a class of proteins that have been determined intrinsically unstructured or (disordered) proteins (Tompa 2002). Unfolded or unstructured proteins are known to be enriched in amino acids (Ser, Pro, Gln, Glu, Lys) that promote disorder and depleted of order-promoting residues Cys, Trp, Tyr, Ile, Phe, Leu, His, Thr, and Asn (Macarena et al, 2012).

Specificity protein 1

Western Blot Analysis of Sp1

Western blot analysis was used to confirm SP1 binding in the CpG-rich promoter area of a canine lymphoma cell line (OSW), a canine lymph node, and canine liver, with a

human lymphoma cell line (MEC1) serving as a control. Western blot analysis confirmed that the antibody used for the chromatin immunoprecipitation (ChIP) reacts with the canine Sp1 protein as expected (Figure 10).

Chromatin Immunoprecipitation (ChIP)

Quantitative PCR revealed approximately equal copy numbers of *GAPDH* (encodes Glyceraldehyde-3Phosphate Dehydrogenase) and *DLC1* in the immunoprecipitate. This indicates that *DLC1* was not enriched in these samples (Figure 11).

Promoter Analysis of SP1 Binding Sites

DLC1 promoter analysis using MatInspector program included multiple Sp1 binding sites suggesting that Sp1 might be involved in *DLC1* expression (Figure 12). The consensus sequences for Sp1 binding sites in the canine DLC1 are GGGCGG or GGGCGGG.

Discussion and Conclusions

Deleted in liver cancer one (*DLC1*) is a tumor suppressor gene that is frequently inactivated by aberrant CpG methylation and mutations in multiple tumors (Yuan et al. 1998, Shikhar et al. 2009, Esteller 2002, Bryan et al. 2009). Currently, extensive research is being conducted on epigenetics of cancer induction, which can contribute to accurate diagnosis and treatment of cancer. We isolated for the first time the full-length cDNA of canine *DLC1*, submitted to GenBank (Bryan et al. 2009), and compared it to *DLC1* of human and mouse. The canine *DLC1* cDNA and amino acid sequence has high homology to the human counterpart *DLC1* genes (accession numbers AAB81637 and BAB21814) with 88% and 95% identity values, respectively. Analysis of the conserved domains of the canine amino acid sequence of *DLC1* revealed four domains that are similar to the human counterpart. RhoGAP is the main protein encoded by *DLC1*. RhoGAP is a GTPase-activating protein that catalyzes the conversion of the active Rho-GTP, a member of the Rho small GTP-binding subfamily protein, into the inactive Rho-GDP (Yuan et al. 2004). The SAM domain is a protein-to-protein, interacting domain that has been shown to interact with other SAM domains to form homo-dimers; it can also interact with DNA and RNA (Bowie and Kim 2003, Qiao and Bowie 2005). In many proteins, the C-terminal, lipid-binding START domain is identified to have a role in lipid transport or metabolism (Alpy and Tomasetto 2005).

Sp1 (specificity protein 1) is a member of the Krüppel-like factor (Sp/KLF) family of transcription factors, which encodes a zinc finger transcription factor, and binds to GC- rich promoters (Black et al. 2001). These KLF proteins are involved in the regulation of many cellular processes including cell proliferation, cell differentiation,

chromatin remodeling, and apoptosis. SP1 also plays a critical role in the growth and metastasis of many tumor types (Jones and Baylin 2002). SP1 binding sites are usually found in GC- rich promoters of house-keeping genes as well as other regions upstream of the Transcription Start Site (TSS). The presence of SP1 binding sites in the promoter region is one of the mechanisms that is generally considered to protect CpG islands from *denovo* methylation so that gene expression is maintained (Bird 2002). Kim and co-authors (2008) examined the molecular mechanism for induction of *DLC1* using trichostatin A (TSA), a histone deacetylase (HDAC) inhibitor in human gastric cancer cell lines and determined that Sp1-binding sites located in the *DLC1* promoter play important roles in the activation of *DLC1* by TSA (Kim et al. 2008). To determine that the transcription factor Sp1 binds to the promoter region of canine *DLC1*, we conducted western blotting using anti-human-Sp1 antibody. Western blot analysis confirmed that the antibody used for the ChIP reacts with the canine protein Sp1, as expected (Figure 11). In addition *DLC1* promoter analysis using MatInspector program included multiple Sp1-binding sites, suggesting that Sp1 might be involved in *DLC1* expression. Canine Sp1 binding sites are conserved and show similarity of transcription control to that of the human counterpart (Bryan et al. 2009).

Although the promoter sequence was as expected, preliminary ChIP analysis does not support Sp1 binding. With clear evidence that the antibody used in these experiments binds to canine Sp1, it must be considered that other transcription factors may control expression of the canine DLC1 gene. If transcription factors are methylation sensitive, the apparent boundary effect of the *DLC1* binding sites may allow usual binding of the other factors.

The mechanisms of epigenetic gene regulation are unexplored in dogs, and understanding may yield new therapy targets for cancer. Deleted in liver cancer one (*DLC1*) is a tumor suppressor gene found in a variety of cancers, particularly liver, breast, prostate, and colorectal cancers (Yuan et al., 1998, Liao and Lo 2007, Goodison et al., 2005). A better understanding of *DLC1* functions and signaling pathways can help the prevention and treatment of *DLC1*-associated cancers. Bryan (2008) suggested that *DLC1* and its downstream signaling molecules can be useful biomarkers for prognosis, diagnosis, and therapeutic target identification for interventions or treatments of cancers.

Presently, we recognize that human cancer is a disease of molecules and genes that originate in the progeny of a single cell through that accumulation of multiple genetic abnormalities (Teicher and Bagley, 2009). Our research focused on isolating and characterizing canine *DLC1*, which will provide us with a unique opportunity to study *DLC1*-related cancers in dogs.

In addition, the dog represents a good model organism to study human-related cancers, since dog has disease onset mechanisms similar to those of human. Finally, the dog can also serve as a good model organism because it possesses great diversity (morphology and traits) and large number of diseases that are more similar to human than other organisms. For example, non-Hodgkin's lymphoma (NHL) is a common disease of dog and canine NHL shares many features with human NHL. A preliminary study was performed to identify anomalies in the gene promoter and CpG island DNA methylation and the associated modification of gene expression in canine NHL, which is helping to develop a naturally occurring model of human NHL for research in carcinogenesis, biomarker identification, and therapy (unpublished results in our laboratory).

References

Alpy, F, Tomasetto C., 2005. *Give lipids a START: the StAR-related lipid transfer (START) domain in mammals.* J Cell Sci.; 118: 2791–801.

Bartolomei M.S., Ferguson-Smith A.C., 2011. *Mammalian genomic imprinting.* Cold Spring Harb. Perspect. Biol., 3 .

Bird A., 2002. *DNA methylation patterns and epigenetic memory.* Genes Dev.; 16, 6–21.

Black, Adrian R., Black, Jennifer D., Azizkhan-Clifford, Jane, 2001. *Sp1 and krüppel-like factor family of transcription factors in cell growth regulation and cancer.* J. Cell. Physiol., V 188- 2 pp 1097-4652

Bowie JU, Kim CA., 2003. *"SAM domains: uniform structure, diversity of function".* Trends Biochem. Sci. 28:12: 625–628.

Brothman A.R. 1997. *Cytogenetic studies in prostate cancer: are we making progress* Cancer Genet. Cytogenet. 95:116-121.

Bryan J.N., Jabbes M., Berent L.M, Arthur J.L., Taylor K.H., Rissetto K.C., Henry C.J., Rahmatpanah F., Rankin W.V, Villamil J.A., Lewis M.R, and Caldwell C.W., 2009. *Hypermethylation of the DLC1 CpG island does not alter gene expression in canine lymphoma.* BMC Genetics 2009, 10:73

Bryan JN, Taylor KH, Henry CJ, Selting KA, Rahmatpanah F, Lewis MR *et al*., 2008. *DNA methylation in cancer: techniques and preliminary evidence of hypermethylation in canine lymphoma*. *Cancer Therapy*, 6: 137-148.

Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T (2005). *MatInspector and beyond: promoter analysis based on transcription factor binding sites.* Bioinformatics 21, 2933-42

Durkin, ME, Yuan BZ, Thorgeirsson SS, Popescu NC. *Gene structure, tissue expression, and linkage mapping of the mouse DLC-1 gene (Arhgap7).* Gene. 2002; 288: 119–27.

Elledge, S.J., and Amon, A. (2002). *The BRCA1 suppressor hypothesis: an explanation for the tissue-specific tumor development in BRCA1 patients.* Cancer Cell 1, 129–132.

Esteller M, Gaidano G, Goodman SN, Zagonel V, Capello D, Botto B, Rossi D, Gloghini A, Vitolo U, Carbone A, Baylin SB and Herman JG., 2002. *Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients*

with diffuse large B-cell lymphoma. J. Natl. Cancer Inst.; 94, 26 ± 32.

FoldIndex program (http://bioportal.weizmann.ac.il/fldbin/findex).

Gardiner-Garden, M., and Frommer, M., 1987. *CpG islands in vertebrate genomes.* J. Mol. Biol. 196, 261–282.

Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, Scott G, et al. et al, 2004. *Genome sequence of the Brown Norway rat yields insights into mammalian evolution.* Nature: 420:520-562.

Goll MG, Bestor TH 2005. *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem 74: 481–514.

Goodison S, Yuan J, Sloan D, Kim R, Li C, Popescu NC, Urquidi V., 2005. *The RhoGAP protein DLC-1 functions as a metastasis suppressor in breast cancer cells.* Cancer Res.; 65: 6042–53.

Greger V, Passarge E, Hopping W, Messmer E and Horsthemke B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum. Genet., 83, 155 ± 158.

Hackenberg M, Previti C, Luque-Escamilla PL, Carpena P, Martinez-Aroza J, Oliver JL, 2006. *CpG cluster: a distance-based algorithm for CpG-island detection.* BMC Bioinformatics: 7:446.
Han L, Zhao Z: *Contrast features of CpG islands in the promoter and other regions in the dog genome.* Genomics 2009, 94:117-124.

Han, L., Lin, I. G. and Hsieh, C. L. 2001. Protein binding protects sites on stable episomes and in the chromosome from de novo methylation. Mol. Cell. Biol. 21: 3416-3424.

Isola J.J., Kallimoniemi O.P., Chu L.W., Fuqua S.A., Hilsenbeck S.G., Osborne C.K., Waldman F.M., 1995. *Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer.* Am J Pathol.; 147(4): 905–911.

Yoder, J. A., Walsh, C. P. and Bestor, T. H., 1997. *Cytosine methylation and the ecology of intragenomic parasites.* Trends Genet. 13, 335-340.

Jones PA, Baylin SB. *The epigenomics of cancer.* Cell. 2007;128:683–692
Jones PA, Baylin SB. *The fundamental role of epigenetic events in cancer.* Nat Rev Genet 2002; **3**: 415–28.

Kim T.Y., Kim I.S., Jong H.S., Lee J.W., Kim T.Y., Jung M., and Bang Y.J., 2008. *Transcriptional induction of DLC-1 gene through Sp1 sites by histone deacetylase inhibitors in gastric cancer cells.* Exp Mol Med.; 40(6): 639–646.

Kim T.Y., Kim I.S., Jong H.S., Lee J.W., Kim T.Y., Jung M., and Bang Y.J., 2008. *Transcriptional induction of DLC-1 gene through Sp1 sites by histone deacetylase inhibitors in gastric cancer cells.* Exp Mol Med.; 40(6): 639–646.

Liao YC, Shih YP, Lo SH, 2008. *Mutations in the focal adhesion targeting region of deleted in liver cancer-1 attenuate their expression and function.* Cancer Res. 2008 Oct 1;68(19):7718-22.

Liao, Y. -C., Lo, S. H., Deleted in liver cancer-1 (DLC-1): A tumor suppressor not just for liver, International Journal of Biochemistry and Cell Biology. 2007 doi:10.1016/j.biocel.2007.04.008

Macarena Marín, Veronika Thallmair and Thomas Ot., 2012. *Mediates Protein-Protein Interactions Region of AtREM1.3 Remorin Protein The Intrinsically Disordered N-terminal P.* J. Biol. Chem. 287:39982-39991.

Mancini D.N., Singh S. M., Archer T. K., and Rodenhiser D.I., 1999. *Site-specific DNA methylation in the neurofibromatosis (NF1) promoter interferes with binding of CREB and SP1 transcription factors.* Oncogene, 18:28, 4108-4119

Marchio A., Meddeb M., Pineau P., Danglot G., Tiollais P., Bernheim A., Dejean A., 1997. *Recurrent chromosomal abnormalities in hepatocellular carcinoma detected by comparative genomic hybridization.* Genes Chromosomes Cancer;18(1):59-65.

Maruyama K, Sugano S, 1994. *Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides.* Gene.;138(1-2):171-4.

Nagase T, Kikuno R, Hattori A, Kondo Y, Okumura K, Ohara O, 2000. *Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.* DNA Res.; 7:347–55.

Novik KL, Nimmrich I, Genc B, Maier S, Piepenbrock C, Olek A, Beck S., 2002. *Epigenomics: genome-wide study of methylation phenomena.* Curr Issues Mol Biol.; 4(4):111-28.

Osborne, C., Wilson P., and Tripathy D., 2004. *Oncogenes and Tumor Suppressor Genes in Breast Cancer: Potential Diagnostic and Therapeutic Applications.* The Oncologist July 2004vol. 9 no. 4 361-377.

Qiao F., Bowie J., 2005. *The many faces of SAM.* Sci. STKE 2005(286),P.re7. Rivlin N, Brosh R, Oren M, Rotter V. *Mutations in the p53 tumor suppressor gene: important milestones at the various steps of tumorigenesis. Genes Cancer* 2011; **2**: 466–474.

Schaefer, B. C. (1995). *Revolutions in Rapid Amplification of cDNA Ends: New Strategies for Polymerase Chain Reaction Cloning of Full-Length cDNA Ends.* Anal. Biochem. 227: 255 - 273.

Shikhar Sharma, Theresa K. Kelly, and Peter A. Jones., 2009. *Epigenetics in cancer.* Carcinogenesis; 31 (1):27-38

Sherr CJ., 2004. *Principles of tumor suppression.* Cell 116:2: 235–46
Teicher B.A., BAGLEY R.J., 2009. *Cancer Drug Discovery and Development Series.*

Humana Press, a part of Springer Science+Business Media, LLC 2009.
Thangapazham R, Saenz F, Katta S, Mohamed AA, Tan S-H, Petrovics G, Srivastava S and Dobi A., 2014. *Loss of the NKX3.1 tumorsuppressor promotes the TMPRSS2-ERG fusion gene expression in prostate cancer.* BMC Cancer. 2014: Jan 13;14:16.

Tompa P., 2002. *Intrinsically unstructured proteins.* Trends Biochem Sci.;27(10):527-33.

Uversky, V.N., et al. 2000. *Why are 'natively unfolded' proteins unstructured under physiologic conditions?* Proteins 41415–427.

Venter JC, Adams MD, et al., 2001: *The sequence of the human genome.* Science; 16;291(5507):1304-51.

Wong C. M., Lee J. M. F., Ching Y. P., Jin D.Y., and Ng I. O. L., 2003. *"Genetic and epigenetic alterations of DLC-1 gene in hepatocellular carcinoma,"* Cancer Research, vol. 63, no. 22, pp. 7646–7651.

Xue W, Krasnitz A, Lucito R, Sordella R, Vanaelst L, Cordon-Cardo C, Singer S, Kuehnel F, Wigler M, Powers S, Zender L, Lowe SW: *DLC1 is a chromosome 8p tumor suppressor whose loss promotes hepatocellular carcinoma.* Genes Dev 2008,22:1439-1444.

Yoder, J. A., Walsh, C. P. and Bestor, T. H., 1997. *Cytosine methylation and the ecology of intragenomic parasites.* Trends Genet. 13, 335-340.

Yuan BZ, Miller MJ, Keck CL, Zimonjic DB, Thorgeirsson SS, Popescu NC., 1998. *Cloning, characterization, and chromosomal localization of a gene frequently deleted in human liver cancer (DLC-1) homologous to rat RhoGAP.* Cancer Res. 58:2196–9.

Yuan, B, Jefferson A.M., Baldwin K. T., Thorgeirsson S.S., Popescu N. C., Reynolds S. H., 2004. *DLC-1 operates as a tumor suppressor gene in human non-small cell lung Carcinomas.* Oncogene; 23: 1405–1411

Zhaoa Z., and Hana L., 2009. *CpG islands: Algorithms and applications in methylation studies.* Biochemical and Biophysical Research Communications; 382(4) 643–645.

Figure 1: *Canis lupus familiaris* deleted in liver cancer 1 mRNA, complete cds. sequenced from canine lymph node DNA samples.

```
   1 ctctgccctc gcctgagcgc gctctcctgc cgagctctgg aaacccctgc gtccttcccg
  61 gggccgcggc acggagcggc gagccgggat gcgctgagga gccgccagcg cgcgcgtgtc
 121 gggcggctcc gtgggtccag gcgcgggagc tgagccggcc gggcgctgcc tcgacttccc
 181 tcccgcgccg ccgccgcccc cgccccccgc cgcagttgat gtgcaggagg aagccggaca
 241 tcatgatcct aacacaaatt gaagccaagg aagcttgtga ttggctacgg gcaacaggtt
 301 tcccccagta tgcacagctt tacgaagatc ttctgttccc cattgatatt tcctcggtca
 361 agagagagca tgacttttg gacagagatg ccattgaggc tttatgcagg cgtctaaata
 421 ctctaaacaa atgtgcggtg atgaagctgg aaattagtcc tcatcggaag agaagtgagg
 481 attcggatga ggacgagcct tgtgcaataa gcggcaagtg gaccttccag agagacagca
 541 agaggtggtc acggctggag gagtttgatg tcttttcttc aaagctggac ccagcccccg
 601 gggcgcccgc cgaggccccc ctgaagaccg cggcgagcca cgaaagcatg ctgacagagc
 661 tcagcgagcg ccaggaggtg gcttcggtcc tcagcctcag cagcaccggc agtctccccg
 721 tgcacgcgcc ccacgcgggg gacgcggcaa cgccccggac caactcggtc atcagcgtct
 781 gctcctccgg ccacttcgtg ggcaacgacg actccttctg cagcctgccc tctcccaagg
 841 aactgtccag cttcagcttc agcatgaaag gccaagagaa gaacgccaag tccaagacgc
 901 gcagcctgct gaagcgcatg gagagcctga agctcagggg ctcgccccac agcaagcaca
 961 aggcgccttc caagctgggg ctgatcatca gcgggcccat cctgcaggag gggatggacg
1021 aggagaagct gaagcagctc aactgtgtgg agatctccgc gctcaacggc aaccgcatca
1081 acgtgccggc cgtgcgcaag cggagcgtct ccaactccac gcagaccagc agcagcagca
1141 gccagtcgga gaccagcagc gccgtgagca cgcccagccc ggtcaccagg actcggagcc
1201 tcagtgcctg caacaagcgg gtgggcatgt acttagaggg ctttgacccg ttcaatcagt
1261 ccacgttcaa caacgttatg gagcagaact gtaagaaccg cgagagctac ccagaggaca
1321 cagtgttcta catcccagaa gatcacaagc ctggcacttt ccccaaggcc ctctccaatg
1381 gcagtttccc tccctcaggg aataacagct ctgtcaactg gaggactgga agcttccacg
1441 gccccggcca catcagcctg aggagggaaa acagcagtga cagccccaag gagctcaaga
```

1501 ggcgcaattc ctccagttcc atgagcagcc gcctgagtat ctatgacaac gtgccaggct

1561 ccatcctcta ttccagttca ggggacctgg ccgacctgga gaacgaggac atcttccccg

1621 agctggatga catcctctac catgtgaagg ggatgcaaag gatagtcaac cagtggtccg

1681 agaagttttc agacgagggg gactctgact cagccctgga ctcggtctct ccatgcccct

1741 cctctcccaa acagatacac ctggatgtgg acaacgaccg agccacgccc agtgacctag

1801 acagtacggg caactcactg aatgaaccag aagagccctc agacatcccg gagaggaggg

1861 attctggggt tggggcctcc ctgacaaggt ccaacaggca caggctgaga tggcacagtt

1921 tccagagctc ccatcggccg agcttaaact ctgtatccct gcagatcaac tgccagtctg

1981 tggcccagat gaacctgctg cagaaatact cgctcctaaa gttaactgcc ctgctggaga

2041 aatacacacc ctctaataag catggtttta gctgggctgt gcccaagttc atgaaaagga

2101 tcaaggttcc agactacaag gaccggaatg tattcggggt ccctctgaca gtcaatgtgc

2161 agcgcacagg acagcccctg ccccagagca ttcagcaggc catgcgctac ctccgcaacc

2221 actgtttgga tcaggttggg ctcttcagaa aatcaggcgt caaatcccga attcaggctc

2281 tgcgccagat gaatgaaagt gccatagatt gtgtcaacta tgagggacaa tctgcttatg

2341 atgtagcaga catgttgaag cagtattttc gagaccttcc tgagccgcta atgacgaaca

2401 aactctcaga aacctttctg cagatctacc agtgtgtgcc caaggaccag cgcctgcagg

2461 ccatgaaggc cgccattatg ctcctgcctg atgagaaccg agaggtccta cagacacttc

2521 tttatttctt gagcgatgtc acagcagccg taaaagaaaa ccagatgact cccaccaacc

2581 tggccgtgtg cttagcacct tccctttcc atctcaacac cctgaagaga gagaattctt

2641 ctccaagggt aatgcaaaga aaacaaagtt tgggcaaacc agatcagaaa gatttgaatg

2701 aaaaccttgc tgccactcag ggtctggccc atatgattgc tgaatgcaag aagctttcc

2761 aggttcctga ggaaatgagt cgatgtcgga attcctacac ggaacaggag ctgaagcccc

2821 tcactctaga agcattagga cgcctgcgta atgatgagtc tgctgactac cagcacttcc

2881 tccaggactg tgtggatagc ctgtttaaag aagtcaagga gaagtttaaa ggctgggtca

2941 gctactctac atcggagcaa gctgagctgt cctacaagaa ggtgagcgaa ggaccacctc

3001 tgaggctttg gcggtcgacc attgaagtcc ctgcaatgcc tgaggaaatc ctaaagcgcc

3061 tacttaaaga gcagcacctc tgggatgtag acctgttaga ttcaaaagtg attgaaattc

3121 tagacagcca aactgaaatc taccagtatg tccaaaacag tatggcgccc catcccgctc

3181 gggactacgt tgttttgaga acatggagga ctaatttacc aaaaggggca tgtgcccttt

3241 tactcacctc tgtggatcac gaccgggcac ctgtggtggg ggtgagagtc aatgtgctcc

3301 tctccaggta tctgatcgaa ccctgtggat cagggaaatc taaactcacc tacatgtgca

3361 gagctgactt aaggggccac atgccagagt ggtacacaaa atctttcgga catctgtgtg

3421 cagctgaagt tgtgaagatc cgagactctt tcagcaatca gaacaccgaa accaaagaca

3481 ccaaatctag gtgattcctg aagctaagcg gccacatctg ccatctggat gtttgtttct

3541 ggaagcccat gccggtcctt gaaagagtgg attctgagtc tagtcctgaa acaaagaaaa

3601 ctaatttaga gtgtggggat caactagcaa ttgaagacgt ttttaaagct gcttcctgtt

3661 tgttgaaggt ctagatttta ggccttgact ggaatatata agactgtgca cgaaaaaaaa

3721 aaaa

Figure 2: *Canis lupus familiaris* deleted in liver cancer 1 amino acid sequence.

MCRRKPDIMILTQIEAKEACDWLRATGFPQYAQLYEDLLFPIDI

SSVKREHDFLDRDAIEALCRRLNTLNKCAVMKLEISPHRKRSEDSDEDEPCAISGKWT

FQRDSKRWSRLEEFDVFSSKLDPAPGAPAEAPLKTAASHESMLTELSERQEVASVLSL

SSTGSLPVHAPHAGDAATPRTNSVISVCSSGHFVGNDDSFCSLPSPKELSSFSFSMKG

QEKNAKSKTRSLLKRMESLKLRGSPHSKHKAPSKLGLIISGPILQEGMDEEKLKQLNC

VEISALNGNRINVPAVRKRSVSNSTQTSSSSSQSETSSAVSTPSPVTRTRSLSACNKR

VGMYLEGFDPFNQSTFNNVMEQNCKNRESYPEDTVFYIPEDHKPGTFPKALSNGSFPP

SGNNSSVNWRTGSFHGPGHISLRRENSSDSPKELKRRNSSSSMSSRLSIYDNVPGSIL

YSSSGDLADLENEDIFPELDDILYHVKGMQRIVNQWSEKFSDEGDSDSALDSVSPCPS

SPKQIHLDVDNDRATPSDLDSTGNSLNEPEEPSDIPERRDSGVGASLTRSNRHRLRWH

SFQSSHRPSLNSVSLQINCQSVAQMNLLQKYSLLKLTALLEKYTPSNKHGFSWAVPKF

MKRIKVPDYKDRNVFGVPLTVNVQRTGQPLPQSIQQAMRYLRNHCLDQVGLFRKSGVK

SRIQALRQMNESAIDCVNYEGQSAYDVADMLKQYFRDLPEPLMTNKLSETFLQIYQCV

PKDQRLQAMKAAIMLLPDENREVLQTLLYFLSDVTAAVKENQMTPTNLAVCLAPSLFH

LNTLKRENSSPRVMQRKQSLGKPDQKDLNENLAATQGLAHMIAECKKLFQVPEEMSRC

RNSYTEQELKPLTLEALGRLRNDESADYQHFLQDCVDSLFKEVKEKFKGWVSYSTSEQ

AELSYKKVSEGPPLRLWRSTIEVPAMPEEILKRLLKEQHLWDVDLLDSKVIEILDSQT

EIYQYVQNSMAPHPARDYVVLRTWRTNLPKGACALLLTSVDHDRAPVVGVRVNVLLSR

YLIEPCGSGKSKLTYMCRADLRGHMPEWYTKSFGHLCAAEVVKIRDSFSNQNTETKDT

KSR

Figure 3. A Schematic diagram showing the composition of protein domains for the canine *DLC1* (accession number FJ602870): SAM domain, GAP domain, START domain, and Serine rich domain.

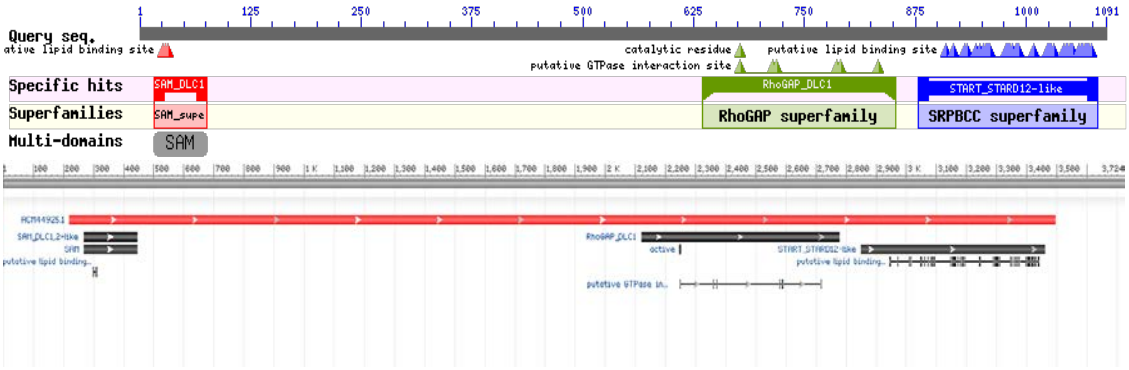


Figure 4. Predicted secondary structure of the canine DLC-1 protein, performed using the FoldIndex program (http://bioportal.weizmann.ac.il/fldbin/findex). The canine DLC1 protein contains serine-rich region between the SAM and RhoGAP domains, which is predicted to have a largely unfolded conformation.

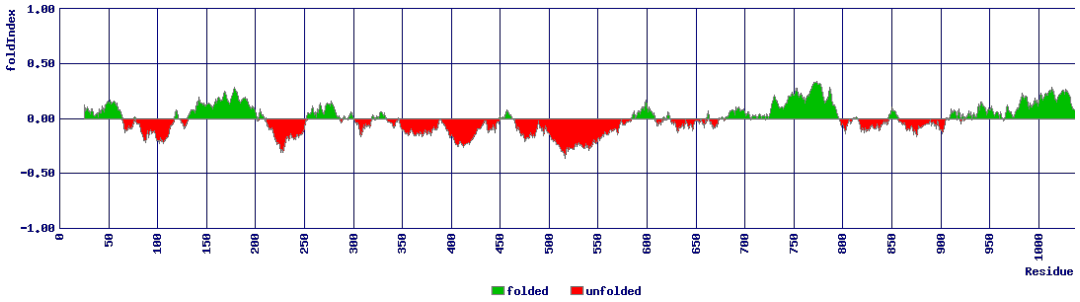


Figure 5. Canine *DLC1* protein sequence showing features of the serine-rich region and the amino acid sequence of the region between the SAM and RhoGAP domains (approximate residues 79–638) of canine *DLC*.

```
   1 MCRRKPDIMI LTQIEAKEAC DWLRATGFPQ YAQLYEDLLF PIDISSVKRE
  51 HDFLDRDAIE ALCRRLNTLN KCAVMKLEIS PHRKRSEDSD EDEPCAISGK
 101 WTFQRDSKRW SRLEEFDVFS SKLDPAPGAP AEAPLKTAAS HESMLTELSE
 151 RQEVASVLSL SSTGSLPVHA PHAGDAATPR TNSVISVCSS GHFVGNDDSF
 201 CSLPSPKELS SFSFSMKGQE KNAKSKTRSL LKRMESLKLR GSPHSKHKAP
 251 SKLGLIISGP ILQEGMDEEK LKQLNCVEIS ALNGNRINVP AVRKRSVSNS
 301 TQTSSSSSQS ETSSAVSTPS PVTRTRSLSA CNKRVGMYLE GFDPFNQSTF
 351 NNVMEQNCKN RESYPEDTVF YIPEDHKPGT FPKALSNGSF PPSGNNSSVN
 401 WRTGSFHGPG HISLRRENSS DSPKELKRRN SSSSMSSRLS IYDNVPGSIL
 451 YSSSGDLADL ENEDIFPELD DILYHVKGMQ RIVNQWSEKF SDEGDSDSAL
 501 DSVSPCPSSP KQIHLDVDND RATPSDLDST GNSLNEPEEP SDIPERRDSG
 551 VGASLTRSNR HRLRWHSFQS SHRPSLNSVS LQINCQSVAQ MNLLQKYSLL
 601 KLTALLEKYT PSNKHGFSWA VPKFMKRIKV PDYKDRNVFG VPLTVNVQRT
 651 GQPLPQSIQQ AMRYLRNHCL DQVGLFRKSG VKSRIQALRQ MNESAIDCVN
 701 YEGQSAYDVA DMLKQYFRDL PEPLMTNKLS ETFLQIYQCV PKDQRLQAMK
 751 AAIMLLPDEN REVLQTLLYF LSDVTAAVKE NQMTPTNLAV CLAPSLFHLN
 801 TLKRENSSPR VMQRKQSLGK PDQKDLNENL AATQGLAHMI AECKKLFQVP
 851 EEMSRCRNSY TEQELKPLTL EALGRLRNDE SADYQHFLQD CVDSLFKEVK
 901 EKFKGWVSYS TSEQAELSYK KVSEGPPLRL WRSTIEVPAM PEEILKRLLK
 951 EQHLWDVDLL DSKVIEILDS QTEIYQYVQN SMAPHPARDY VVLRTWRTNL
1001 PKGACALLLT SVDHDRAPVV GVRVNVLLSR YLIEPCGSGK SKLTYMCRAD
1051 LRGHMPEWYT KSFGHLCAAE VVKIRDSFSN QNTETKDTKS R
```

Figure 6. Predicted secondary structure of the human *DLC1* protein, performed using the FoldIndex program (http://bioportal.weizmann.ac.il/fldbin/findex). The Human *DLC1* protein contains serine-rich region between the SAM and RhoGAP domains, which is predicted to have a largely unfolded conformation.

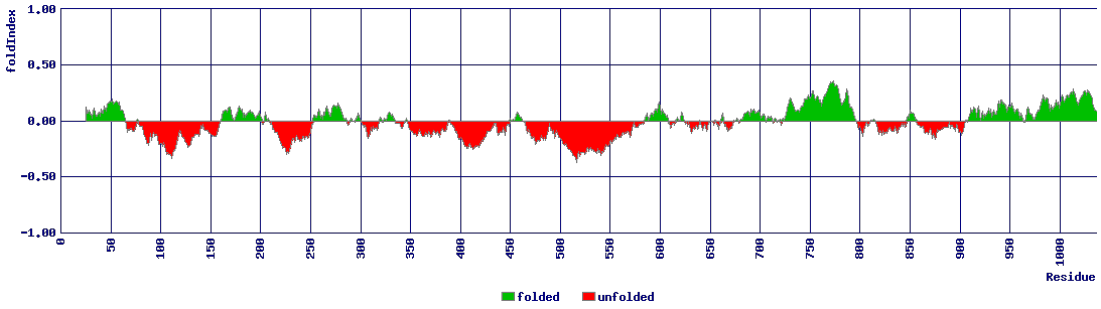


Figure 7. Human *DLC1* protein sequence showing features of the serine-rich region and the amino acid sequence of the region between the SAM and RhoGAP domains (approximate residues 78–638) of human *DLC1*.

```
   1 MCRKKPDTMI LTQIEAKEAC DWLRATGFPQ YAQLYEDFLF PIDISLVKRE
  51 HDFLDRDAIE ALCRRLNTLN KCAVMKLEIS PHRKRSDDSD EDEPCAISGK
 101 WTFQRDSKRW SRLEEFDVFS PKQDLVPGSP DDSHPKDGPS PGGTLMDLSE
 151 RQEVSSVRSL SSTGSLPSHA PPSEDAATPR TNSVISVCSS SNLAGNDDSF
 201 GSLPSPKELS SFSFSMKGHE KTAKSKTRSL LKRMESLKLK SSHHSKHKAP
 251 SKLGLIISGP ILQEGMDEEK LKQLNCVEIS ALNGNRINVP MVRKRSVSNS
 301 TQTSSSSSQS ETSSAVSTPS PVTRTRSLSA CNKRVGMYLE GFDPFNQSTF
 351 NNVVEQNFKN RESYPEDTVF YIPEDHKPGT FPKALTNGSF SPSGNNGSVN
 401 WRTGSFHGPG HISLRRENSS DSPKELKRRN SSSSMSSRLS IYDNVPGSIL
 451 YSSSGDLADL ENEDIFPELD DILYHVKGMQ RIVNQWSEKF SDEGDSDSAL
 501 DSVSPCPSSP KQIHLDVDND RTTPSDLDST GNSLNEPEEP SEIPERRDSG
 551 VGASLTRSNR HRLRWHSFQS SHRPSLNSVS LQINCQSVAQ MNLLQKYSLL
 601 KLTALLEKYT PSNKHGFSWA VPKFMKRIKV PDYKDRSVFG VPLTVNVQRT
 651 GQPLPQSIQQ AMRYLRNHCL DQVGLFRKSG VKSRIQALRQ MNEGAIDCVN
 701 YEGQSAYDVA DMLKQYFRDL PEPLMTNKLS ETFLQIYQYV PKDQRLQAIK
 751 AAIMLLPDEN REVLQTLLYF LSDVTAAVKE NQMTPTNLAV CLAPSLFHLN
 801 TLKRENSSPR VMQRKQSLGK PDQKDLNENL AATQGLAHMI AECKKLFQVP
 851 EEMSRCRNSY TEQELKPLTL EALGHLGNDD SADYQHFLQD CVDGLFKEVK
 901 EKFKGWVSYS TSEQAELSYK KVSEGPPLRL WRSVIEVPAV PEEILKRLLK
 951 EQHLWDVDLL DSKVIEILDS QTEIYQYVQN SMAPHPARDY VVLRTWRTNL
1001 PKGACALLLT SVDHDRAPVV GVRVNVLLSR YLIEPCGPGK SKLTYMCRVD
1051 LRGHMPEWYT KSFGHLCAAE VVKIRDSFSN QNTETKDTKS R
```

Figure 8. Predicted secondary structure of the mouse *DLC1* protein, performed using the FoldIndex program (http://bioportal.weizmann.ac.il/fldbin/findex). The mouse *DLC1* protein contains serine-rich region between the SAM and RhoGAP domains, which is predicted to have a largely unfolded conformation.

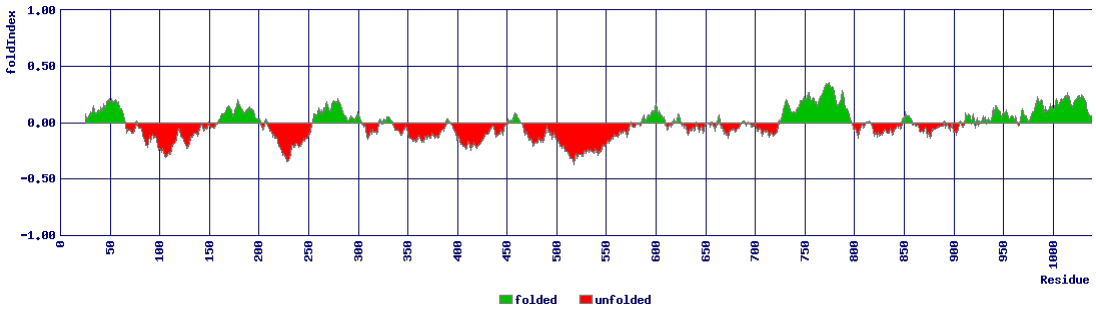
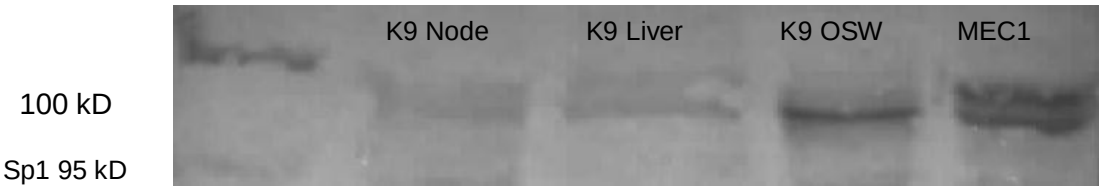


Figure 9. Mouse *DLC1* protein sequence showing features of the serine-rich region and the amino acid sequence of the region between the SAM and RhoGAP domains (approximate residues 79–639) of mouse *DLC1*.

```
   1 MCRDEPDTMI LTQIEAKEAC DWLRVTGFPQ YAQLYEDLLF PIDIALVKRE
  51 HDFLDRDAIE ALCRRLNTLN KCAVMKLEIS PHRKRSEDSD EDEPCAISGK
 101 WTFQRDSKRW SRLEEFDVFS PKQDPIPGSP DNSRLQSATS HESMLTDLSE
 151 HQEVASVRSL SSTSSSVPTH AAHSGDATTP RTNSVISVCS SGHFVGNDDS
 201 FSSLPSPKEL SSFSFSMKGH HEKNTKSKTR SLLKRMESLK LKGSHHSKHK
 251 APSKLGLIIS APILQEGMDE AKLKQLNCVE ISALNGNHIN VPMVRKRSVS
 301 NSTQTSSSSS QSETSSAVST PSPVTRTRSL STCNKRVGMY LEGFDPFSQS
 351 TLNNVTEQNY KNRESYPEDT VFYIPEDHKP GTFPKALSHG SFCPSGNSSV
 401 NWRTGSFHGP GHLSLRRENS HDSPKELKRR NSSSSLSSRL SIYDNVPGSI
 451 LYSSSGELAD LENEDIFPEL DDILYHVKGM QRIVNQWSEK FSDEGDSDSA
 501 LDSVSPCPSS PKQIHLDVDH DRRTPSDLDS TGNSLNEPEE PTDIPERRDS
 551 GVGASLTRCN RHRLRWHSFQ SSHRPSLNSV SLQINCQSVA QMNLLQKYSL
 601 LKLTALLEKY TPSNKHGFSW AVPKFMKRIK VPDYKDRSVF GVPLTVNVQR
 651 SGQPLPQSIQ QAMRYLRNHC LDQVGLFRKS GVKSRIQALR QMNESAEDNV
 701 NYEGQSAYDV ADMLKQYFRD LPEPLMTNKL SETFLQIYQY VPKDQRLQAI
 751 KAAIMLLPDE NREVLQTLLY FLSDVTAAVK ENQMTPTNLA VCLAPSLFHL
 801 NTLKRENSSP RVMQRKQSLG KPDQKDLNEN LAATQGLAHM IAECKKLFQV
 851 PEEMSRCRNS YTEQELKPLT LEALGHLNSD QPADYRHFLQ DCVDGLFKEV
 901 KEKFKGWVSY PTSEQAELSY KKVSEGPPLR LWRSTIEVPA APEEILKRLL
 951 KEQHLWDVDL LDSKVIEILD SQTEIYQYVQ NSMAPHPARD YVVLRTWRTN
1001 LPRGACALLL TSVDHDRAPV AGVRVNVLLS RYLIEPCGSG KSKLTYMCRA
1051 DLRGHMPEWY SKSFGHLCAA EVVKIRDSFS NQNTESKDTR SR
```

Figure 10: Western blot of canine and human tissues. The transcription factor Sp1 is a 95-kD protein. The antibody used for the ChIP was used for a western blot of canine lymph node, liver, a lymphoma cell line, and a human cell line served as a control. The antibody recognizes canine Sp1.



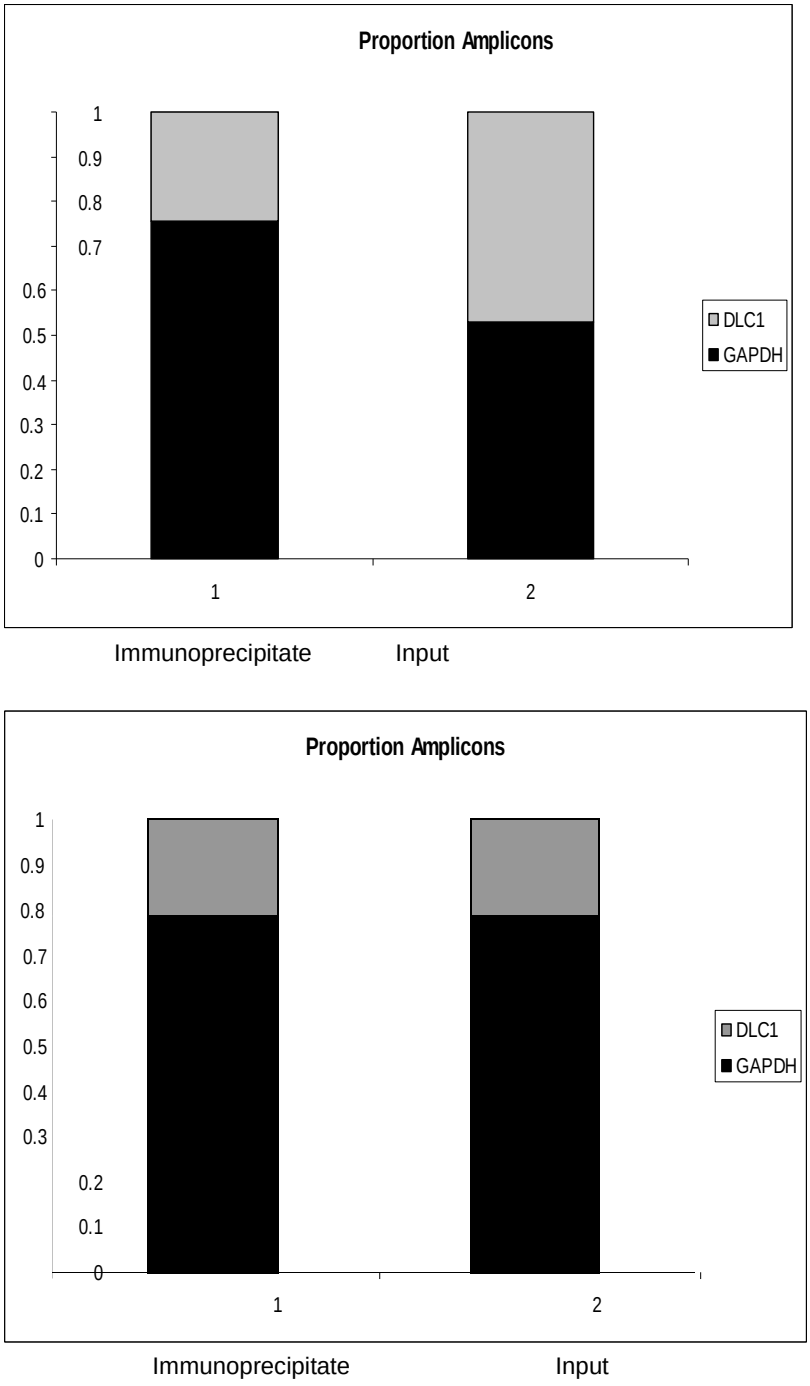


Figure 11: qRT-PCR of the relative binding of Sp1 to DLC1 as compared to GAPDH expressed in relative proportion of total amplicons. *DLC1* was present in relatively equal or fewer amplicons than GAPDH in the immunoprecipitate than in the input DNA sample. This suggests that expression of this gene is controlled to a small degree by Sp1.

Figure 12. Identification of the Sp1 binding sites in the canine *DLCJ* promoter region comprising 601 bp in chromosome 16 (start 36578234 and ends 36578834). The consensus sequences for Sp1 binding sites in the canine DLC1 are GGGCGG or GGGCGGG.

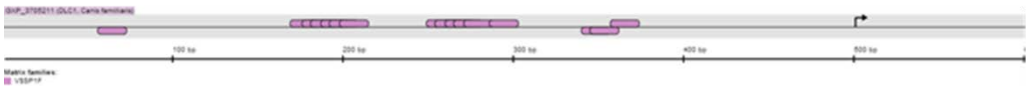


Table 1. RLM-RACE primers and RACE adapters used to amplify canine *DLC1* cDNA.

5' RACE Adapter	5'-GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA-3'
3' RACE Adapter	5'-GCGAGCACAGAATTAATACGACTCACTATAGGT12VN-3'
5' RACE Outer Primer	5'-GCTGATGGCGATGAATGAACACTG-3
5' RACE Inner Primer	5'-CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG-3'
3' RACE Outer Primer	5'-GCGAGCACAGAATTAATACGACT-3'
3' RACE Inner Primer	5'-CGCGGATCCGAATTAATACGACTCACTATAGG-3'

Table 2. Canine, Human, and Mouse DLC1 accession numbers, cDNA and amino acid sequence Sizes.

DLC1	cDNA Sequence (GenBank #)	cDNA size	Protein Size (Amino acid)	Protein Sequence
Canine	FJ602870	3724	1091	ACM44925.1
Human	AF026219	3766	1083	AAB81637.1
Human	AB051510	7365	1554	BAB21814 (AB051510.1)
Mouse	NM_015802	6184	1092	NP_056617

Table 3. Percent nucleotide identity of the *DLC1* in Canine, Human, and Mouse created by Clustal 12.1.

Organisms	Mouse	Canine	Human AF026219	Human AB051510
Mouse	100	81	81	75
Canine	81	100	88	85
Human AF026219	81	88	100	95
Human AB051510	75	85	95	100

Table 4. Percent amino acid identity matrix of the *DLC1* in Canine, Human, and Mouse created by Clustal 12.1.

Organisms	Mouse	Canine	Human AF026219	Human AB051510
Mouse	100	93	93	90
Canine	93	100	95	94
Human AF026219	93	95	100	99
Human AB051510	90	94	99	100