THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOLOGY

POTENTIAL IMPLICATIONS ON THE GIRAFFE DNA REPAIR PATHWAY THROUGH THE MODIFICATION OF MDC1

BRENDAN WOOD
SPRING 2015

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biology
with honors in Biology

Reviewed and approved* by the following:

Douglas Cavener
Professor of Biology & Interim Dean of the Eberly College of Science
Thesis Supervisor

Gong Chen
Professor of Biology & Holder of the Verne M. Willaman Chair in Life Sciences
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

MDC1 is a mediator DNA-Damage Checkpoint 1 (MDC1) and is instrumental in determining the course of double-stranded break (DSB) repair pathways. Its interactions with H2AX and BRCA1 as a scaffold protein make it necessary for recruiting DNA Double Strand Break Repair proteins. Without the control and signaling from MDC1, cells will divide uncontrollably with errors resulting in tumorous growths. This indicates that MDC1 is necessary in tumor suppression. Interestingly, we have recently discovered that the giraffe and okapi have a novel splice site in exon 5 of MDC1 which results in a truncated mRNA and ultimately 250 fewer amino acids in the protein. This change in tertiary structure of MDC1 may have promoted the formation of Robertsonian translocations and led to giraffe evolving to have a karyotype of 15 chromosome pairs rather than a karyotype of 30 chromosomes as found in fellow ruminant, the cow. Giraffe and okapi modifications of the MDC1 protein could contribute to our understanding of its function in carcinogenesis and its close interactions with H2AX and BRCA1.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Cavener for being an excellent mentor over the last three years and allowing me to pursue research in his lab. He has always provided me with the necessary resources to carry out my experiments and has allowed me to pursue developmental biology topics that I have found interesting. He has taught me to think about science critically and has spread his passion for science to me.

I would also like to thank Barbara McGrath for teaching me all the technical skills in the lab and for providing me with reference materials to teach me the different techniques necessary for becoming a successful scientist. She has been incredibly patient and understanding.

I would like acknowledge the graduate students in my lab who have become some of my best friends over my college experience. I would like to thank Siying Zhu for guiding me through my first two years in the lab. Sharing our successes and failures instilled in me the importance of team work in a research setting. I would also like to thank Carrie Lewis for aiding me in my scientific endeavors in almost every way throughout my undergraduate years. Her passion for science made me look forward to coming into the lab every day. Finally, I would like to acknowledge Carrie Lewis and Chelsea Hans for their guidance in the writing of my thesis.

INTRODUCTION

Giraffe (*Giraffa Camelopardalis*)

As the tallest land vertebrate in the animal kingdom, the giraffe has held a special place in African folklore as well as biology. Biologists like Jean-Baptiste Lamarck and Charles Darwin marveled at the long necks of the giraffe and based their theories of evolution on this feature. With necks extending to two meters, body masses of around 2,000 kg and distinctive coat patterns, giraffes can be seen browsing in the acacia trees across the African savannah.

Currently, all giraffe are grouped under the extant species *Giraffa Camelopardalis*. In recent years, the classification of giraffes as one or multiple species has been hotly debated. This is due to the geographic variation in pelage color patterns of giraffe. Through mitochondrial DNA sequences and nuclear microsatellite loci, it has been shown that there are at least six genealogically distinct lineages of giraffe in Africa with little interbreeding among the different lineages. One hypothesis is that giraffe lineages have become isolated due to the giraffe's preference for birthing in the wet season. Giraffe above and below the equator are therefore at a disadvantage if they breed with giraffes in the other hemisphere because of the reversal of seasons (Brown et al., 2007).

The giraffe is one of two species belonging to the family Giraffidae. This family also contains the *Okapia johnstoni* which is also known as okapi. The okapi contrasts the giraffe with a different morphology and existence in a different geographic location. The okapi is found exclusively in the forested environment of the north-eastern Democratic Republic of the Congo

(Fennessy, 2008). Due to its limited habitat and small population, the okapi is considered endangered according to the International Union for the Conservation of Nature (Mallon et al., 2013). The okapi is a medium sized animal with an average adult size of 250 kg and an average height of 1.5 m at the shoulders (Gijzen, 1958). As a comparison, zebra are about 100 kg heavier and 0.5-1 meter taller. Okapi can be identified by their dark velvety pelage with white horizontal stripes on the front and rear legs (Gijzen, 1959). As the closest extant ancestor to the giraffe, the okapi is important in regards to determining gene divergence in the giraffe.

The lifespans of giraffe and okapi tend to be similar. According to documented records, giraffe have been found to live for 38 years, a particularly long life span for free-ranging animals (Carey & Judge, 2000). Okapi has been found to live equally long with lifespans of 33 years (EAZWV, 2008). Researchers at the Clinic for Zoo Animals, Exotic Pets and Wildlife at the University of Zurich in Switzerland speculate that giraffids' (giraffe and okapi) longevity is due to their low metabolic rates. A low metabolic rate is also indicative of giraffe and okapi's extremely long gestation periods and low intrauterine growth rates (Muller et al., 2011). The average gestational period, time in which a fetus develops to birth, of the giraffe is on average around 460 days (Skinner & Hall-Martin, 1975). This tends to be a much longer pregnancy than similarly-sized land vertebrate pregnancies.

Similar to other ruminants, male and female giraffes' daily food consumption is 1.6% and 2.1% percent of their body weight, respectively (Pellow, 1984). During the wet season, the diet of both male and female giraffes is dominated by *Acacia tortilis*; however, during the dry season, this source of food is replaced by *Grewia* bushes, *G. bicolor* and *G. fallax* (Pellow, 1984). The change of preference is due to the proximity of *Grewia* bushes to sources of water, not a change in preference for taste or nutrients (Pellow, 1984). Through their elongated body form, giraffe

are able to avoid competition with smaller browsers by reaching into tall trees, therefore allowing them to ingest more leaf mass per bite (Cameron & du Toit, 2007). With giraffe's advanced ability to reach into tall trees come certain physiological challenges.

One of these challenges is delivering blood around the body. To remedy this issue, the giraffe requires a unique cardiovascular system. For one, when a giraffe has its neck at a natural position of 55° to the vertical, the giraffe has a carotid mean arterial pressure of 200 mm. Hg (Goetz et al., 1960). This represents twice the blood pressure of similar-sized mammals under normal physiological conditions (Goetz & Budtz-Olsen O, 1955). In order to pump blood with enough magnitude to the rest of the body, the giraffe's heart must have special features. Although the heart has a relative ventricular mass that is no different from other mammals (0.5% of body mass), it has a significantly thicker left ventricle wall (Brøndum et al., 2009). This equates to a stronger force of contraction, but a smaller stroke volume.

High blood pressure causes two potential issues in the giraffe. First, due to the height of the giraffe, an extremely high blood pressure would be created in the legs which would likely cause interstitial edema (Hargens et al., 1987). However, due to the lack of tissue compliance in the legs, the thick arteries in the legs acting as resistance arteries, and the structural adaptation of small muscle arteries to withstand extraordinary tension, the giraffe is able to prevent interstitial edema from occurring (Petersen et al., 2013). Second, due to the long neck of the giraffe, a substantial increase in blood pressure occurs in the neck and head when the neck is bent to drink water. To combat this rise in blood pressure, the giraffe has an anastomosis between the carotid and vertebral arteries that directs the blood into cerebral circulation (Goetz & Keen, 1957). In addition, the jugular vein contains valves that prevent backflow of blood into cranial circulation (Mitchell et al., 2009).

Giraffe have been known to have excellent vision. From birth to adulthood, vision changes drastically. Giraffe eye volume increases from 33 cm^3 to 65 cm^3, the focal length increases from 40 to 48 mm and the retina surface increases from 3000 mm^2 to 4320 mm^2 (Mitchell et al., 2013). In addition, the orbital axis angle becomes more acute, therefore, transforming their visual field to become more binocular and allowing them to see more three dimensionally. These factors are comparatively more impressive than any other ungulate (Mitchell et al., 2013). Figure 1 represents an image of the angle in which the eyes sit in the eye sockets. As one can see, through development, the eyes point more towards the front of the body allowing for a greater crossover of visual fields and ultimately a more three dimensional image.

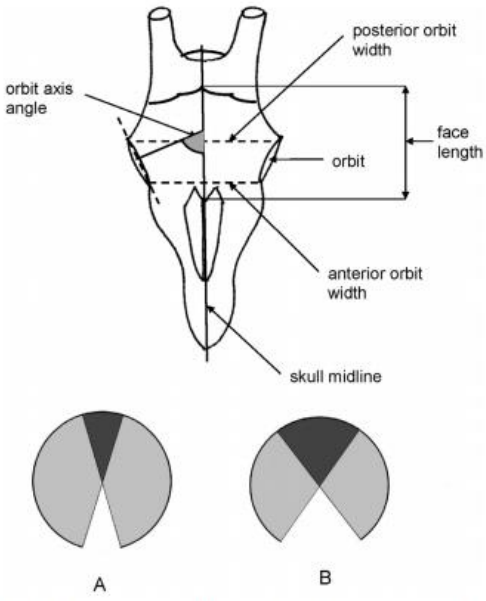


Fig. 1. Dorsal aspect of the skull of a giraffe showing the measurements made to determine face length and width and orbit orientation (see text) and diagrams showing the estimated changes in monocular (pale grey) and binocular (dark grey) fields of vision throughout life: A = neonate and B = mature (>1000 kg) giraffe. Note that in adult giraffe the horizontal field of vision is less monocular and more binocular than it is in neonatal animals, and the total field of vision is reduced.

Figure 1: Giraffe Binocular Vision Development (Mitchell et al., 2013)

In addition, like other artiodactyls, the giraffe contains a horizontal streak and temporal area in its vision, which means that it can improve resolution on the horizon and in the frontal visual field (Coimbra et al., 2013). Unlike other species, however, the giraffe has an advanced ability to view motion. Through its unique distribution of alpha ganglion cells in the eye, the giraffe can acquire motion information for controlling its tongue for browsing while identifying predators in the distance (Coimbra et al., 2013).

Although giraffes have extraordinarily long necks, like most vertebrates, the giraffe has only seven cervical vertebrae (van Sittert et al., 2010). The disproportionate cervical vertebrae extension in giraffe is not a fetal process. It occurs after birth throughout development into adulthood. The growth of each cervical vertebrae from C2-C7 occurs at the same rate (van Sittert et al., 2010), resulting in necks that can be 2 meters in length. The remaining vertebrae seem to be consistent in length will other ungulates (Badlangana et al., 2009).

Debate over the evolutionary explanation for the long necks of giraffe has occurred over the years. The most commonly acknowledged theories are those of natural selection and sexual selection. Charles Darwin hypothesized that those individuals with long necks were able to reach the tops of trees and browse more effectively than shorter necked individuals; therefore, during draughts, longer necked giraffes could survive and sire more offspring. This would have enhanced the genetic propensity for taller height (Gould, 1996). The sexual selection hypothesis is based on giraffe's behavior where they will combat other males in the presence of reproductively fertile females. Through combat known as 'neck sparring,' males will stand side by side and alternate swinging their heads at each other. Those giraffe with longer necks tend to win this competition due to the greater amount of torque that they create (Kingdon, 1979).

Ultimately these are the giraffe that copulate with fertile females and pass their genes to a new generation of giraffes (Kodric-Brown & Brown, 1984).

As expected, we found genes involved in axial skeletal and cardiovascular pathways in the giraffe that underlie their unique skeletal and cardiovascular attributes. Unexpectedly, however, we have found a number of genes that have diverged in double-strand break repair pathways. The focus of my thesis will be presenting this data and speculating the possible function of these genes.

DNA Double-Strand Break (DSB) Repair

DNA double-strand breaks (DSB) are thought to be the most dangerous type of DNA damage that can occur within the cell (Khanna & Jackson, 2001). As a result, genome integrity signaling and repair pathways have evolved to combat these aggressive and at times cancerous changes to the cell (Bohgaki et al., 2010). The cell responds to DNA double-strand breaks through a variety of means depending on the severity of the break. Some possible responses to breaks are cell cycle arrest, DNA repair and apoptosis (Khanna & Jackson, 2001).

DNA double-strand breaks can occur through a variety of fashions, some intentional and others unintentional. Some intentional developmentally programmed DSBs are V(D)J recombination, class switch recombination (CSR) or meiosis (Goodarzi & Jeggo, 2013). V(D)J recombination is a way in which antigen receptor genes are diversified. This diversification is critical to the adaptive immune system. RAG1 and RAG2 proteins initiate the double-strand breaks and through a non-homologous end joining mechanism (to be discussed later), the strands

are recombined (Schatz, 2004). Class switch recombination functions by replacing immunoglobulin genes which results in altering immunoglobulin isotypes from IgM to IgG, IgE, or IgA. This process is critical for the proper functioning of mature B lymphocytes (Honjo et al., 2002). Double-strand breaks are important in meiosis because it allows for genetic variation in gametes. Spo11 is a significant player in this process, for it is a catalytic subunit for meiotic DNA cleavage activity (Keeney et al., 1997).

Some unintentional sources of DNA DSBs are ionizing radiation like UV radiation and ROS (Reactive Oxygen Species) and replication fork collapse (Goodarzi & Jeggo, 2013). Replication fork collapse is triggered by an upstream single-strand break or base damage. When the replication fork collides with this damaged segment of DNA, replication stops and the fork collapses (Goodarzi & Jeggo, 2013).

When an unintentional DSB happens, a signal transduction process occurs to alert the cell of DNA damage. First, the MRN complex, consisting of MRE11/RAD50/NBS1, senses the DNA double-strand break (Goodarzi & Jeggo, 2013). The process begins when the kinase ATM interacts with the C-terminus of the NBS1 protein, resulting in the binding of ATM to the DSB location (Goodarzi & Jeggo, 2013). ATM will then phosphorylate H2AX, a histone variant specifically charged with the function of double-strand break repair. If needed, DNA-PK and ATR proteins can fulfill ATM's function in phosphorylating H2AX; however, downstream steps in the signal transduction pathway require ATM (Stiff et al., 2004). The phosphorylated form of H2AX, γH2AX, initiates the assembly of the DNA damage response (DDR) proteins near the break and also influences protein assembly at the DSB. The DDR proteins restructure histones around the DSB to prepare for the repair process. A series of other modifications occur under the influence of γH2AX including ubiquitylation (addition of ubiquitin), sumoylation (addition

of proteins similar to ubiquitin) and methylation (addition of a methyl group) (Goodarzi & Jeggo, 2013).

γH2AX enables recruitment of MDC1, the first of many mediator proteins. MDC1 acts as a scaffold protein, providing another way to tether MRN at the DSB and also recruiting two ubiquitin ligases, RNF8 and RNF168, to ubiquitylate histone H2A. The ubiquitylated histone H2A promotes exposure to H4K20me3 which ultimately aids in the recruitment of 53BP1 (Goodarzi & Jeggo, 2013). The role of 53BP1 is to tether ATM at DSBs, to release CHD3, a chromatin remodeling protein, and to relax heterochromatin (Noon et al., 2010). In late S/G2 phase cells, another set of proteins accumulate at the DSB site due to the actions of BRCA1. BRCA1 has a unique relationship with 53BP1. It has been speculated that during the G2 phase of the cell cycle, BRCA1 displaces 53BP1 to bring about a DSB repair process known at homologous recombination (HR) (Kakarougkas et al., 2013). However, 53BP1 is still required to carry out HR in G2. On the other hand, 53BP1 is dispensable during the late S phase of homologous recombination. In the G1 and G0 phases of the cell cycle, where BRCA1 is not introduced, 53BP1 promotes a repair process known as non-homologous end joining (Kakarougkas et al., 2013).

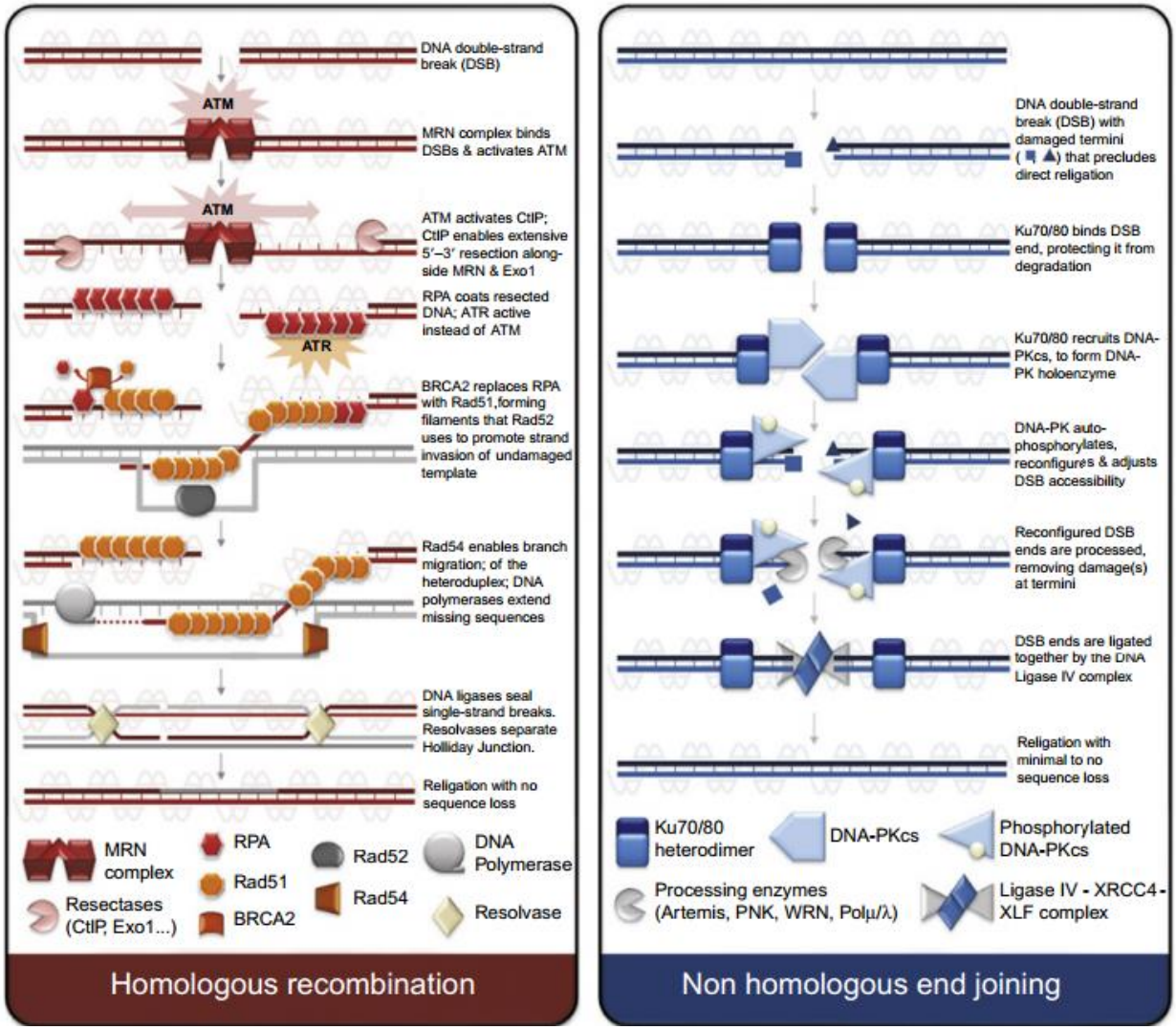


Figure 2: Homologous Recombination and Non Homologous End Joining Diagrams (Goodarzi & Jeggo, 2013)

Represented in Figure 2 above, Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ) are the two major methods of ligating a DNA double-strand break (DSB). These two methods are functionally different. Homologous recombination is capable of restoring any lost sequence information (Goodarzi & Jeggo, 2013). In order to function, it requires an undamaged template; however, this template cannot be a homologous chromosome. Only a sister chromatid can function in this role because the heteroduplex, the hybrid DNA,

cannot be assembled under the amount of variation that exists between homologous chromosomes (Goodarzi & Jeggo, 2013). This means that homologous recombination will only occur during late S and G2 phases (Johnson & Jasin, 2000).

HR begins with the generation of 3' single-stranded DNA tail through the process of resection. Resection consists of an initiating event and an elongation event. The initiating event involves proteins CtIP and Mre11 and the elongation step involves exonuclease I (ExoI) or a complex of Sgs1-Top3-RMI1(STR)/DNA2/BLM (Goodarzi & Jeggo, 2013). These proteins function together to release nucleotides in the 5' to 3' direction, forming a single-strand 3' end at each end of the DNA where the DSB has occurred. Following resection, RPA coats the resected DNA for stabilization. RPA will activate the kinase, ATR. BRCA2 will then replace RPA with RAD51, a necessary component for strand exchange and heteroduplex formation. BRCA2 interacts with RAD51 through its BRCT domain and C-terminal RAD51 interaction domain (Jensen, Carreira, & Kowalczykowski, 2010). RAD52 promotes strand invasion of an undamaged template with the assistance of RAD54, BLM, WRN, RECQ1, and RECQ5 which act in helicase manners (Larsen & Hickson, 2013). The heteroduplex is formed, and DNA polymerases extend the 3' end of the DNA strand using the undamaged stand as a template. Finally, DNA ligase anneals the single-strand breaks and resolvases separate the Holliday junctions, the junctions formed by the DNA-strands crossing over (Goodarzi & Jeggo, 2013).

NHEJ is the common pathway utilized for DNA double-stand break repair brought about due to ionizing radiation like UV radiation and reactive oxygen species (ROS). It occurs during G1 and G0 of the cell cycle. Due to the nature of ionizing radiation, double-strand breaks tend to occur with damaged terminal ends at the break. An exposed single strand of DNA might appear or nucleotide or sugar structures could be altered. NHEJ is initiated by the Ku70-Ku80 (Ku)

heterodimer. Ku is a basket-shaped molecule that allows double-strand DNA to thread through it via interactions with the DNA backbone (Walker et al., 2001). The Ku heterodimer prevents nucleolytic degradation and recruits the DNA-PK catalytic subunit, forming the DNA-PK holoenzyme. DNA-PK is activated through autophosphorylation and reconfigures DSB assembly. Next, end processing of the DSB occurs where non-ligatable ends such as 5' OH or 3' phosphates and damaged nitrogenous bases are cleaved or repaired. One of the DNA repair proteins is polynucleotide kinase 3' phosphatase (PNKP). PNKP acts as phosphatase and kinase by removing 5' phosphates and adding 3' phosphates, respectively (Bernstein et al., 2009). Finally, with XRCC4 for stability, DNA ligase IV binds the two DNA double strands (Riballo et al., 2009).

Considering all of the previously mentioned proteins, H2AX and MDC1 in giraffe showed more divergence from closely related species than any of the other DNA double-strand break repair proteins. The central role that these two proteins play in the process and their degree of divergence warranted further exploration and analysis.

MATERIALS AND METHODS

Investigation of Giraffe Genes

In order to examine the divergence of the genes in giraffe, a whole genome sequencing of *Giraffa camelopardalis* and *Okapia johnstoni* was accomplished by the Genomics Core Facility at the Pennsylvania State University. These sequences were compiled by Webb Miller and his group. In addition, his group identified protein coding sequences. Following the protein coding sequence identification, Douglas Cavener and his group at Pennsylvania State University along with collaborators at the Nelson Mandela African Institute of Science and Technology in Tanzania did comparative genomic analysis of okapi and giraffe to identify genes that have uniquely diverged in giraffe. The genes from giraffe with the top 15% of genetic divergence from the most closely related vertebrates were identified and grouped according to function. These sequences were then translated to amino acid sequences to determine the amount of functional change that had occurred.

Amplification and Sequencing

Douglas Cavener was responsible for identifying genes in which a significant amount of genetic changes had occurred. Although some interesting changes in cardiovascular genes as well as developmental genes were found, genes involved in cell division and double strand break repair appeared to be some of the most interesting.

MDC1, a scaffold protein, was the first gene from DNA double-strand break repair that was examined. Table 1 depicts the cDNA samples that were chose for sequencing and their concentrations. Nzoo Giraffe samples were taken from the Nashville Zoo in Tennessee.

Table 1: Giraffe and Okapi cDNA Samples

Number Designation	Sample	RNA ng/µl
1	Nzoo Masai Fetal Giraffe Heart	398.2
2	Nzoo Masai Fetal Giraffe Liver	27.9
3	Nzoo Masai Adult Giraffe Heart	31.6
4	Nzoo Masai Adult Giraffe Liver	191.5
5	Adult Okapi Liver	367.4
6	Fetal Okapi Liver	557.2

The second DNA double-strand break repair gene analyzed was H2AX. For this gene, the genomic DNA of Masai Giraffe, Reticulated Giraffe, and Rothschild Giraffe were utilized. For H2AX and MDC1, the cDNA and genomic DNA were amplified using primers designed by Douglas Cavener. These consisted of 20 nucleotide strands that were specific to the gene in which I was amplifying. Listed in Table 2 is a list of primers for H2AX and MDC1.

Table 2: MDC1 and H2AX Primers

Gene and Primer Pair #	Forward or Reverse	Template strand 5' → 3'	Melting Temp. (Tm) °C
MDC1 Primer Pair: 8	Forward	GCTTCTGTTGGCTGAGGACT	57.3
	Reverse	CGAGGGCAGTTTCTTGGTCT	57.3
MDC1 Primer Pair: 10	Forward	GGTCCTGCCCTTTTCATCCA	57.5
	Reverse	GGCCACTGCTAAGGATGTGT	57.4
MDC1 Primer Pair: 13	Forward	AGGAGCTGCCAGAAGAGGTA	57.7
	Reverse	CCTTCCGAGGGCAGTTTCTT	57.4
MDC1 Primer Pair: 24	Forward	TCTGTGGTCCTGCCCTTTTC	57.4
	Reverse	GTTGCTCTTGCTGTGCTGTC	56.9
MDC1 Primer Pair: 29	Forward	CTCTGTGGTCCTGCCCTTTT	57.4
	Reverse	TCTCTGTCCTGTATCCCCCG	58.0
MDC1 Primer Pair: 7	Forward	ACGAAGCTGGTACACAGAGC	57.2
	Reverse	CTGTGTCGACAAGACCTCCC	57.5
MDC1 Primer Pair: 9	Forward	AGGAGAGTCCAGCTGATGGT	57.7
	Reverse	TCGGACAGAATGGCTGTGTC	57.2
H2AX Primer Pair: 1	Forward	GGGTGGCTGTTCCTGTTTTG	56.9
	Reverse	CTTGTTGAGCTCCTCGTCGT	57.1
H2AX Primer Pair: 2	Forward	CCTCATACAGCATGTCGGGC	58.2
	Reverse	GCTGGTCTTCTTGGGTAGCA	57.1

For PCR, a 20 µl concoction was predominantly used to isolate the correct gene fragment and a 50 µl concoction was used to amplify the fragment sufficiently so that the concentration after purification was high enough to send for sequencing. Table 3 presents approximations for the amount of each solution needed for PCR. The amount of water compared to the amount of cDNA or genomic DNA varies depending on the genetic material's concentration.

Table 3: PCR Concoction Components

Solution	20 µl PCR Concoction	50 µl PCR Concoction
KOD Hot Start Master Mix	10 µl	25 µl
Forward Primer	0.6 µl	1.5 µl
Reverse Primer	0.6 µl	1.5 µl
Nuclease Free H_2O	4.8 µl	20 µl
Genomic DNA or cDNA	4 µl	2 µl

KOD Hot Start Master Mix (Cat: 71842, Novagen)

The KOD Hot Start Master Mix contained important components for DNA replication like KOD Hot Start DNA polymerase, two monoclonal antibodies, ultrapure deoxynucleotides, and reaction buffer with $MgSO_4$ (Novagen). After running PCR according to the time and temperature specifications given by the KOD Hot Start Master Mix recommendations, gel electrophoresis was conducted. Each DNA sample was mixed with one microliter of 6X gel loading dye (CAT: 1B1010) so that the locations of the samples were visible in the gel. After the gel was run for 40 minutes under low voltage (75V) with an adjacent ladder for reference, an image was taken under ultraviolet light. The bands of this image were analyzed according to length and were compared to the predicted length of the gene. The bands that most closely

matched with the proposed fragment length were consequently excised from the gel with a razor. These "cubes" of gel were grinded using mortar and pestle and mixed with 1X TAE buffer.

After vortexing and spinning this sample, the resulting genetic material was amplified again using a 50 µl PCR concoction, described in the Table 3. A gel was run with 5 µl of DNA and a microliter of loading dye to confirm that the band still existed, and the remaining amplified DNA was purified. DNA was purified using the QIAquick PCR Purification Kit (Qiagen Cat: 28106). After following Qiagen's procedure, the DNA concentration was identified with a biophotometer. Blanks of H_2O were used as reference points in calculating concentration of purified samples. The guidelines for concentration required for sequencing was made available by dnaLIMS (Penn State DNA Sequencing Ordering and Sequence Retrieval System).

Analysis of Changes

Once the DNA data sequences from the Penn State DNA Sequencing and Ordering Sequence Retrieval System were returned, the National Center for Biotechnology Information's program called BLAST (Basic Local Alignment Search Tool) was utilized by inserting the newly sequenced genes. This program was able to determine whether the gene fragment aligned with the same gene in other organisms or if the gene was in any way altered from closely related species. Since amino acid information is more important functionally in the body than DNA sequences, the DNA sequences would be translated into amino acid sequences. After this, amino acid sequences were copied and pasted into a BLAST and alignment software was run. After alignment, a phylogenetic tree was built to determine the amount of genetic divergence and to

discover which organisms were most closely related to each other. Differences in amino acid sequences are often times the product of non-synonymous SNPs (single nucleotide polymorphisms). These occur when one nucleotide is switched for another nucleotide. Some amino acid substitutions are more significant than others. For example, replacing a nonpolar group with a very acidic group can cause conformational changes in the tertiary structure of the protein. This could ultimately cause the protein to be ineffective or cause the protein's function to change.

RESULTS

Unique Divergence of H2AX in Giraffe

The initial analyses of H2AX indicated that giraffe from the Nashville Zoo contained seven unique amino acid substitutions not present in other mammals. As one can see in the alignment in Figure 3, H2AX is highly diverged in the Nashville Zoo giraffe.

Figure 3: H2AX Protein Alignment for Cow and Related Species

```
human       1    MSGRGKTGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGHYAERVGAGAPVYLAAVLEYLT  60
NZoogiraffe 1    ................W........S....M.......RH...A.................  60
Okapi       1    ..........S.....................................................  60
Cow         1    ..........S.....................................................  60
gerenuk     1    ..........S.....................................................  60
Watbuff     1    ..........S.....................................................  60
Manatee     1    ..........S.....................................................  60
Orca        1    ..........S.....................................................  60
Bushbaby    1    ..........S.....................................................  60
Hedgehog    1    ..........S.....................................................  60
mouse       1    ..........S.....................................................  60

Query   61   AEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGGVTIAQGGVLPNIQAVLLPKK  120
156022  61   ..................G.....................................  120
156024  61   ..........................................................  120
156025  61   ..........................................................  120
156026  61   ..........................................................  120
156028  61   ..........................................................  120
156029  61   ..........................................................  120
156030  61   ..........................................................  120
156031  61   ..........................................................  120
156032  61   ..........................................................  120
156033  61   ..........................................................  120

Query   121  TSATVGPKAPSGGKKATQASQEY  143
156022  121  ..........A............  143
156024  121  ..........A............  143
156025  121  ..........A............  143
156026  121  ..........A............  143
156028  121  ........T.A............  143
156029  121  ..........A....T.......  143
156030  121  ...A......A............  143
156031  121  ........T.A....T.......  143
156032  121  ......A...A....S.......  143
156033  121  S.........AV....S......  143
```

For a histone protein variant, which is highly conserved across animals and plants because of its function in DNA folding and chromatin formation, it seemed unusual that so many substitutions had occurred in giraffe.

To confirm these amino acid differences, H2AX was amplified from genomic DNA in two Masai giraffe (one being the same sample as originally sequenced), one reticulated giraffe and one Rothschild giraffe. I began by amplifying the genomic DNA using H2AX primers 1 and 2. After PCR amplification, electrophoresis was conducted to decide whether or not the amplification was successful. The results are listed in Table 4.

Table 4: H2AX PCR Results after Genomic DNA Amplification

Primer Pair	Sample	Band Present?
1	Nzoo Masai Giraffe Genomic DNA	Present
2	Nzoo Masai Giraffe Genomic DNA	Absent
1	Reticulated Giraffe Genomic DNA	Present
2	Reticulated Giraffe Genomic DNA	Absent
1	Rothschild Giraffe Genomic DNA	Present
2	Rothschild Giraffe Genomic DNA	Absent

All samples using primer pair 1 were successfully amplified while the samples using primer pair 2 were not amplified. The bands that were present (Primer Pair 1 with samples Nzoo Masai, reticulated, and Rothschild Giraffe) were then amplified again in a 50 µl PCR concoction. Each of these samples created bands in gel electrophoresis. Then, each sample was sent to the Penn State University Nucleic Acid Facility for sequencing. The sequences obtained from the nucleic acid facility would be compared to the original H2AX genomic sequencing results from Nzoo giraffe. The genomic sequencing results of H2AX from the original Nzoo giraffe genome are presented in Figure 4. This sequence is consistent with Nzoo giraffe from Table 4.

Figure 4: H2AX Protein from Original Genome Sequencing of Nzoo

MSGRGKTGGKARAKAKWRSSRAGLQSPVGRMHRLLRKGRHAERAGAGAPV
YLAAVLEYLTAEILELAGNAARDNKKTGIIPRHLQLAIRNDEELNKLLGG
VTIAQGGVLPNIQAVLLPKKTSATVGPKAPAGGKKATQASQEY*

Listed here are the raw DNA sequencing results from the Penn State University Nucleic Acid Facility and their corresponding amino acid chains starting with the start codon, methionine. Note that the DNA before methionine is not present in the amino acid sequence because it is a part of the introns.

Figure 5: H2AX Nucleotide and Amino Acid Sequences

>Nzoo_H2AX_2-GCAH2AX_R1 508 25 287 0.05
TATCCGCGTCTTCTTGTTGTCGCGGGCCGCGTTGCCCGCCAGCTCCAGGATCTCAGCGGTGAGGTACTCGAGTACCG
CCGCCAGGTAGACCGGCGCTCCGGCGCCCACCCTCTCGGCGTAGTGGCCCTTCCGCAGCAGCCGGTGCACGCGGCCC
ACCGGAAACTGGAGGCCCGCTCGCGAAGAGCGCGACTTAGCCTTGGCGCGAGCCTTGCCGCCGGTCTTGCCGCGGCC
CGACAT
>Nzoo H2AX Amino Acid
MSGRGKTGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGHYAERVGAGAPVYLAAVLEYLTAEILELAGNAARDNKKT
RI
>Reticulated_H2AX_2-GCAH2AX_R1 593 29 283 0.05
TATCCGCGTCTTCTTGTTGTCGCGGGCCGCGTTGCCCGCCAGCTCCAGGATCTCAGCGGTGAGGTACTCGAGTACCG
CCGCCAGGTAGACCGGCGCTCCGGCGCCCACCCTCTCGGCGTAGTGGCCCTTCCGCAGCAGCCGGTGCACGCGGCCC
ACCGGAAACTGGAGGCCCGCTCGCGAAGAGCGCGACTTAGCCTTGGCGCGAGCCTTGCCGCCGGTCTTGCCGCGGCC
CGACAT
>Reticulated H2AX Amino Acid
MSGRGKTGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGHYAERVGAGAPVYLAAVLEYLTAEILELAGNAARDNKKT
RI

>Rothschild_H2AX_2-GCAH2AX_R1 589 28 40 0.05
TATCCGCGTCTTCTTGTTGTCGCGGGCCGCGTTGCCCGCCAGCTCCAGGATCNAAGCGGTGAGGTACTCGAGTACCG
CCGCCAGGTAGACCGGCGCTCCGGCGCCCACCCTCTCGGCGTAGTGGCCCTTCCGCAGCAGCCGGTGCACGCGGCCC
ACCGGAAACTGGAGGCCCGCTCGCGAAGAGCGCGACTTAGCCTTGGCGCGAGCCTTGCCGCCGGTCTTGCCGCGGCC
CGACAT
>Rothschild H2AX Amino Acid
MSGRGKTGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGHYAERVGAGAPVYLAAVLEYLTAXILELAGNAARDNKKT
RI

An alignment of the previous amino acid sequences with the H2AX histone sequence from the original genome sequencing was carried out. The result of this comparison is displayed in Figure 6.

Figure 6: H2AX Alignment of Giraffe subspecies and Cow

Bos Taurus H2AX Sequence: A
Original MA1 (Masai) Giraffe: B
Nzoo (Masai) Giraffe: C
Reticulated Giraffe: D
Rothschild Giraffe: E

```
A 1    MSGRGKTGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGHYAERVGAGAPVYLAAVLEYLT  60
B 1    .................W........S....M.......RH...A.................  60
C 1    ............................................................  60
D 1    ............................................................  60
E 1    ............................................................  60

A 61   AEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGGVTIAQGGVLPNIQAVLLPKK  120
B 61   ..................G.......................................  120
C 61   ...................                                         79
D 61   ...................                                         79
E 61   .X.................                                         79

A 121  TSATVGPKAPAGGKKATQASQEY  143
B 121  ........................  143
```

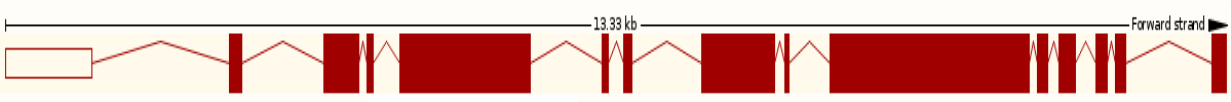
The comparison of H2AX demonstrated that every giraffe sample had the same amino acid sequence as cow. This included the reticulated giraffe, the Rothschild giraffe and the same Nzoo Masai giraffe that had been sequenced before in the full genome sequencing. It can therefore be determined that the amino acid changes found in H2AX from the full genome

analysis were inaccurate. The giraffe has not had any significant changes to H2AX over its evolutionary history and is consistent with its most closely related species.

Truncated Exon 5 of MDC1

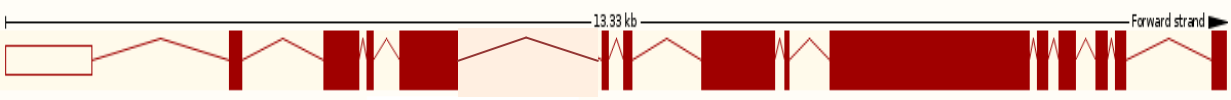
MDC1 initially interested our group because of its close association as a scaffold protein with H2AX. After further inspection of MDC1, we found a series of stop codons inserted into exon 5 of the giraffe MDC1 gene. The implications of this are substantial. Encoding for these stop codons would eliminate most of the amino acids in the MDC1 protein. This motivated us to find a novel splice site within MDC1. Finding this novel splice site would mean that the stop codons were no longer within an exon and would be in an intron instead. At 750 nucleotides past what we believed to be the start of exon 5 was a nucleotide substitution. The nucleotide substitution altered the sequence and transformed it into a location where a splice would occur. Figure 7 shows what the giraffe MDC1 was originally thought to exist as in terms of exons and introns. Figure 8, in contrast, is the new model of giraffe MDC1with a truncated exon 5. Exons are represented as boxes and introns are represented as lines.

Figure 7: Sheep MDC1 Structure (ensembl.org)



-- Exon 5 --

Figure 8: Proposed Giraffe MDC1 Structure (ensembl.org)



Exon 5

As seen in Figure 8, exon 5 is truncated in half due to the novel splice site. The truncation was a probable alteration for two reasons. The reading frame was maintained with the truncation; therefore, it did not affect the translation of downstream codons. Also, the remaining segment of the protein contained a higher proportion of synonymous nucleotide substitutions to non-synonymous nucleotide substitutions demonstrating that the protein still retained its function. If more non-synonymous substitutions were present, then the creation of an intron sequence would have been considered.

My goal for MDC1 was to find out whether this novel splice site actually existed. The way in which I was going to prove that was by analyzing the mRNA of giraffe MDC1. If the mRNA did not contain the 750 nucleotides before the proposed splice site, I would determine that indeed a novel splice site existed; however, if the 750 nucleotides before the splice site did exist in mRNA, then I would consider that those stop codons truly existed and that MDC1 protein was missing a majority of its amino acids. In order to prove the existence of the splice site or lack thereof, I PCR amplified MDC1 in a variety of giraffe and okapi cDNA samples (found in Table 1). Previous experiments conducted had demonstrated that these cDNA samples were extremely difficult to amplify, therefore obtaining enough DNA for sequencing to prove or dismiss my hypothesis required a great deal of time and resources.

After amplification of MDC1 using giraffe and okapi cDNA, gel electrophoresis was conducted. The results are listed in Table 6. The numbers listed across the top axis correspond to the giraffe and okapi samples listed in Table 1.

Table 5: MDC1 Gel Electrophoresis Results after Initial Amplification

Primer Pair	1: Fetal Gir. Heart	2: Fetal Gir. Liver	3: Adult Gir. Heart	4: Adult Gir. Liver	5: Adult Okap. Liver	6: Fetal Okap. Liver
8	Absent	Absent	Absent	Absent	Absent	Absent
10	Absent	Absent	Absent	Absent	Absent	Absent
13	Absent	Absent	Absent	Absent	Present	Present
24	Absent	Absent	Absent	Absent	Absent	Present
29	Absent	Absent	Absent	Present	2 Present	Present

Each box designates whether a band of the predicted size was present in the gel or if no band similar to the predicted size was found. The predicted band lengths from each primer pair varied. Primer pair 8 was expected to have a product length of 1963 nucleotides, primer pair 10 was expected to have a product length of 1513 nucleotides, primer pair 13 was expected to have a product length of 1767 nucleotides, primer pair 24 was expected to have a product length of 2110 nucleotides, and primer pair 29 was expected to have a product length of 2325. Products that were under the predicted product length were considered because of the possibility of a deletion in exon 5.

As expected, very few of the PCR amplifications of giraffe and okapi cDNA yielded any bands in gel electrophoresis. The few bands that did appear were removed from the gel and amplified again in an effort to obtain enough cDNA for sequencing. The results of this experiment can be found in Table 7.

Table 6: MDC1 Gel Electrophoresis Results after Amplification of Amplified cDNA

Primer Pair	Sample & Number	Band Present?
24	6: Fetal Okapi Liver	Present
29	4: Adult Giraffe Liver	Absent
29	5: Adult Okapi Liver - Longer	Absent
29	5: Adult Okapi Liver - Shorter	Absent
29	6: Fetal Okapi Liver	Present
13	5: Adult Okapi Liver	Present
13	6: Fetal Okapi Liver	Present

As Table 7 indicates, the giraffe sample and two okapi samples no longer showed any bands; therefore, it can be assumed that these sequences were either too fragile to be amplified

further, or that the bands were not actually MDC1. The okapi cDNA bands that showed successful amplification were amplified again for a third time using the 50 µl PCR concoction. The purpose of this experiment was to create enough replicates of MDC1 cDNA for sequencing. The gel electrophoresis results in Table 8 are based on 5 µl of cDNA with 1 µl of loading dye.

Table 7: MDC1 Gel Electrophoresis Results after Third Amplification

Primer Pair	Sample & Number	Band Present?
24	6: Fetal Okapi Liver	Present
29	6: Fetal Okapi Liver	Absent
13	5: Adult Okapi Liver	Absent
13	6: Fetal Okapi Liver	Present

After this round of amplification, I purified and sequenced the fetal okapi liver sample amplified by primer pair 24 and the fetal okapi liver sample amplified by primer pair 13. In Figure 9, I have presented the raw cDNA sequence produced by the Penn State University Nucleic Acid Facility's DNA Sequencing Ordering and Sequence Retrieval System. Due to the amount of undetermined nucleotides from the okapi sample using primer pair 24, only primer pair 13 is shown.

Figure 9: Primer Pair 13 MDC1 Fragment of Fetal Okapi Liver

```
>#6: Fetal Okapi Liver – MDC1 Fragment – Primer 13 Forward
GCCGACTCCAGATTGAGATGGATCAGTGTTCCGTGAAGGAGAGGAACAGGGACACGGAAATCAAGAGGGATGTGAGG
AATGGGGTGGTTCCAATTGGAGTGATTCTGAAGAGGAGCCAGTCTTCTGGGGAGGACAGTGACACATATGTGAGTGA
TGAGAGTGGGCCTAAAAAAAGGCTTGGTGGGGTCCCTCCGAAAAGGGTCTGGTCTTGTCTCTTCATAGGCAGTGATA
CCAGTGGGAAAGACGAGGGGATCCCTGCTACCCCAGCAGTGGTTCCCGTGAAAGAGAGACAGATTTTCCACGAATCT
GGTACACAGAGCCCCCAGGCACCTGGTGTGGCACGTCGGCAGGAGAGCCCATCTGATGGTGATACATATATAGAGGA
GGGGGACGCCCCCCCACACAGAAGCCAACCCTCCGTGGTGATAGACCCCAATACATATGAAGTCGCCGCACAATTGA
CGCTGACAGATCAAAGAGACACCCACTCTGATTTTGAAGATCTGCACATAACAGATGTAGTGTGCTTTATGAAGAGA
CAGAATCAAAGACTGGTAGGAGACCGCCACTGCGAGGAAGCAGCTGAAGACTTTATATGTTATCACTGTCGGAGGAG
CAGTTCGTTCTTTGCAATCAATCCACTCTGGCCAGCTGACGAATACTGGAAAGCGATCTGTCAGATACCAGAAATAT
TGCATACAGTAAATGAAAGACACATACCGACGTGCCAACCCGGGTATCCCCGAGCAGTGGATAATTATTCAGTCACT
```

CTACGTTGATAACCGAGCACTACACCGCACTATGAAACCAGCTGTAGCAACCCACAAGAAAAATACTGGTTTAGCAC
TCCGACAGTACTGGATCCGCTAGAGCACCAAGGTATGAAGATAATAGTGAAGCCAACTCATCCATTGCAATGAA

In order to prove that the fetal okapi liver MDC1 was consistent with the novel splice site hypothesis, the sequence was compared to the giraffe truncated exon 5 model. In the Appendix, one can find a model of giraffe MDC1 before the truncation and after the truncation. The giraffe sequences are based on the exon portions in cow. A comparison was made between the giraffe model with truncated exon 5 and the fetal okapi liver MDC1 sequence to determine whether or not the two sequences aligned. Alignment would demonstrate that the proposed portion of exon 5 was missing and that the novel splice site truly existed. Figure 10 is the alignment of these two sequences.

Figure 10: Okapi CDS Compared to Model of Giraffe MDC1 with Truncated Exon 5

Giraffe MDC1 Model with Truncated Exon 5: A Okapi MDC1 coding sequence: B

```
A 145  GCCGAAATCCAGATTGAGAAGGATCAGTGTTCTGTGAAAGAGAAGAACAGAGACACAGAA  204
B 31   ....-.C............T............C.....G....G......G.....G...  89

A 205  ATCGAGAGGGATGCAAGGAATGGGGTGGTTCCAACTGGAGTGATTCTGAAGAGGAGCCAG  264
B 90   ...A.........TG..................T.......................  149

A 265  CCTTCTGGGGAGGACAGTGACACAGATGTGAGTGATGAGAGTGGGCCTCCAAGAAGGCTT  324
B 150  T......................T......................AA..A.......  209

A 325  GCTGGGGTCCCTCCGAAAAGGGCCTGGTCTTGTAACTTCATAGGCAGTGATACCAATGGG  384
B 210  .G...................T..........CT...................G....  269

A 385  AAAGACGAAGGGATCCCTGCTACCCCAGCAGTGGTTCCCGTGAAGGAGAGGCAGATTTTC  444
B 270  ........G.................................A.....A.........  329

A 445  CACGAAGCTGGTACACAGAGCCCCCAGGCACCTGGTGTGGCACGTCAGCAGGAGAGTCCA  504
B 330  ......T.......................................G.........C...  389

A 505  GCTGATGGTGATACAGATATAGAGGAAGGGGAGGcccccccGGACAGAAGCCAAGCCTCC  564
B 390  T.............T..........G.....C........AC...........C.....  449

A 565  GTGGTGATGGACAGCAATACAGACGAAGTCGCGGCAGCATTGACACTGGCACATCTAAGA  624
B 450  ........A...CC.......T.T........C...CA......G...A..G...A....  509

A 625  GAGAGCCACGCTGATTTTGAAGATCTGGACCTACCAGCTACCCAGTGCTTTGTAGACAGA  684
B 510  ..C.C....T................C..A..A...A.GTAGT.......A.-..-...  567
```

```
A 685  GAGAATCAGAGCCTGGAAGCAGTCCCCAGCGTGGAGGA  722
B 568  ........A..A....T..G..A.....-.-.......  607
        \                        \       \
        |                        |       |
       AC                       G       C
```

This alignment shows the proper matching of the fetal okapi liver MDC1 cDNA with the proposed Giraffe model. Although there are many nucleotide differences (due most probably to the large number of rounds of replication), these sequences match. In conclusion, the novel splice site does exist in okapi, and okapi MDC1 mRNA is missing 750 nucleotides. Since there are three nucleotides per codon, it can be assumed that the okapi MDC1 protein is missing 250 amino acids.

Although the okapi is missing 750 nucleotides in MDC1 mRNA, it was still unknown whether this same deletion had occurred in giraffe MDC1. Therefore, I amplified fetal giraffe heart, fetal giraffe liver, adult giraffe heart, and adult giraffe liver cDNA using primer pairs 8, 10, and 13. After amplification through PCR, I conducted gel electrophoresis. The results of this gel electrophoresis are found in Table 9.

Table 8: MDC1 Gel Electrophoresis Results of Giraffe Only Amplification

Primer Pair	1: Fetal Gir. Heart	2: Fetal Gir. Liver	3: Adult Gir. Heart	4: Adult Gir. Liver
8	Absent	Absent	Absent	Absent
10	2 Present	2 Present	Absent	Absent
13	Absent	Absent	Present	Present

Unsurprisingly, only a few of the cDNA amplifications yielded any bands. The bands that did appear were removed from the gel and amplified again using PCR. The results can be found in Table 10. The purpose of this was to confirm that the amplification was actually of

MDC1 and to produce more MDC1 replicates to obtain a concentration high enough for sequencing.

Table 9: MDC1 Gel Electrophoresis Results of Giraffe Second Amplification

Primer Pair	Sample & Number	Band Present?
10	1: Fetal Gir. Heart -Longer	Absent
10	1: Fetal Gir. Heart -Shorter	Absent
10	2: Fetal Gir. Liver -Longer	Absent
10	2: Fetal Gir. Liver -Shorter	Absent
13	3: Adult Giraffe Heart	Present
13	4: Adult Giraffe Liver	Present

Due to the lack of quality amplifications using primers 8, 10, and 13, new primer pairs 7 and 9 were synthesized and utilized. Primers 8, 10, and 13 targeted a large portion of MDC1. Therefore, it can be speculated that because primers 8, 10, and 13 targeted a large portion of MDC1, the amplifications were weak and the sequences were especially fragile. Therefore, we developed primers 7 and 9 to specifically target a small region of the MDC1 transcript around the truncated region of exon 5. These primer sequences can be found in the Materials and Methods section. The primer pairs were used to amplify Fetal Giraffe Heart, Fetal Giraffe Liver, Adult Giraffe Heart, and Adult Giraffe Liver. In addition, primers 7 and 9 were used to amplify the specific region of exon 5 of MDC1 using the Adult Giraffe Heart and Adult Giraffe Liver samples that had been amplified previously using primer pair 13. The results are shown in Table 11.

Table 10: MDC1 Gel Electrophoresis Results of Giraffe Only Using Primers 7 & 9

Primer Pair	3: Adult Gir. Heart Amplif. By PP 13	4: Adult Gir. Liver Amplif. By PP 13	1: Fetal Gir. Heart	2: Fetal Gir. Liver	3: Adult Gir. Heart	4: Adult Gir. Liver
7	Present	Present	Absent	Absent	Absent	Absent
9	Present	Present	Absent	Absent	Absent	Absent

Due to successful amplification, adult giraffe heart and liver initially amplified by primer pair 13 and then amplified by primer pairs 7 and 9 were submitted for sequencing to the Penn State DNA Sequencing Ordering and Sequence Retrieval System. The raw sequence data returned to the Cavener group is presented in Figure 11.

Figure 11: Giraffe MDC1 Fragments Utilizing Primer Pairs 7 and 9

>Adult Giraffe Heart – MDC1 Fragment – Primer 7 Reverse
GGCTGGGAGCTGCAACAGGAAGGGCCAGGCTCTTGGGGCAGAGGAAGTAGGAAGGCCTGGGTGGGCTCATCCTCCAT
GCTGGGGACTGCTTCCAGGCTCTGATTCTCTCTGGCTACAAAGCACTGGGTAGCTGGTAGGTCCAGATCTTCAGAAT
CAGCGTGGCTCTCTCTTAGATGTGCCAGTGTCAATGCTGCCGCGACTTCATCTGTATTGCTGTCTATCACCACGGAG
GCTTGGCTTCTGTCTGGGGGGGCCTCCCCCTCCTCTATATCTGTATCACCATCAGCTGGGCTCTCCTGCCGACGTGC
CACACCAGGTGCCTGGGGGCTCTGTGTACCAGCTTCG

>Adult Giraffe Liver – MDC1 Fragment – Primer 7 Forward
GTCGGCAGGAGAGCCCAGCTGATGGTGATACAGATATAGAGGAGGGGGAGGCCCCCCCAGACAGAAGCCAAGCCTCC
GTGGTGATAGACAGCAATACAGATGAAGTCGCGGCAGCATTGACACTGGCACATCTAAGAGAGAGCCACGCTGATTC
TGAAGATCTGGACCTACCAGCTACCCAGTGCTTTGTAGCCAGAGAGAATCAGAGCCTGGAAGCAGTCCCCAGCATGG
AGGATGAGCCCACCCAGGCCTTCCTACTTCCTCTGCCCCAAGAGCCTGGCCCTTCCTGTTGCAGCTTCCCAGCCACA
GGTTCCCTGGATGAGGCATGGGAGGTCTTGTCGACACAG

>Adult Giraffe Liver – MDC1 Fragment – Primer 7 Reverse
GCTGGGAGCTGCAACAGGAAGGGCCAGGCTCTTGGGGCAGAGGAAGTAGGAAGGCCTGGGTGGGCTCATCCTCCATG
CTGGGGACTGCTTCCAGGCTCTGATTCTCTCTGGCTACAAAGCACTGGGTAGCTGGTAGGTCCAGATCTTCAGAATC
AGCGTGGCTCTCTCTTAGATGTGCCAGTGTCAATGCTGCCGCGACTTCATCTGTATTGCTGTCTATCACCACGGAGG
CTTGGCTTCTGTCTGGGGGGGCCTCCCCCTCCTCTATATCTGTATCACCATCAGCTGGGCTCTCCTGCCGACGTGCC
ACACCAGGTGCCTGGGGGCTCTGTGTACCAGCTTCG

>Adult Giraffe Heart – MDC1 Fragment – Primer 9 Forward
CCCAGACAGAAGCCAAGCCTCCGTGGTGATAGACAGCAATACAGATGAAGTCGCGGCAGCATTGACACTGGCACATC
TAAGAGAGAGCCACGCTGATTCTGAAGATCTGGACCTACCAGCTACCCAGTGCTTTGTAGCCAGAGAGAATCAGAGC
CTGGAAGCAGTCCCCAGCATGGAGGATGAGCCCACCCAGGCCTTCCTACTTCCTCTGCCCCAAGAGCCTGGCCCTTC
CTGTTGCAGCTTCCCAGCCACAGGTTCCCTGGATGAGGCATGGGAGGTCTTGGCGACACAGCCATTCTGTCCGA

```
>Adult Giraffe Heart – MDC1 Fragment – Primer 9 Reverse
TGTGGCTGGGAAGCTGCAACAGGAAGGGCCAGGCTCTTGGGGCAGAGGAAGTAGGAAGGCCTGGGTGGGCTCATCCT
CCATGCTGGGGACTGCTTCCAGGCTCTGATTCTCTCTGGCTACAAAGCACTGGGTAGCTGGTAGGTCCAGATCTTCA
GAATCAGCGTGGCTCTCTCTTAGATGTGCCAGTGTCAATGCTGCCGCGACTTCATCTGTATTGCTGTCTATCACCAC
GGAGGCTTGGCTTCTGTCTGGGGGGGCCTCCCCCTCCTCTATATCTGTATCACCATCAGCTGGACTCTCC

>Adult Giraffe Liver – MDC1 Fragment – Primer 9 Forward
CCCCAGACAGAAGCCAAGCCTCCGTGGTGATAGACAGCAATACAGATGAAGTCGCGGCAGCATTGACACTGGCACAT
CTAAGAGAGAGCCACGCTGATTCTGAAGATCTGGACCTACCAGCTACCCAGTGCTTTGTAGCCAGAGAGAATCAGAG
CCTGGAAGCAGTCCCCAGCATGGAGGATGAGCCCACCCAGGCCTTCCTACTTCCTCTGCCCCAAGAGCCTGGCCCTT
CCTGTTGCAGCTTCCCAGCCACAGGTTCCCTGGATGAGGCATGGGAGGTCTTGGCGACACAGCCATTCTGTCC

>Adult Giraffe Liver – MDC1 Fragment – Primer 9 Reverse
CTGTGGCTGGGAAGCTGCAACAGGAAGGGCCAGGCTCTTGGGGCAGAGGAAGTAGGAAGGCCTGGGTGGGCTCATCC
TCCATGCTGGGGACTGCTTCCAGGCTCTGATTCTCTCTGGCTACAAAGCACTGGGTAGCTGGTAGGTCCAGATCTTC
AGAATCAGCGTGGCTCTCTCTTAGATGTGCCAGTGTCAATGCTGCCGCGACTTCATCTGTATTGCTGTCTATCACCA
CGGAGGCTTGGCTTCTGTCTGGGGGGGCCTCCCCCTCCTCTATATCTGTATCACCATCAGCTGGACTCTCCT
```

Through the utilization of the giraffe MDC1 model with a portion of exon 5 missing, an alignment was constructed using BLAST with each of these seven cDNA sequences. The successful alignment of the sequences would indicate that each of the seven sequences is MDC1 with a truncated exon 5. Figure 12 shows this alignment.

Figure 12: Comparison of Giraffe CDS to Model of Giraffe MDC1 with Truncated Exon 5

Giraffe MDC1 Model with Truncated Exon 5:	A
Adult Giraffe Heart – MDC1 Fragment – Primer 7 Reverse:	C
Adult Giraffe Liver – MDC1 Fragment – Primer 7 Forward:	D
Adult Giraffe Liver – MDC1 Fragment – Primer 7 Reverse:	E
Adult Giraffe Heart – MDC1 Fragment – Primer 9 Forward:	F
Adult Giraffe Heart – MDC1 Fragment – Primer 9 Reverse:	G
Adult Giraffe Liver – MDC1 Fragment – Primer 9 Forward:	H
Adult Giraffe Liver – MDC1 Fragment – Primer 9 Reverse:	I

```
A 447  CGAAGCTGGTACACAGAGCCCCCAGGCACCTGGTGTGGCACGTCAGCAGGAGAGTCCAGC  506
C 357  ..............................................G.........C.....  298
D 18                                              ...G.........C.....  36
E 358  ..............................................G.........C.....  299
G 323                                                   ............  312
I 326                                                   ............  314

A 507  TGATGGTGATACAGATATAGAGGAAGGGGAGGcccccccGGACAGAAGCCAAGCCTCCGT  566
C 297  ........................G..............A.....................  238
D 37   ........................G..............A.....................  96
E 298  ........................G..............A.....................  239
```

```
F 21                                          ...A...................  44
G 311  .......................G.............A.....................  252
H 17                                        ....A...................  41
I 313  .......................G.............A.....................  254

A 567  GGTGATGGACAGCAATACAGACGAAGTCGCGGCAGCATTGACACTGGCACATCTAAGAGA  626
C 237  ......A.............T.....................................  178
D 97   ......A.............T.....................................  156
E 238  ......A.............T.....................................  179
F 45   ......A.............T.....................................  104
G 251  ......A.............T.....................................  192
H 42   ......A.............T.....................................  101
I 253  ......A.............T.....................................  194

A 627  GAGCCACGCTGATTTTGAAGATCTGGACCTACCAGCTACCCAGTGCTTTGTAGACAGAGA  686
C 177  ..............C.....................................C......  118
D 157  ..............C.....................................C......  216
E 178  ..............C.....................................C......  119
F 105  ..............C.....................................C......  164
G 191  ..............C.....................................C......  132
H 102  ..............C.....................................C......  161
I 193  ..............C.....................................C......  134

A 687  GAATCAGAGCCTGGAAGCAGTCCCCAGCGTGGAGGATGAGCCCACCCAGGCCTTCCTACT  746
C 117  ............................A...............................  58
D 217  ............................A...............................  276
E 118  ............................A...............................  59
F 165  ............................A...............................  224
G 131  ............................A...............................  72
H 162  ............................A...............................  221
I 133  ............................A...............................  74

A 747  TCCTCTGCCCCAAGAGCCTGGCCCTTCCTGTTGCAGCTTCCCAGCCACAGGTTCCCTGGA  806
C 57   ....................................-.......                  13
D 277  ...........................................................  336
E 58   ....................................-......                  15
F 225  ...........................................................  284
G 71   ................................................             23
H 222  ...........................................................  281
I 73   ................................................            24

A 807  TGAGGCATGGGAGGTCTTGTCGACACAGCCATTCTGTCCGA  847
D 337  ..........................                  364
F 285  ...................G.....................  325
H 282  ...................G...................    320
```

As seen in Figure 12, each of the seven sequences shows a perfect alignment. This indicates that the giraffe, like the okapi, has an MDC1 protein with a truncated exon 5. The giraffe is therefore missing 750 nucleotides in its mRNA and 250 amino acids in MDC1. This demonstrates that the previously shown model in Figure 8 of the giraffe exons and introns is

accurate. This significant alteration of the MDC1 structure could potentially have profound impacts on its function and its ability to interact with closely associated proteins like H2AX.

DISCUSSION

Giraffe DSB Repair Pathway Impact on Karyotype

The removal of 250 amino acids from the structure of MDC1 could have important implications on the proper functioning of the DNA double-strand break repair pathway. MDC1 (mediator of DNA damage checkpoint 1) promotes the recruitment of repair proteins to the sites of DNA double-strand breaks (Stewart et al., 2003). It functions in the DSB Repair pathway as a scaffold protein and is important for protein recruitment and cell signaling. With 250 amino acids missing from a protein of approximately 1800 amino acids, there are going to be changes in its tertiary structure. One can speculate that this change of structure influences other proteins' ability to bind to MDC1. Whether proteins can bind more closely or less closely to MDC1in giraffe is unknown. I speculate that it reduces a protein's ability to bind to MDC1 because the protein has a smaller surface area to bind on. Certain proteins have changed in accordance with MDC1. Through the whole genome analysis of the giraffe, it has been discovered that other proteins important to DSB repair like Nibrin, MRE11, and SOSB2 have also had amino acid changes in giraffe. These proteins interact with MDC1 and therefore it is likely that their unique substitutions were selected for to adapt to the altered structure of MDC1.

Closely linked to double-strand breaks is the process of chromosomal translocations. These translocations can result in a reduced number of chromosomes which are displayed in karyotypes. In particular, the giraffe exhibits an exceptionally small number of chromosomes in its karyotype. As seen in the karyotype in Figure 15, the giraffe has 15 pairs of chromosomes.

This may not seem obscure until one considers closely related pecoran species like the pronghorn. The pronghorn (*Antilocapra americana*) has a total of 29 pairs of chromosomes (Cernohorska et al., 2013) and is thought to be very similar to the pecoran ancestral state in terms of chromosome number.

Figure 13: Giraffe Karyotype (Huang et al., 2008) (Kulemzina et al., 2009)

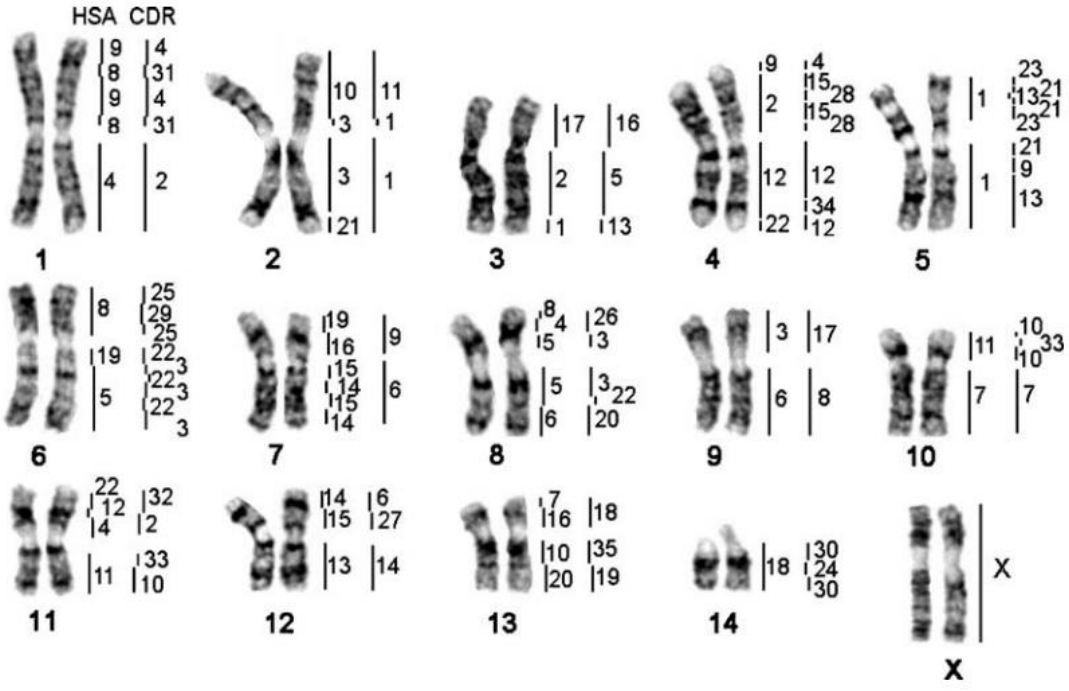


Fig. 3 G-banded karyotype of the giraffe (GCA) with homologies to human (HSA) (based on previously published map of Huang et al. 2008) and dromedary (CDR) painting probes

Unlike the pecoran ancestral state which consists of many acrocentric chromosomes, the giraffe exhibits metacentric chromosomes. This indicates that ancestral giraffe acrocentric chromosomes have repositioned themselves to form metacentric chromosomes over the giraffe's evolutionary history, therefore reducing its chromosome number. This reduction of chromosomes in the giraffe is the result of Robersonian translocations. Robertsonian translocation is a process initiated by a DNA double-strand break event. In this translocation,

arms of acrocentric chromosomes are cleaved from the chromosome's centromere and are bound to cleaved arms of other non-homologous acrocentric chromosomes. This creates new chromosomes that consist of two non-homologous chromosomes. The remaining two small arms that have not been translocated are, often times, redundant in the genome and therefore the loss of this genetic material is insignificant.

A correlation between MDC1 and Robertsonian translocations does exist. When the MDC1 protein was knocked down in B and T cells of mice, 49 out of 240 cells observed had chromosome and chromatid breaks (Franco et al., 2006). Ultimately 20.4% had aberrations while only 1 of 132 cells had aberrations in MDC1 wild type mice cells (Franco et al., 2006). Based on this study, I speculate that the altered MDC1 contributes to the high amount of Robertsonian translocations in giraffe. Due to MDC1's lack of 250 amino acids, proteins that would normally bind can no longer do so. I hypothesize that a protein involved in DNA double-strand break repair that originally interacted in this location on the protein can no longer do so, resulting in sticky ends of chromosomes binding together and forming metacentric chromosomes. Over a period of thousands of years, translocation events occurred gradually, thereby altering the chromosomes of the giraffe to how they are today.

Further research needs to be conducted to decide whether or not the MDC1 exon 5 truncation is the reason behind the unique karyotype. With CRISPR/Cas9, scientists can test whether or not the MDC1 is the responsible for Robertsonian translocations. CRISPR/Cas9 can be used for targeted site specific mutagenesis of MDC1 in the cow. By truncating the portion of exon 5 that is truncated in giraffe, we can test whether cow acrocentric chromosomes form metacentric chromosomes. The cow is a good model because of how closely related it is to the giraffe and because 58/60 chromosomes are acrocentric. Through several subsequent

generations of cow cells, one can observe whether chromosomal fusions have occurred through karyotyping.

Ultimately, giraffe, with a reduced number of chromosomes, had a selective advantage which allowed them to sire more offspring. I speculate that a reduced number of chromosomes could lead to fewer nondisjunction events during meiosis, and ultimately more capable offspring. However, many other animals have not evolved to form metacentric chromosomes. Therefore, it is a possibility that the small arms of acrocentric chromosomes contain important genes that are required for cell growth and functioning.

Okapi DSB Repair Pathway Impact on Karyotype

The giraffe's closest living relative, the okapi, is morphologically much different from the giraffe; however, it shares very similar genes. Like the giraffe, the okapi has an altered MDC1 protein with a 250 amino acid deletion. Based on my earlier conclusions, one might assume that okapi would therefore have a similar karyotype to the giraffe; however, the karyotype of the okapi is very different from than that of the giraffe. Instead of 15 pairs of chromosomes, the okapi has 22 pairs of chromosomes. This makes the okapi's chromosome count halfway between the giraffe and the ancestral pecoran state. Figure 16 properly displays the karyotype of the okapi. In the okapi, Robertsonian translocations have occurred, but not to the degree in which they have occurred within the giraffe.

Figure 14: Okapi Karyotype (Kulemzina et al., 2009)

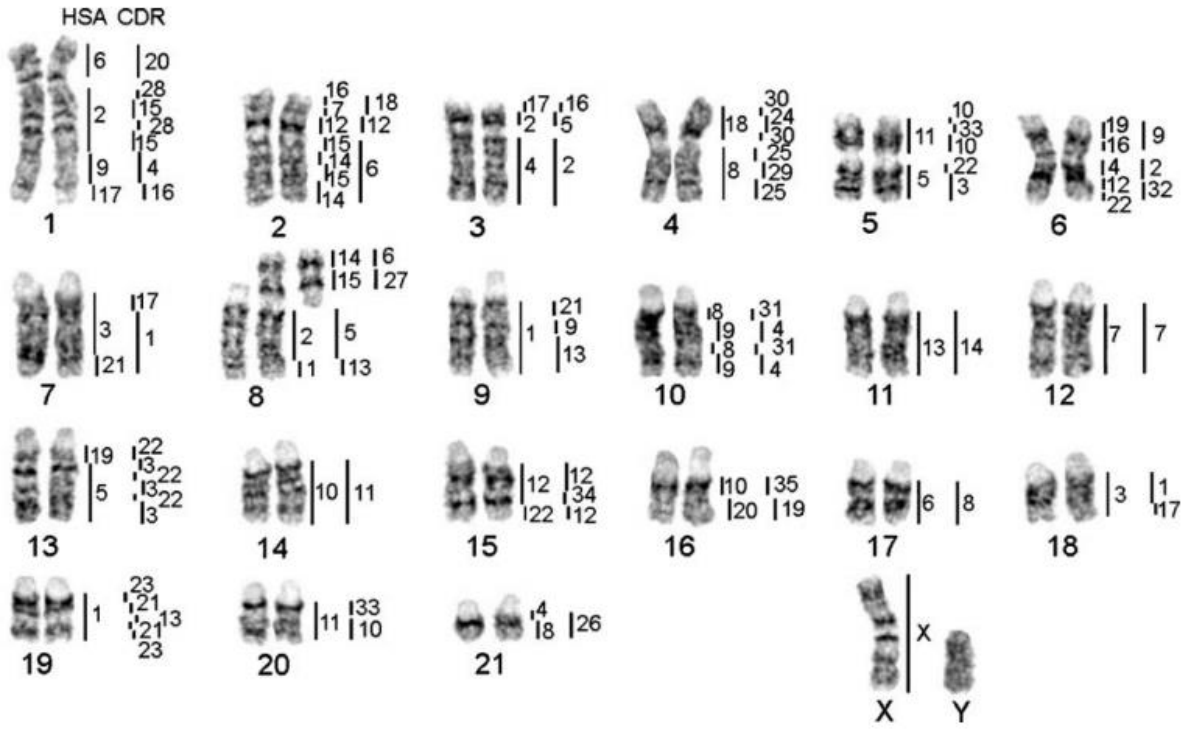
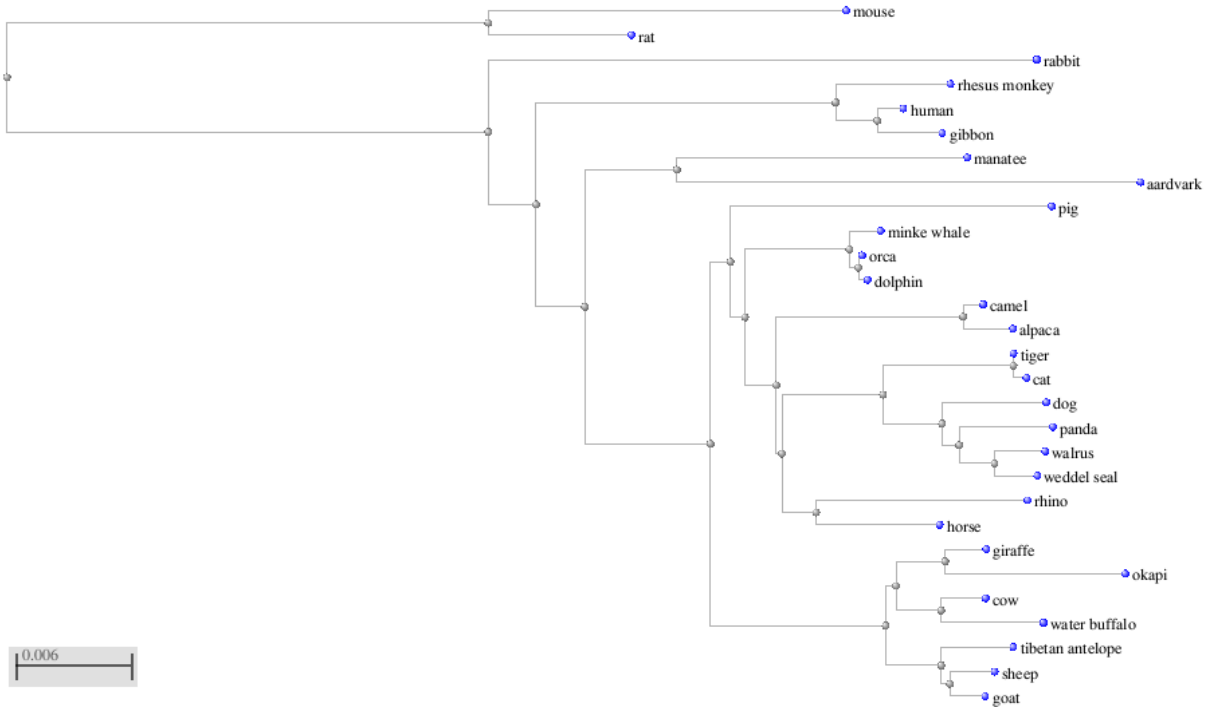


Fig. 4 G-banded karyotype of the okapi (OJO) with homologies to human (HSA) and dromedary (CDR) painting probes

Since the okapi karyotype is so much different than the giraffe karyotype, it is necessary to distinguish other players in the DNA double-strand break repair pathway that are unique to the okapi but not significantly changed in giraffe. Okapi has two additional proteins that are highly altered compared to their most recent common ancestors. These are ATR and BAZ1B. ATR is a serine/threonine kinase which activates checkpoint signaling during stress on the genome such as ionizing radiation or ultraviolet light or DNA replication stalling (uniprot.org). This protein is responsible for phosphorylating BRCA1, RAD17, and p53 among other proteins in the DSB repair process (uniprot.org). According to our results from sequencing the okapi genome, the okapi is the most highly diverged mammal in terms of this protein.

BAZ1B is a tyrosine protein kinase that phosphorylates at Tyr 142 of H2AX (uniprot.org). It plays a major role in chromatin remodeling and acts as a transcription regulator. It determines whether DNA repair enzymes should be recruited or apoptosis should occur (uniprot.org). Like ATR, BAZ1B is highly diverged in okapi. There are 12 additional non-synonymous substitutions when comparing okapi to cow than when comparing giraffe to cow. Using genetic information from okapi BAZ1B and BAZ1B from other mammals, this phylogenetic tree was created by the Cavener group, showing the protein's genetic divergence.

Figure 15: Phylogenetic Tree for BAZ1B Divergence



I speculate that genetically diverged BAZ1B and ATR in okapi may have influenced the okapi's ability to properly carry out DNA double-strand break repair, just as the diverged MDC1 has in the species. Okapi has been exposed to Robertsonian translocations throughout its evolution; however, the okapi has more chromosomes than so it can be assumed that translocation has not occurred as frequently as it has in the giraffe. I

speculate that BAZ1B and ATR have accumulated these non-synonymous substitutions over its evolution to more efficiently bind to the MDC1 scaffolding protein. By changing its amino acid composition, it can more closely fit along MDC1 to carry out its function. This could have recovered some of the functions of MDC1 that were lost in the giraffe.

Potential Cancer Implications

Altering the structure of MDC1 through the removal of 250 amino acids could have major implications on tumor growth. When MDC1 is knocked down in cells, they are hypersensitive to ionizing radiation and fail to activate the intra-S phase and G2/M phase cell-cycle checkpoints (Stewart et al., 2003), which can lead to uncontrollable cell division within the cell.

The most important segments of MDC1 are its two conserved domains. One is the Forkhead associated domain (FHA) and the other is the Breast Cancer Suppressor Protein (BRCA1) carboxy-terminal domain (BRCT) (ncbi.org). The FHA domain is critical in the cell for establishing and maintaining cell cycle checkpoints and DNA repair, two of the main roles of MDC1. The BRCT domain is also found in DNA damage repair and cell cycle checkpoint proteins. In addition, the BRCT domain forms homo/hetero multimers and interacts within DNA double-strand breaks (ncbi.org). The BRCT domain binds to the phosphorylated tail of H2AX where the COOH-terminal Tyr residue lies (Lee et al., 2005). The FHA and BRCT domains are located towards the ends of the MDC1 protein. The FHA domain is located within the first 200 amino acids, while the BRCT domain is located within the last 300 amino acids (exact locations depend on species). In giraffe MDC1, the 250 amino acid deletion occurs between amino acids

324 and amino acid 674 in relation to the cow MDC1. Although this area does not immediately affect each of these two domains, I speculate that the sheer size of the deletion changes the tertiary structure of the protein and the proteins overall ability to interact through its FHA and BRCT domains.

Changing the structure of a highly conserved protein like MDC1 heavily impacts the cell's ability to recruit the proper DNA double-strand break repair proteins. Data suggests that MDC1 is critical to tumor suppression in humans (Bartkova et al., 2007). When MDC1 is knocked down in mice, tumor incidence increases significantly (Minter-Dykhouse et al., 2008). Interestingly, no mutations of MDC1 have been found in human cancer samples (Minter-Dykhouse et al., 2008). I speculate that this is because of its necessity in maintaining DNA stability. The Giraffe and okapi deletion of 250 amino acids could potentially yield interesting scientific cancer studies and help to assess whether the giraffe and okapi version of MDC1 is better or worse at reducing tumor growth. Knowing that such a massive change in the highly conserved MDC1 exists could allow scientists to test its scaffolding functions for BRCA1 and H2AX.

CONCLUSION

In conclusion, MDC1 is missing 250 amino acids in both giraffe and okapi. In exon 5, a large portion has become intronic sequences and is no longer encoded into a protein. In addition to this change in MDC1, BAZ1B and ATR is highly altered in the okapi. Due to the how closely these proteins interact with each other in DNA double strand break repair pathway, it can be speculated that the giraffe and okapi display a unique DSB repair system. Ultimately, the alterations of MDC1 could have influenced the evolution the giraffe and okapi karyotypes and resulted in Robertsonian translocations. Researchers in the near future will have to carry out a variety of experiments to determine how exactly this MDC1 variant operates. The genome of the giraffe could reveal some interesting information that could shine light on their unique physiology as well as the physiology of other vertebrates.

APPENDIX

Coding Sequence Model of Giraffe MDC1 based on *Bos Taurus*

Note that the region in green is proposed truncated portion of MDC1.

ATGGAAGACACCCAGGTTTTAAACTGGGAAGTTGAAGAAGAGGAGGTGGTTGAAGAGAGC
CCCAGTGAATCTTTGGGGTATAGCTTGGAGCCCCTAGGGCAGCTGCATGTCTTCAGTAGT
TCCTATGGACCAGAAAAAGACTTCCCACTCTACCTCGGGAAGAACATGATAGGCCGAATG
CCTGATTGCTCTGTGGTCCTGCCCTTTTCATCCATCTCCAAACAACATGCAGTGATTGAA
ATTTCTGCCTGGGACAAGGCACCTGTCCTCCAGGATTGCGGGAGCCTCAATGGCACTCAA
ATCCTGCGGCCTCCTAAGGTCCTGGGCCCTGGGGTGAGTCATCGTTTGCGGGACCAGGAG
TTGATTCTCTTTGCTGACTTGCCCTGCCAGTACCATCGCTTGGATGTCCCCCGGCCTTTT
GTCTCCCGGGGCCGCCCTCTAACTGTAGAGGAGACACCCAGGGTACAGGGAGGAACTCAA
CCCCCCAGGCTTCTGTTGGCTGAGGACTCAGAGGAAGAAGTAGATTCCTTTTTGGACAAG
TGTGTGGTGAAAGGACCAAGGACCTCCTCTTTGGCAACAGTCGTTCCAGAGAGTGATGAA
GAAGGGCCTTCCCCTGCCCCGGATGGGCCTAGGCCACCTTCTGCCTTCAACTTGAACAGT
GACACGGATGAGGAAGAAAGTCAGGAATCAGGAGCTGCCAGAAGAGGTACTGCTGCAGAG
ACAGAACAGCCTAAACCTGTCATGGCCGAAATCCAGATTGAGAAGGATCAGTGTTCTGTG
AAAGAGAAGAACAGAGACACAGAAATCGAGAGGGATGCAAGGAATGGGGTGGTTCCAACT
GGAGTGATTCTGAAGAGGAGCCAGCCTTCTGGGGAGGACAGTGACACAGATGTGAGTGAT
GAGAGTGGGCCTCCAAGAAGGCTTGCTGGGGTCCCTCCGAAAAGGGCCTGGTCTTGTAAC
TTCATAGGCAGTGATACCAATGGGAAAGACGAAGGGATCCCTGCTACCCCAGCAGTGGTT
CCCGTGAAGGAGAGGCAGATTTTCCACGAAGCTGGTACACAGAGCCCCCAGGCACCTGGT
GTGGCACGTCAGCAGGAGAGTCCAGCTGATGGTGATACAGATATAGAGGAAGGGGAGGCC
CCCCCGGACAGAAGCCAAGCCTCCGTGGTGATGGACAGCAATACAGACGAAGTCGCGGCA
GCATTGACACTGGCACATCTAAGAGAGAGCCACGCTG
GTAAGTGGAACCGAGATCCAGATGCAGAAGAGGACAGGGCCCAACTAGTGGCCCTTCTGC
GAAGCCAGGCCTCTGCTAGGGGAGACAGTGACACAGATGTGGAGGAAGAGGGGCTCCCA
GTGGAAAGGAAGGGAGTGGTTCCCAAGGGTCACATGGACAGGGAATATTCAAAAAAGAGC
CAGCATCCTCCCAGGGACAGTGATACAGAGGGGAAAGAAGATGAGAGCTCACCGGGGGTC
CACCTGGAGAGAAGCCAGGCCTCTGCACAAGTGGAGGACGAGGTCCCACTGGGGCCAGCT
GTTGCACTTCCATAGAAGCGTCAGGTACGAGGCATGGCGTGGACACATCACACCAATGCA
GAGGCAGAAGGGGGCCCGTCACAGCTCCCCATGCTGCATCTAGAGGAAGCCCGACCTCCT
CTAGCTGGGGATTGTGAACCGGATGCGGAGAACACATCCTTAGCAGTGGCCGGTGTCAGA
AAGAGCCAGCTTCCAGCGGAAAAAGATGCTGGGACCACGTGGGCCGTGGCTGTTCCTGAA
CAGGACAGATTTGCCACCGGGATCCAGGGTGGGTCATTCATGGCACCAGGGGAGCAAGAC
CTTCTCCCAGTCTCAAGGGAAAACCTAGCAGATCTGGTGGTGGACACAGGCACTCCAGGG
GAGCCCCAACCGCAGAAAGAGGGGGCCCAGACCACCACAGGAAGGGAGAGAGAACCACAT
GGGGATAGGGCCACAGACTCTCGAGAGAACCTCCATGATTTTGAAGATCTGGACCTACCA
GCTACCCAGTGCTTTGTAGACAGAGAGAATCAGAGCCTGGAAGCAGTCCCCAGCGTGGAG

GATGAGCCCACCCAGGCCTTCCTACTTCCTCTGCCCCAAGAGCCTGGCCCTTCCTGTTGC
AGCTTCCCAGCCACAGGTTCCCTGGATGAGGCATGGGAGGTCTTGTCGACACAGCCATTC
TGTCCGAGAGAGTCTGAGGCCTCCGAGACCCAAACTGCTGTCACCCGCCTTGACACCTGT
GCATCTTGCCCTCCTCCATCTAGGACAGCACAGCAAGAGCAACATCCAGAGAGCCCAGTC
CACGCAGAGCCACTGGGGATGGAAGGCAGAGGGATGCAGACTGTGGAGAAAGACATGGGT
ACCCCAAGAGAAACAGCAGAGAGGGTGATACCTGAGGGAGGGCCACTGCAGAAGGAGACC
AAGAAACTGCCCTCGGAAGGAGAGAGGGAAGATGTGACGGGAGAGGAAGAATTAATCCGG
GGGATACAGGACAGAGAACAAAATCAGGTGTTAGCTGGAGATACTCAGAGCCAACAATCT
GACAAAAAAAATGAAAAGTGCAAGTACTGGAAGGGGAATGGAGATTGTGAAGGTAGAAATT
GAGACACCCAAGGAAACACAGGAGAAAGAGAGAGAAAAGCAGACTCTTGCAGGGGAAATA
TTTGAGAGTGAAGCAGGGAAACTGGTAGCAGAGAGAGAGAATGAGGCAGGTGGTTTAGAA
GTCAAGACAGAGGCGGGGGGCCAGGACCAGAAAAGCCAAGCCTCTGGTCCGCCACCAGAG
CTTGGAGCAGGGGCGGGAGACCTTCAGGGACTTGCTTCAGACCCAGTAGCTTCTGGGAGA
CAGGCAGGTGGAGGAAGGGGAGCCCCAGGGAGCCCCAGGAGGCAGCAGAGAGGTGACTTG
AATTGCAAGATGCCACCTGCTGAGAAGGCTTCCAGGGTGAGAACTGCTGTCTCTACCCCC
ACCACTCCACAGGGTGGTCAGGAATCCTCAGATGCTTGTCTGCCTCCTGCAGCGCCTGAAGCC
TCAGCCGCACTCCCAAACTCCCTCATCTCTCAGATCCAAAAACATCCCACACCTCAGTCC
CTCCTTTTTCCCTCTCCAGCTCCTTTAGAACTGCCCATTCCCAGGACCAGACAAAATGAG
AGTCAGGAAGCTCCAGAGACTCCCTTCTTCTCAGAGCTGAACTCTGTCCACCCAAAACCC
AAAGTCAGGCCCCAGGGGTCCTCTCCAATTTCTTCTCTACCCCTTGAGCCCCACCCTACC
ACCCCCACAGGCCAGCATATTGCCCCTGAACCCACATCTGGGGTCTCTCGGAGCAGGACA
CATAGTTCCTTCGACGTAACTGCCTCATTCATTGTCCCCACAGCCCTTGCGTTGCAGCCA
TCCACCTCCACAGACCAGCCTATCACCCCTAAGCCCAGACTTCGGGCCCCTCGGGGCAGG
GCACACAGGTCTTCTGTCAAAACCCCTGAACCCAGTGTCTCCACAGACCAGCCTGTTGCC
CCTGAGCTCACAGCTAAGGCCACTCGGGGCAGGGCACAGAGGTCTTCTGTCAAGACTCCC
AAACCAGATAACCCCACAACACTCAAGCCCCAGCCTTCCACTTCCACAGACCAGCCTGTC
ACCCCCAAACCCACATCTCTGGCCCCTCGGGGCAGGACACCTAGGTCTTCTGTCAAGACT
CCTGAGCCAGTTGTCCCCACAGCCTCTGAGCTCCAGCCTGCTGCCCCTAAAGACCAGCCT
CCTGTTGCTCCTGAGCTCACATCTAGAGCCACTCGAGGCAGGACACAGAGGTCTTCTGTC
AAGACTTCCAAACCAGATATGTCCACAACCCCCGAGCCCCAGCCTTCCACTTCCACAGAC
CAGCCTGTCACCCCGAAACCCACATCTCGGGCCCCTCGGGGCAGGACACCTAGGTCTTCC
ATCAAGACTCCTGAACCAGTTGTCTCCACAGCCTCTGAGCTCCAGCCTTCTGCCCTCACA
GACCAGCTTGTCACTCCTGAGCTCACATCTAGGGCTACTCGGGGCAGGGCACAGAGGTCC
TCTGTCAAAACCTCTGATCCAGTTACCACCACAACACCTGAGCTCCAGCCTTCCACCTCC
ACAGACCAGCTTGTTACCCCCAAACCCACATCTCGGGCCCCTAGAGGCAGGACACTTAGG
TCGTCTGCCAAGACTCCTGAATCAGTTGTTCCCACAGCCTCTGAGCTCCAGCCTTCTGCC
CCTGCAAACCAGCCTGTTGGTCCTTGGGCCACTCAGTGTAGAAGACGTAGGTCTTCTGTC
AAGACCCCAGAACCAGTTGTCCCTGCAGTCCCTGAGCCTCAGCCTTCCACTTCTAAAGAC
CAGTCTGTCGCTCCTGAACCCACATCTCAGGCCACTCAGGGCCAAACACATAAGTCCTCT
GTCAAGACATCCCAGCCAACTGAACCCACAGCCCCTGACCTCAAACCTTCCTCCCCCACA
GACCAGCCTGTCACTCCCAAGGTCATAGCTCAGGGTGGTCCAAGCAGGACACGAAGGTCT
TCTACAGCAAGTGCTGTGCTGGTTCCTACTACCCCTGAATTCCAGCCTCCAGTCCCCTCA
GAACAGCCTCTTCTCCCTGACCCCATCCCCGAAGTCAACTGCAGCAGGAGGCCGAGGGCC
ACTAGGAAACATGGGTCTCCCACAGCTCATGTTCATGAGCCCTACTCCGCACCCCCTGAA
CCTAACTCCCGCTCCTCAAGGAATCAAAGACGAGGGGCAGTGAAAGCAGCGGAGCCCCTT
AGCACCATTCCTGAGCCTGCCTTTGCCCAGCTCCCCGAGGTACCCCCTCACATTCCCCAG

ATCCCAAAGGAGGATGCAGCAGATGGATCAGGCTTCACCCCAGAGCCCCAGCCTAAGGCC
TCTCAAAACCACAAGAGGCCTGCAGCTATTACACATTCACCTCCACTTCAAAAACGGTTC
CAGAGAGGGAAAGTCCCTCAGAAGGCAGCAACCATTAAGGAAGAAGAAGAAAATCCAGCG
GGGAGGCCAAGGAAGGAAGAGGGTGTAGTGATTCCAGGTCCAGGCAAGAGAAAGAGAGAG
CAGACAGAGGAGGAGTCCCAGGGAAGACCAAGCCGCAGCCTGCGACGGACCAAACTTATC
CAGGAGTCCACGGCCCCCAAAGTGCTCTTCACGGGCGTGGTAGATGCTCGCGGAGAGAGG
ACGGTGCTGACCCTGGGGGGCAGCCTGGCCAGCTCAGTGGCTGAAGCTTCTCACCTCGTG
ACTGATCGGATCCGCCGGACGGTCAAGTTTCTGTGTGCCCTGGGCCGGGGCATCCCCATC
CTCTCCCTGGCCTGGCTGCACGAGTCCCGCAAGGCAGGCTGCTTCTTGCCCCCGGACGAA
TATTTGGTGACTGATCCTGAGCAGGAGAAGAACTTCGGCTTCAGCCTTCGGGAGGCCCTG
AGCCGAGCTCAGGAGCGAAGGCTGCTGGAGGGCTATGAGATTTATGTGACACCTGGAGTC
CAGCCACCGCCACCTCAGATGGGAGAAATCATCAACTGCTGTGGAGGCGCCATCCTACCC
AGCATGCCCCGGTCCTACAAGCCTCAGAGGGTCGTGATCACTTGTTCCCAGGACTTCCCT
CGATGTGCCATTCCATATCGGGTTGGGCTGCCCATCCTCTCACCCGAGTTCCTGCTGACG
GGAGTACTGAAGCAGGAAGTCAAGCCAGAGGCCTTTGCCTTCTCCACTGTGGAAATGTCA
TCCACC

Coding Sequence Model of Giraffe MDC1 with Truncated Exon 5

ATGGAAGACACCCAGGTTTTAAACTGGGAAGTTGAAGAAGAGGAGGTGGTTGAAGAGAGC
CCCAGTGAATCTTTGGGGTATAGCTTGGAGCCCCTAGGGCAGCTGCATGTCTTCAGTAGT
TCCTATGGACCAGAAAAAGACTTCCCACTCTACCTCGGGAAGAACATGATAGGCCGAATG
CCTGATTGCTCTGTGGTCCTGCCCTTTTCATCCATCTCCAAACAACATGCAGTGATTGAA
ATTTCTGCCTGGGACAAGGCACCTGTCCTCCAGGATTGCGGGAGCCTCAATGGCACTCAA
ATCCTGCGGCCTCCTAAGGTCCTGGGCCCTGGGGTGAGTCATCGTTTGCGGGACCAGGAG
TTGATTCTCTTTGCTGACTTGCCCTGCCAGTACCATCGCTTGGATGTCCCCCGGCCTTTT
GTCTCCCGGGGCCGCCCTCTAACTGTAGAGGAGACACCCAGGGTACAGGGAGGAACTCAA
CCCCCCAGGCTTCTGTTGGCTGAGGACTCAGAGGAAGAAGTAGATTCCTTTTTGGACAAG
TGTGTGGTGAAAGGACCAAGGACCTCCTCTTTGGCAACAGTCGTTCCAGAGAGTGATGAA
GAAGGGCCTTCCCCTGCCCCGGATGGGCCTAGGCCACCTTCTGCCTTCAACTTGAACAGT
GACACGGATGAGGAAGAAAGTCAGGAATCAGGAGCTGCCAGAAGAGGTACTGCTGCAGAG
ACAGAACAGCCTAAACCTGTCATGGCCGAAATCCAGATTGAGAAGGATCAGTGTTCTGTG
AAAGAGAAGAACAGAGACACAGAAATCGAGAGGGATGCAAGGAATGGGGTGGTTCCAACT
GGAGTGATTCTGAAGAGGAGCCAGCCTTCTGGGGAGGACAGTGACACAGATGTGAGTGAT
GAGAGTGGGCCTCCAAGAAGGCTTGCTGGGGTCCCTCCGAAAAGGGCCTGGTCTTGTAAC
TTCATAGGCAGTGATACCAATGGGAAAGACGAAGGGATCCCTGCTACCCCAGCAGTGGTT
CCCGTGAAGGAGAGGCAGATTTTCCACGAAGCTGGTACACAGAGCCCCCAGGCACCTGGT
GTGGCACGTCAGCAGGAGAGTCCAGCTGATGGTGATACAGATATAGAGGAAGGGGAGGCC
CCCCCGGACAGAAGCCAAGCCTCCGTGGTGATGGACAGCAATACAGACGAAGTCGCGGCA
GCATTGACACTGGCACATCTAAGAGAGAGCCACGCTG
ATTTTGAAGATCTGGACCTACCA
GCTACCCAGTGCTTTGTAGACAGAGAGAATCAGAGCCTGGAAGCAGTCCCCAGCGTGGAG
GATGAGCCCACCCAGGCCTTCCTACTTCCTCTGCCCCAAGAGCCTGGCCCTTCCTGTTGC
AGCTTCCCAGCCACAGGTTCCCTGGATGAGGCATGGGAGGTCTTGTCGACACAGCCATTC
TGTCCGAGAGAGTCTGAGGCCTCCGAGACCCAAACTGCTGTCACCCGCCTTGACACCTGT

GCATCTTGCCCTCCTCCATCTAGGACAGCACAGCAAGAGCAACATCCAGAGAGCCCAGTC
CACGCAGAGCCACTGGGGATGGAAGGCAGAGGGATGCAGACTGTGGAGAAAGACATGGGT
ACCCCAAGAGAAACAGCAGAGAGGGTGATACCTGAGGGAGGGCCACTGCAGAAGGAGACC
AAGAAACTGCCCTCGGAAGGAGAGAGGGAAGATGTGACGGGAGAGGAAGAATTAATCCGG
GGGATACAGGACAGAGAACAAAATCAGGTGTTAGCTGGAGATACTCAGAGCCAACAATCT
GACAAAAAAAATGAAAAGTGCAAGTACTGGAAGGGGAATGGAGATTGTGAAGGTAGAAATT
GAGACACCCAAGGAAACACAGGAGAAAGAGAGAGAAAAGCAGACTCTTGCAGGGGAAATA
TTTGAGAGTGAAGCAGGGAAACTGGTAGCAGAGAGAGAGAATGAGGCAGGTGGTTTAGAA
GTCAAGACAGAGGCGGGGGGCCAGGACCAGAAAAGCCAAGCCTCTGGTCCGCCACCAGAG
CTTGGAGCAGGGGCGGGAGACCTTCAGGGACTTGCTTCAGACCCAGTAGCTTCTGGGAGA
CAGGCAGGTGGAGGAAGGGGAGCCCCAGGGAGCCCCAGGAGGCAGCAGAGAGGTGACTTG
AATTGCAAGATGCCACCTGCTGAGAAGGCTTCCAGGGTGAGAACTGCTGTCTCTACCCCC
ACCACTCCACAGGGTGGTCAGGAATCCTCAGATGCTTGTCTGCCTCCTGCAGCGCCTGAAGCC
TCAGCCGCACTCCCAAACTCCCTCATCTCTCAGATCCAAAAACATCCCACACCTCAGTCC
CTCCTTTTTCCCTCTCCAGCTCCTTTAGAACTGCCCATTCCCAGGACCAGACAAAATGAG
AGTCAGGAAGCTCCAGAGACTCCCTTCTTCTCAGAGCTGAACTCTGTCCACCCAAAACCC
AAAGTCAGGCCCCAGGGGTCCTCTCCAATTTCTTCTCTACCCCTTGAGCCCCACCCTACC
ACCCCCACAGGCCAGCATATTGCCCCTGAACCCACATCTGGGGTCTCTCGGAGCAGGACA
CATAGTTCCTTCGACGTAACTGCCTCATTCATTGTCCCCACAGCCCTTGCGTTGCAGCCA
TCCACCTCCACAGACCAGCCTATCACCCCTAAGCCCAGACTTCGGGCCCCTCGGGGCAGG
GCACACAGGTCTTCTGTCAAAACCCCTGAACCCAGTGTCTCCACAGACCAGCCTGTTGCC
CCTGAGCTCACAGCTAAGGCCACTCGGGGCAGGGCACAGAGGTCTTCTGTCAAGACTCCC
AAACCAGATAACCCCACAACACTCAAGCCCCAGCCTTCCACTTCCACAGACCAGCCTGTC
ACCCCCAAACCCACATCTCTGGCCCCTCGGGGCAGGACACCTAGGTCTTCTGTCAAGACT
CCTGAGCCAGTTGTCCCCACAGCCTCTGAGCTCCAGCCTGCTGCCCCTAAAGACCAGCCT
CCTGTTGCTCCTGAGCTCACATCTAGAGCCACTCGAGGCAGGACACAGAGGTCTTCTGTC
AAGACTTCCAAACCAGATATGTCCACAACCCCCGAGCCCCAGCCTTCCACTTCCACAGAC
CAGCCTGTCACCCCGAAACCCACATCTCGGGCCCCTCGGGGCAGGACACCTAGGTCTTCC
ATCAAGACTCCTGAACCAGTTGTCTCCACAGCCTCTGAGCTCCAGCCTTCTGCCCTCACA
GACCAGCTTGTCACTCCTGAGCTCACATCTAGGGCTACTCGGGGCAGGGCACAGAGGTCC
TCTGTCAAAACCTCTGATCCAGTTACCACCACAACACCTGAGCTCCAGCCTTCCACCTCC
ACAGACCAGCTTGTTACCCCCAAACCCACATCTCGGGCCCCTAGAGGCAGGACACTTAGG
TCGTCTGCCAAGACTCCTGAATCAGTTGTTCCCACAGCCTCTGAGCTCCAGCCTTCTGCC
CCTGCAAACCAGCCTGTTGGTCCTTGGGCCACTCAGTGTAGAAGACGTAGGTCTTCTGTC
AAGACCCCAGAACCAGTTGTCCCTGCAGTCCCTGAGCCTCAGCCTTCCACTTCTAAAGAC
CAGTCTGTCGCTCCTGAACCCACATCTCAGGCCACTCAGGGCCAAACACATAAGTCCTCT
GTCAAGACATCCCAGCCAACTGAACCCACAGCCCCTGACCTCAAACCTTCCTCCCCCACA
GACCAGCCTGTCACTCCCAAGGTCATAGCTCAGGGTGGTCCAAGCAGGACACGAAGGTCT
TCTACAGCAAGTGCTGTGCTGGTTCCTACTACCCCTGAATTCCAGCCTCCAGTCCCCTCA
GAACAGCCTCTTCTCCCTGACCCCATCCCCGAAGTCAACTGCAGCAGGAGGCCGAGGGCC
ACTAGGAAACATGGGTCTCCCACAGCTCATGTTCATGAGCCCTACTCCGCACCCCCTGAA
CCTAACTCCCGCTCCTCAAGGAATCAAAGACGAGGGGCAGTGAAAGCAGCGGAGCCCCTT
AGCACCATTCCTGAGCCTGCCTTTGCCCAGCTCCCCGAGGTACCCCCTCACATTCCCCAG
ATCCCAAAGGAGGATGCAGCAGATGGATCAGGCTTCACCCCAGAGCCCCAGCCTAAGGCC
TCTCAAAACCACAAGAGGCCTGCAGCTATTACACATTCACCTCCACTTCAAAAACGGTTC
CAGAGAGGGAAAGTCCCTCAGAAGGCAGCAACCATTAAGGAAGAAGAAGAAAATCCAGCG

GGGAGGCCAAGGAAGGAAGAGGGTGTAGTGATTCCAGGTCCAGGCAAGAGAAAGAGAGAG
CAGACAGAGGAGGAGTCCCAGGGAAGACCAAGCCGCAGCCTGCGACGGACCAAACTTATC
CAGGAGTCCACGGCCCCCAAAGTGCTCTTCACGGGCGTGGTAGATGCTCGCGGAGAGAGG
ACGGTGCTGACCCTGGGGGGCAGCCTGGCCAGCTCAGTGGCTGAAGCTTCTCACCTCGTG
ACTGATCGGATCCGCCGGACGGTCAAGTTTCTGTGTGCCCTGGGCCGGGGCATCCCCATC
CTCTCCCTGGCCTGGCTGCACGAGTCCCGCAAGGCAGGCTGCTTCTTGCCCCCGGACGAA
TATTTGGTGACTGATCCTGAGCAGGAGAAGAACTTCGGCTTCAGCCTTCGGGAGGCCCTG
AGCCGAGCTCAGGAGCGAAGGCTGCTGGAGGGCTATGAGATTTATGTGACACCTGGAGTC
CAGCCACCGCCACCTCAGATGGGAGAAATCATCAACTGCTGTGGAGGCGCCATCCTACCC
AGCATGCCCCGGTCCTACAAGCCTCAGAGGGTCGTGATCACTTGTTCCCAGGACTTCCCT
CGATGTGCCATTCCATATCGGGTTGGGCTGCCCATCCTCTCACCCGAGTTCCTGCTGACG
GGAGTACTGAAGCAGGAAGTCAAGCCAGAGGCCTTTGCCTTCTCCACTGTGGAAATGTCA
TCCACC

BIBLIOGRAPHY

Badlangana N. L., Adams J. W., Manger P. R. (2009) The giraffe (*Giraffa Camelopardalis*) cervical vertebral column: a heuristic example in understanding evolutionary processes? *Zool. J. Linnean. Soc.* 155, 736-757.

Bartkova J., Horej Sbrever I. Z., Sehested M., Nesland J. M., Rajpert-De Meyts E., Skakkebaek N. E., Stucki M., Jackson S., Lukas J., Bartek J., (2007) DNA damage response mediators MDC1 and 53BP1: constitutive activation and aberrant loss in breast and lung cancer, but not in testicular germ cell tumours. *Oncogene* 26, 7414-7422.

Bernstein N. K., Hammel M., Mani R. S., Weinfeld M., Pelikan M., Tainer J.A., Glover J. N. (2009) Mechanism of DNA substrate recognition by the mammalian DNA repair enzyme, Polynucleotide Kinase. *Nucleic Acids Res.* 37, 6161–6173.

Bohgaki T., Bohgaki M., Hakem R. (2010) DNA double-stand break signaling and human disorders. *Genome Integr.* 1, 15.

Brøndum E., Hasenkam J. M., Secher N. H., Bertelsen M. F., Grøndahl C., Petersen K. K., Buhl R., Aalkjaer C., Baandrup U., Nygaard H., Smerup M., Stegmann F., Sloth E., Ostergaard K. H., Nissen P., Runge M., Pitsillides K., Wang T. (2009) Jugular venous pooling during lowering of the head affects blood pressure of the anesthetized giraffe. *Am. J. Physiol.-Reg. I.* 297, R1058–R1065.

Brown D. M., Brenneman R. A., Koepfli K. P., Pollinger J. P., Mila B., Georgiadis N. J., Louis E. E., Grether G. F., Jacobs D.K., Wayne R. K. (2007) Extensive population genetic structure in the giraffe. *BMC Evol. Biol.* 5, 57.

Cameron E. Z., du Toit J. T. (2007) Winning by a neck: tall giraffes avoid competing with shorter browsers. *Am. Nat.* 169(1), 130-135.

Carey J. R., Judge D. S. (2000) *Longevity records; life spans of mammals, birds, amphibians, reptiles, and fish*. Odense: Odense University Press.

Cernohorska H., Kubickova S., Kopecna O., Kulemzina A. I., Perelman P. L., Elder F. B., Robinson T. J., Graphodatsky A. S., Rubes J. (2013) Molecular cytogenetic insights to the phylogenetic affinities of the giraffe (*Giraffa Camelopardalis*) and pronghorn (*Antilocapra americana*). *Chromosome Res.* 21, 447-460.

Coimbra J. P., Hart N. S., Collin S. P., Manger P. R. (2013) Scene From Above: Retinal Ganglion Cell Topography and Spatial Resolving Power in the Giraffe (*Giraffa Camelopardalis*). *J. Comp. Neurol.* 521, 2042-2057.

EAZWV (2008) Evaluation of okapi (*Okapia johnstoni*) necropsy reports and studbook data as part of the EAZWV Summer School. *Proc EAZWV* 7, 323-327.

Fennessy J. (2008) *An Overview of Giraffe Giraffa Camelopardalis Taxonomy, Distribution and Conservation Status, with a Namibian Comparative and Focus on the Kunene Region*. Namibia Scientific Society.

Franco S., Gostissa M., Zha S., Lombard D. B., Murphy M. M., Zarrin A. A., Yan C., Tepsuporn S., Morales J. C., Adams M. M., Lou Z., Bassing C. H., Manis J. P., Chen J., Carpenter P. B., Alt F. W. (2006) H2AX Prevents DNA Breaks from Progressing to Chromosome Breaks and Translocations. *Mol. Cell* 21, 201-214.

Gijzen, A. (1958) Notice sur la reproduction de l'okapi (*Okapia johnstoni* Sclater) au Jardin Zoologique d'Anvers. *Bull. Soc. R. Zool. Anvers* 8, 3-62.

Gijzen, A. (1959) Das Okapi. Die Neue Brehmbücherei, Wittenberg, Lutherstadt, 231, 1-120.

Goetz R. H., Budtz-Olsen O. (1955) Scientific Safari – the circulation of the giraffe. *S. A. Med. J.* 29, 773-776.

Goetz R. H., Keen E. N. (1957) Some aspects of the cardiovascular system in the giraffe. *Angiology* 8, 542-564.

Goetz R. H., Warren J. V., Gauer O. H., Patterson J. L., Doyle J. T., Keen E. N., McGregor M. (1960) Circulation of the Giraffe. *Circ. Res.* 8, 1049-1058.

Goodarzi A. A., Jeggo P. A. (2013) The Repair and Signaling Responses to DNA Double-Strand Breaks. *Adv. Genet.* 82, 1-45.

Gould S. J. (1996) The tallest tale. *Nat. Hist.* 105, 18-27.

Hargens A. R., Millard R. W., Pettersson K., Johansen K. (1987) Gravitational haemodynamics and oedema prevention in the giraffe. *Nature* 329, 59-60.

Honjo T., Kinoshita K., Muramatsu M. (2002) Molecular Mechanism of Class Switch Recombination: Linkage with Somatic Hypermutation. *Annu. Rev. Immunol.* 20, 165-196.

Huang L., Nesterenko A., Nie W., Wang J., Su W., Graphodatsky A. S., Yang F. (2008) Karyotypic evolution of giraffes (*Giraffa Camelopardalis*) revealed by cross-species chromosome painting with Chinese muntjac (*Muntiacus reevesi*) and human (*Homo sapiens*) paints. *Cytogenet. Genome Res.* 122, 132-138.

Jensen R. B., Carreira A., Kowalczykowski S. C. (2010) Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* 467, 678–683.

Kakarougkas A., Ismail A., Klement K., Goodarzi A. A., Conrad S., Freire R., Shibata A., Lobrich M., Jeggo P. A. (2013) Opposing roles for 53BP1 during homologous recombination. *Nucleic Acids Res.* 41(21), 9719-9731.

Keeney S., Giroux C. N., Kleckner N. (1997) Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. *Cell* 88.3, 375-384.

Khanna K. K., Jackson S. P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.* 27, 247-254.

Kingdon J. (1997) *The Kindon field guide to African mammals*. London: Academic Press.

Kodric-Brown A., Brown H. (1984) Truth in advertising: the kinds of traits favoured by sexual selection. *Am. Nat.* 124, 309-323.

Kulemzina A. I., Trifonov V. A., Perelman P. L., Rubtsova N. V., Volobuev V., Ferguson-Smith M. A., Stanyon R., Yang F., Graphodatsky A. S. (2009) Cross-species chromosome painting in Cetertiodactyla: Reconstructing the karyotype evolution in key phylogenetic lineages. *Chromosome Res.* 17, 419-436.

Larsen N. B., Hickson I. D. (2013) RecQ helicases: Conserved guardians of genomic integrity. *Adv. Exp. Med. Biol.* 767, 161–184.

Lee M. S., Edwards R. A., Thede G. L., Mark Glover J. N. (2005) Structure of the BRCT Repeat Domain of MDC1 and Its Specificity for the Free COOH-terminal End of the γ-H2AX Histone Tail. *J. Biol. Chem.* 280, 32053-32056.

Mallon D., Kümpel N., Quinn A., Shurter S., Lukas J., Hart J., Mapilanga J. (2013) *Okapia johnstoni*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 26 March 2015.

Minter-Dykhouse K., Ward I., Huen M. S. Y., Chen J., Lou Z. (2008) Distinct versus overlapping functions of MDC1 and 53BP1 in DNA damage response and tumorigenesis. *J. Biol. Chem.* 181, 727-735.

Mitchell G., Roberts D. G., van Sittert S. J., Skinner J. D. (2013) Orbit orientation and eye morphometrics in giraffes (*Giraffa Camelopardalis). Afr. Zool.* 48(2), 333-339.

Mitchell G., van Sittert S. J., Skinner J. D. (2009) *S. Afr. J. Wildl. Res.* 39(2), 175-180.

Muller D. W., Zerbe P., Codron D., Clauss M., Hatt J. M. (2011) A long life among ruminants: giraffids and other special cases. *Schweiz Arch Tierheilkd* 11, 515-519.

Pellow R. A. (1984) Food Consumption and Energy Budgets of the Giraffe. *J. Appl. Ecol.* 1, 141-159.

Pellow, R. A. (1984) The feeding ecology of a selective browser, the giraffe (*Giraffa Camelopardalis tippelskirchi). J. Zool.* 202, 57-81.

Petersen K. K., Hørlyck A., Østergaard K. H., Andresen J., Broegger T., Skovgaard N., Telinius N., Laher I., Bertelsen M. F., Grøndahl C., Smerup M., Secher N. H., Brøndum E., Hasenkam J. M., Wang T., Baandrup U., Aalkjaer C. (2013) Protection against high intravascular pressure in the leg of giraffes. *Am. J. Physiol.* 305(9), R1021-1030.

Riballo E., Kuhne M., Rief N., Doherty A., Smith G. C., Recio M. J. (2004) A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Mol. Cell* 16, 715-724.

Schatz D. G. (2004) V(D)J recombination. *Immunol. Rev.* 200.1, 5-11.

Skinner J. D., Hall-Marin A. J. (2009) A note on foetal growth and development of the giraffe *Giraffa camelopadalis giraffa. J. Zool.* 177, 73-79.

Stewart G. S., Wang B., Bignell C. R., Taylor A. M. R., Elledge S. J. (2003) MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature* 421, 961-966.

Stiff T., O'Driscoll M., Rief N., Iwabuchi K., Lobrich M., Jeggo P. A. (2004) ATM and DNA-PK function redundantly to phosphorylate H2AX following exposure to ionising radiation. *Cancer Res.* 64, 2390–2396.

van Sittert S. J., Skinner J. D., Mitchell G. (2010) From Fetus to Adult – An Allometric Analysis of the Giraffe Vertebral Column. *J. Exp. Zool. B. Mol. Dev. Evol.* 314B, 469-479.

Walker J. R., Corpina R. A., Goldberg J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412, 607–614.

ACADEMIC VITA

Brendan Wood
2021 Farmview Drive, Newtown Pa 18940 / wood.brendan1026@gmail.com

---

Education

Pennsylvania State University – Schreyer Honors College

Bachelors of Science in Biology – Vertebrate Physiology Option

Honors and Awards

The President's Freshman Award	2012
The President Sparks Award	2013
The Evan Pugh Scholar Award – Junior	2014
The Evan Pugh Scholar Award – Senior	2015
Edward C. Hammond Jr. Memorial Scholarship	2013

Association Memberships/ Activities

Phi Eta Sigma Honor Society	2012 – current
THON Rules and Regulations Committee Member	2011 – 2012
THON Rules and Regulations Security Leader	2012 – 2013
THON Donor & Alumni Relations Tour Guide	2013 – 2014
THON Captain - Donor & Alumni Relations - Company Relation	2014 - 2015
Penn State Intramural Soccer	2012, 2013, 2014, 2015
Learning Assistant for ECON 104: Macroeconomics	2013

Professional Experiences

Cavener Lab Undergraduate Research Assistant	2012-2015
Yardley-Makefield EMS – Emergency Medical Technician	2013
St. Mary Medical Center Internship – Pre-Medical Volunteer	2013