2013

Chondrodysplasia-Like Dwarfism in the Miniature Horse

John E. Eberth
University of Kentucky, john.eberth@uky.edu

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

John E. Eberth, Student

Dr. Earnest Bailey, Major Professor

Dr. Daniel Howe, Director of Graduate Studies

CHONDRODYSPLASIA-LIKE DWARFISM IN THE MINIATURE HORSE

THESIS

A thesis submitted in partial fulfillment of the
requirements of the degree of Masters in the
College of Agriculture at the University of Kentucky

By

John Edmund Eberth

Lexington, Kentucky

Director: Dr. Ernest Bailey, Professor of Veterinary Science

Lexington, Kentucky

2013

ABSTRACT OF THESIS

CHONDRODYSPLASIA-LIKE DWARFISM IN THE MINIATURE HORSE

Dwarfism is considered one of the most recognized congenital defects of animals and humans and can be hereditary or sporadic in cause and expression. There are two general morphologic categories within this vastly diverse disease. These categories are disproportionate and proportionate dwarfism and within each of these there are numerous phenotypes which have been extensively described in humans, and to a lesser extent in dogs, cattle, mice, chickens, and other domestic species. Ponies and Miniature horses largely differ from full size horses only by their stature. Ponies are often defined as those whose height is not greater than 14.2 hands; however the maximum height for Miniature horses is constitutionally defined as 8.2 hands. Dwarfism is not considered a desirable genetic trait for Miniature horses. A majority of these conformationally inferior horses showed consistent physical abnormalities typical of disproportionate dwarfisms as seen in other mammal species. A whole genome scan with the Illumina Equine SNP50 chip clearly implicated a region on ECA1 as being associated with dwarfism of horses. The region implicated on the horse chromosome 1 (Equus Caballus; ECA1) contained a candidate gene for dwarfism, aggrecan (ACAN). Mutations were found in Exons 2, 6, 11 and 15 with each mutation associated with a distinct type of dwarfism. These mutations are independently transmitted throughout the population. Absence of normal homozygotes for these mutations and absence of normal horses which were heterozygous for these mutations indicated that these alleles caused dwarfism in those genotypes. These genotypes did not explain all observed dwarves in this population.

KEYWORDS: Dwarfism, *Aggrecan*, *Equus Caballus*, Miniature Horse, Genetics, JPEG

John Edmund Eberth
Student's Signature

September 11, 2013
Date

CHONDRODYSPLASIA-LIKE DWARFISM IN THE MINIATURE HORSE

By

John Edmund Eberth

Earnest Bailey
Director of Thesis

Daniel K. Howe
Director of Graduate Studies

September 11, 2013
Date

This thesis is dedicated to….

My mother for the inspiration to pursue a dream,

My father for his knowledge and drive I aspire to,

My family for their ability to lighten the load of life,

My wife for her endless love and devotion,

My teachers for igniting my fire to learn,

My professors for stoking that fire,

My friends for lightheartedness,

Most importantly my children, to look in your eyes I know God’s love.

Thank You for your patience and understanding.

ACKNOWLEDGMENTS

First and foremost, I would like to thank all the individuals who supported or helped me through my research and education. I am forever grateful to each one of you.

I would particularly like to give my utmost gratitude to my advisor Dr. Ernest Bailey. I could not have accomplished this research without his unwavering support and encouragement. His patience with me is an attribute I hope to someday be able to emulate. His willingness to share frankly his knowledge and life experiences will be deeply cherished. He has motivated me to be the best scientist I could be; pushing me to go beyond what is acceptable to demand exceptional.

I would also like to thank the members of my thesis committee, Dr. Teri Lear and Dr. Kathy Graves.I am eternally appreciative of their willingness to contribute their time and advice to my research. Their mentoring and care to ensure it would be well done is support I am truly grateful to have had.

I would like to thank all my fellow members(past and present) and associates of the Bailey Lab, specifically, Judy Lundquist, Dr. Samantha Brooks, Rose Magee, Debbie Cook, and Allison Sparling. Without Judy I probably would have never accomplished anything in the lab. Sam is a wonderful friend; she challenged me to critically analyze even the most minute information. Rose and Debbie are great friends and I enjoyed all of our numerous rants and frustrations we shared during our time together in the lab. I am grateful for all of their help at times when I was at my wits end. Allison is the newbie to the lab and quickly has become someone I can trust and rely on to be there to help whenever it is needed, without question.

My sincere thanks go to all my friends and colleagues at the Gluck Equine Research Center. I give special thanks to Mr. Roy Leach, Drs. Mats Troedsson and Peter Timoney as Department Chairs. I am especially grateful to Drs. Barry Fitzgerald and Dan Howe as Directors of Graduate Studies for their assistance with my tenure here.

And to all my fellow graduate students, my time here has been highlighted by numerous experiences that I will cherish for life. I also would like to give special thanks to Brock Daugherty and Mandy Eakins in the Animal Genetics Testing Lab for their assistance. Finally, special thanks go to Drs. Tom Swerczek and James MacLeod for their mentoring advice and passionate assistance in my research endeavors.

I would like to thank my family for their committed support over the many, many years it has taken to accomplish this work. My parents and siblings for their love and encouraging support during all the highs and lows of graduate school and research. I am eternally grateful for the love and devotion of my wife, Melinda. I could not have accomplished any of this without her by my side. I am grateful for my three daughters, Mariah, Lauren and Marissa. They always showed interest in what I was doing and have the ability to always make me smile.

"Gratitude is not only the greatest of virtues, but the parent of all others." ― Marcus Tullius Cicero

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1
The Miniature Horse Breed

Summary

The Miniature horse breed is defined by a registry for horses of small stature. Miniature horses are a breed of horse in which the only defining characteristic is a maximum height up to 34 inches (AMHA, American Miniature Horse Association 2013) or up to 38 inches (AMHR, American Miniature Horse Registry 2013) at the base of the mane, depending on the registry. Any color combination and any physical type of Miniature is acceptable to be registered as long as the height requirement is met. This type of flexibility in the breed's characteristics requirements has allowed the Miniature horse industry to explode in popularity all over the world.

Until very recently, open studbooks for Miniature registries allowed the use of other pony breed crosses into the gene pool of the Miniature breed as long as the resulting offspring was under the height restriction when mature. Enthusiasts of all backgrounds and fanciers have been allowed to breed the colors and body types they have so desired without the registries' constraints that would be typical for pure-bred horse registries. Worldwide, there are dozens of Miniature horse registries. Some organizations stress breeding of Miniatures with proportional large horse characteristics, others promote the Miniatures to retain traditional pony characteristics (AMHA 2013; AMHR 2013; BMHS, British Miniature Horse Society 2013).

The two largest registries in the United States (as well as the world) for Miniature Horses are; the American Miniature Horse Association (AMHA) and the American Miniature Horse Registry (AMHR). The AMHA standard of perfection of today suggests that if a person were to see a photograph of a Miniature horse, without any size reference, it would be identical in characteristics, conformation, and proportion to a full-size horse.

General appearance is to be of a well-balanced horse with correct conformation characteristics required of most large horse breeds with any color or marking pattern, as well as any eye color. The head should be in proportion to length of neck and body yet broad and relatively short from eye to muzzle with large eyes set wide apart with medium sized ears carried alertly. The neck is to be flexible and lengthy blending smoothly into the withers with a clean well-defined throat latch allowing flexibility at the poll. Shoulders are to be well angulated and long-sloping that allows a free moving stride with alert head carriage. The body is to be well muscled and balanced with a short back and proportionate with ample bone. The topline should be smooth and level with a trim barrel, deep girth and flank with hindquarters that are long with a well-muscled hip, thigh and gaskin. The croup should be the same height as the withers and tail set smoothly rounding at the rump. The legs should provide ample ability for smooth fluid gates when in motion. When viewed from front or back the legs are to be set straight and parallel with pasterns and hooves to hold a 45 degree angle blending smoothly. Hooves should be round and compact and trimmed appropriately for an unshod horse (AMHA 2013). The standard of perfection for the AMHR Miniature is similar to the AMHA standard, however with a taller height limit. For its size, the Miniature is quite possibly the strongest of all horse and pony breeds. It can pull twice its own weight under circumstances where a draft horse can only pull approximately half its own weight. Miniatures are found worldwide including North America, Europe, Asia, Africa, South America and Australia, though the majority of the breed population is in Europe and North America.

Introduction

History and Foundation of the Miniature Horse

The foundation of the Miniature horse breed originated from a large number of different populations worldwide. Many different pony breeds (originally landraces) have been bred for small size, including the Shetland pony, Exmoor pony, and Dartmoor pony. The Exmoor and Dartmoor have height variations from 12.3 hands down to 44 inches and the Shetland down to 28 inches; however, there is no minimum height for these registries

(Dartmoor Pony Society 2013; Exmoor Pony Society 2013; Shetland Pony Stud-Book Society 2013). The Miniature is the smallest equine breed known to exist with no minimum height (AMHA 2013; AMHR 2013).

In 17th century Europe, according to records of 1650 from the court of the French King Louis XIV (1643-1715), third monarch of the Bourbon family, the presence of tiny horses among the exotic animals in the king's zoo are noted. Paintings and articles featured minute horses by 1765 (*International Museum of the Horse,* Kentucky Horse Park 2013). However, many early Miniature sized horses were not pets of kings and queens. Many were used to work in the English Midlands, Wales and Northern European coal mines as pit ponies until the European industrial revolution (*International Museum of the Horse,* Kentucky Horse Park 2013). Later in other countries such as the United States, these pit ponies were used until the industrial revolution (Naviaux 1999).

In other parts of the world, the Miniature horse has had similar foundations to those in Europe and North America. The Falabella Miniature horse was originally developed in Argentina in 1868 by Patrick Newell. When Mr. Newell died, the herd and breeding methods were given to Newell's son-in-law, Juan Falabella. Juan supposedly added additional bloodlines including the Welsh Pony, Shetland pony, and small Thoroughbreds; however, that could not be verified. Utilizing extensive inbreeding, he was able to produce small size within the herd (Naviaux 1999). South Africa's Miniature horses were developed in that nation and are recognized as the South African Miniature horse. Mr. Wynand de Wet of Lindley, South Africa, started in 1945 with two Shetland pony mares and a Shetland stallion. Through strict selection, the offspring became smaller and in 1991 Mr. de Wet apparently produced a mare that was only 25 inches tall (66cm) (South African Miniature Horse Breeders Society 2012).

The foundation horses of the European Miniature registries exhibited physical characteristics with striking similarities to the European pony breeds listed earlier, including coat color patterns. The Shetland pony breed, originating from the Shetland Isles, was a significant contributor to the foundation of the European Miniature horse. In

appearance, European Shetlands and Foundation and Classic American Shetlands have very similar characteristics to Miniatures which include a small head, sometimes with a dished face, wide spaced eyes and small and alert ears (AMHA 2013; AMHR 2013; ASPC, American Shetland Pony Club 2013). The original United Kingdom Shetland Pony breed, as well as the founding European Miniature horses, had a short, muscular neck, compact, stocky bodies, and short, strong legs and a shorter than normal cannon bone in relation to their size. A short broad back and deep girth are universal characteristics as is a springy stride. UK Shetland ponies have long thick manes and tails and a dense double winter coat to withstand harsh weather. These are identical characteristics to the European Miniature and the early American Miniature horse (Naviaux 1999). In general, modern European Miniatures tend to preserve more of the original characteristics of the breed and are often stockier than their current American Miniature horse relatives (AMHA 2013; AMHR 2013; BMHS 2013).

This earliest history of the Miniature horse in the United States is similar to that of the European Miniature. The first Shetland ponies in the United States, for which there are written records, were imported in 1885 by Eli Elliot. These ponies provided the foundation stock for the development of the American Shetland registry (ASPC 2013) and later the development of the American Miniature Horse Registry (AMHR) (Naviaux 1999). Some of these Shetlands were crossed with ponies of other breeds, including the Hackney pony, Welsh pony, POA (Pony of America), and Harness Show Pony. The breeding of American Shetlands was mainly centered in Indiana, Illinois and Iowa. The Pony of the Americas breed (POA) has its origin in America, where Leslie Boomhower, an Iowa breeder of Shetland Ponies, founded the breed in 1954. He acquired an Arabian/Appaloosa mare that had been accidentally bred to a Shetland pony stallion. Boomhower named the resulting colt Black Hand, and he became the foundation sire for the POA breed. POA breeding was centered in Indiana, Illinois and Iowa like the Shetlands (Dutson 2005). This region has been universally recognized as having some of the founding Miniature horse lines. Most importantly some of those founding Miniature horse breeders were well-known Shetland breeders. This coincidence leads to a vast amount of anecdotal evidence of an infusion of bloodlines from certain large pony breeds

such as the Welsh, American Shetland, and POA, as some of the founding lines of the Miniature horse (ASPC 2013; AMHR 2013). Some American Shetland pony breeders in Virginia started to gather the smallest ponies they could find beginning in the 1930's through the 1960's to breed a "midget pony". Some of these were to be among the first Miniature horses registered in the AMHR studbook in 1972 (Naviuax 1999). In the United States, the Miniature horse breed was refined conformationally even more during the later 20th and early 21st century by utilizing more lines from sources that include the Modern Shetland and Hackney Pony (AMHR 2013).

Miniatures, like the Shetland pony, appear with most colors and color patterns including black, chestnut, bay, brown, gray, palomino, dun, roan, cremello, silver dapple pinto, overo, sabino and tovero. Interestingly, the Miniature breed also has appaloosa and champagne colors. Registered European Shetland ponies are not Appaloosa, nor do they carry the champagne gene. American Shetland ponies do carry appaloosa and champagne. Evaluating the history of some of these other pony breeds mentioned above, modern DNA testing for coat color, and breed registry constrictions of color within each breed, brings about an interesting question. When observing the existence of appaloosa and champagne coat colors in the Miniatures all over the world, how does a relatively new breed of horse with registries less than 40 years old express almost ALL colors seen in equines? This observation quite possibly makes coat color testing a sound and helpful technique to investigate of the Miniature horse's multiple breed history.

Genetic Background and selection for size in the Miniature

The genetic background of the Miniature horse is quite diverse, based on the heterogeneous origin of the foundation stock. Genetic studies have shown extensive variation at the level of blood groups, biochemical polymorphisms and DNA markers (Bailey, personal communication). The major premise for many breeders has been to produce small, well-proportioned horses. Ponies and Miniature horses differ from full size horse only by their stature. Ponies are often defined as those whose height is not greater than 14.2 hands; however, the maximum height for Miniature horses is constitutionally defined as 8.2 hands. The entire premise of the Miniature horse is to

breed the smallest most correct and proportionate horse in Miniature. This reduced stature is usually the cumulative effect of hundreds of genes, each having a small impact on stature. Unfortunately, there are also dwarfism genes which greatly reduce statute and may negatively impact health and reproduction. This is not considered a desirable genetic trait for Miniature horses. Some breeders focused on selection for extreme diminutive size in the mid to latter part of the 20th century at a time when the breed was also experiencing great popularity and rapid population growth. Consequently, any gene which decreased size was selected without regard to its effect on conformation or fitness. While reduced stature can be attained though many generations of breeding for genes which combine to reduce size, other genes can reduce stature rapidly (in a single generation) and produce phenotypes which are commonly referred to as dwarfs.

Genetics of Stature

Stature in all organisms is the additive interaction of dozens if not hundreds of genes within an individual that determines its ultimate mature height. This polygenic trait is primarily due to genetics (80 – 90 %) and only slightly affected by other environmental factors. This mature height has been shown to be quite variable within certain isolated breeding populations of animal species including humans. Genome-wide association studies (GWAS) using SNP (Single Nucleotide Polymorphism) analyses have implicated numerous chromosomal sites involving genes that regulate to some degree mature height in humans (Allen *et al.* 2010).

Gudbjartsson *et al.* (2008) discovered that candidate genes in those chromosomal sites that were known to be involved in growth could be divided into three functional groups. The groups were based on skeletal development, processes involving mitosis and chromosomal segregation, and cartilage composition. A total of 58 loci were found to be significantly associated within Caucasians. It was estimated that individually, each of the candidate genes found in those regions only had a minute effect, (0.4 cm) on mature height. However, collectively these same genes may wield a significant effect and account for roughly 4% of mature stature. The more "tall" alleles an individual may have, the taller the individual ultimately becomes. Weedon *et al.* (2008) discovered a

five cm difference in adult stature between individuals with 17 or fewer "tall" alleles when compared to those individuals with 27 or more. A study of Koreans by Kim *et al.* (2010) found another 15 loci associated with stature. With nearly 200 linked loci identified to stature suggesting almost 700 genes identified having some involvement with height determination in humans, this scenario can only explain approximately 16% of its variation (Allen *et al.* 2010).

Stature variation within various animal species has been shown to exist to some degree. Domesticated animals have been shown to have great variation in height within a species (Clutton-Block 1999). GWAS studies in cattle, horses and dogs utilizing genetic information gleamed from GWAS stature studies in humans have strongly suggested that polymorphic regions affecting human stature also affect stature in these domesticated species. Pryce *et al.* (2011) looked at orthologous genes that could be mapped to Bovine Genome v4.0. Fifty-five genes were found and mapped with 879 SNPs in the Bovine Build that were known to be associated with stature in humans. Dairy and beef cattle populations were tested. It was found that 10 of the genes associated in humans were associated with stature in dairy cattle and 12 with stature in beef cattle. Two of the associated genes for stature in humans were also found to be among the genes in this study; HMGA2 within the Beef cattle and LCORL in the Dairy cattle.

Makvandi-Nejad *et al.* (2012) reported four loci on four different chromosomes associated with determining 83% of height across 16 horse breeds ranging from the smallest known breed (Miniature) to the largest (Shire). Of the genes linked to the GWAS scan SNPs, the highest valued genome-wide associated SNP was on chromosome 3 located 100 kb upstream of the ligand dependent nuclear receptor corepressor-like (LCORL) gene. It is a transcription factor that has repeatedly been associated with human height (Allen *et al.* 2010; Gudbjartsson *et al.* 2008; Weedon *et al.* 2008; Kim *et al.* 2010). The association value at this SNP was almost the maximum possible in the 16 breed scan. The alleles associated with this gene segregate by horse size close to perfect. LCORL and/or HMGA2 have been found to be strongly associated with stature variation, confirming what has been seen in humans, cattle and mice.

One of the most compelling finds in stature determination was done by Sutter *et al.* (2007). Using GWAS, Sutter was able to pinpoint a single allele of IGF-1 as a major determining factor in dog size across 143 breeds. This evidence showed clear evidence that a single gene/allele can have a significant impact on the mature stature of an organism, therefore altering the notion that height is strictly a quantitative trait, especially in certain highly inbred organisms with specific characteristics. Considering the GWAS studies in cattle, dogs, and horses, it is becoming increasingly evident that a short list of genes repeatedly selected is responsible for stature in the domestication of animals.

GWAS in stature has proven an effective tool in giving insight into controlling genetic forces of complex polygenic traits. The issue of stratification is highly unlikely to cause the same genetic artifacts across different species. It also gives tremendous strength to the physiological relationship of a gene in linkage disequilibrium to an associated single nucleotide polymorphism and a specific trait, as seen in dogs. In consideration of GWAS and stature studies in various different human races and animal species, it gives strong argumentative information that there is no known mechanism associating genes found with GWAS and complex traits such as stature, with those genes that show linkage (Pryce *et al.* 2011).

Genetics and Biology of Dwarfism

Dwarfism is considered one of the most recognized congenital defects of animals and humans and can be hereditary or sporadic in cause and expression. There are two general morphologic categories within this vastly diverse disease. These categories are disproportionate and proportionate dwarfism and within each of these there are numerous phenotypes which have been extensively described in humans, and to a lesser extent in dogs, cattle, mice, chickens, and other domestic species. Among humans, over 200 known distinct types of dwarfism have been reported, each with a different underlying cause. The approximate worldwide population of dwarfism numbers 190,000 in humans. Causes of dwarfisms in humans can involve genes responsible for hormone production and/or recognition, metabolism, cartilage development, pituitary gland hormones, thyroid gland hormones, and bone growth plate development. The most common type of

dwarfism in humans is Achondroplasia, which affects about 1 in every 40,000 children worldwide. This number varies widely depending on the country. Eighty percent of all human dwarfisms are Achondroplasia. This disproportionate dwarfism involves mostly the shortening of the limbs with an averaged sized trunk, enlarged head with a prominent forehead with other lesser skeletal malformations. Achondroplasia is caused by a single base substitution of one of two different bases G → A or G → C, resulting in an amino acid change from arginine to glycine (Shiang *et al.* 1994). This FGFR3 (fibroblast growth factor – 3) mutation is most commonly sporadic with a small portion being inherited. It is estimated that 150,000 people have Achondroplasia worldwide. The FGFR3 gene is part of a family of genes responsible for determining cell type, regulation of cell growth and division, blood vessel formation and embryo development. Besides Achondroplasia, different mutations in FGFR3 can cause other more rare types of dwarfisms; Hypochodroplasia, SADDAN (Severe Achondroplasia with Developmental Delay and Acanthosis Nigricans) (Bellus *et al.* 1995), and Thanatophoric Dysplasia (Pannier *et al.* 2009).

Turner Syndrome is a type of dwarfism caused by a female having a severely altered or missing X chromosome. One of the side effects of the missing X is that she is unusually short in stature (4‘8” or less). The male version of this disease is called Noonan Syndrome. Males are usually short as well (50-70% of the cases). Both males and females have other characteristic abnormalities with facial structure, heart defects and other skeletal malformations.

Pituitary dwarfism (hypophysis) is the result of the anterior half of the pituitary gland not producing enough growth hormones involving the gene POU1F1 (formerly *Pit 1*, Pituitary-specific positive transcription factor 1) (Irie *et al*. 1995; Hendriks-Stegeman *et al*. 2001), or no hormones at all (Panhypopituitarism). Pituitary dwarfism causes the person to have proportionate body structures for their decreased height compared to age of development. Other abnormal characteristics may be involved specifically when dealing with mutations in the gene PROP-1 (Pituitary-Specific Homeodomain Factor), and its related CPHD (Combined Pituitary Hormone Deficiency) due to the hormone

cascade necessary for growth and development involving multiple hormones (Bottner *et al.* 2004). Hypopituitarism may be acquired after childbirth from causes other than genetic. It could be caused by trauma to the pituitary gland by injury or from surgery (tumor removal) or even environment.

Cretinism is caused by a decrease in thyroid hormone production by the thyroid gland during childhood resulting in a type of dwarfism characterized by short stature with disproportionally small limbs and other skeletal structures, mental retardation and lack of sexual maturity. This can be caused by environmental (endemic goitre) or genetic factors (congential nongoitrous hypothyroidism) caused by mutations in the Thyroid Hormone Receptor Alpha (THRA) gene (Bochukova *et al.* 2012; van Mullem *et al.* 2012).

The vast number and variety of dwarfism in the humans as compared to all other species is quite interesting and complex. One consideration to the perpetuation and complexity of these types of dwarfisms could be due to the increasing public acceptance of the diseases, modern medicine assisting in survival of individuals affected by a type of dwarfism and the individuals' desire of reproduction. Other mammalian dwarfisms have been shown to exist. In cattle specifically, disproportionate dwarfism has been reported in different breeds which include; Aberdeen, Angus, Dexter, Holstein, Hereford, Japanese Brown and Shorthorn. Mutations causing dwarfism have been reported in Japanese Brown and Australian Dexter cattle, as well as expression data showing down regulation of IGF-1 in Holstein dwarfs (Takeda *et al.* 2002; Cavanagh *et al.* 2007; Blum *et al.* 2006). Bovine Chondrodysplastic Dwarfism (BCD) in Japanese Brown cattle is an autosomal recessive disorder (Takeda *et al.* 2002). The gene responsible for BCD is limbin (Lbn) and is expressed in the epiphysial growth plates of long bones (Moritomo *et al.* 1992). In the mid thru later 20th century, the Hereford, Angus, and Shorthorn cattle breeds experienced a type of dwarfism. Interestingly, during that period of time cattle breeders were producing cattle of a particular type and build with heavy or thick bone and highly muscular but small in height. Physical characteristics used to describe the dwarfism in cattle were similar to human Achondroplasia.

Currently, only one other equine breed with recently documented dwarfism is the Friesian (Back *et al.* 2008). GWAS scan was performed and associated a 2-MB region of chromosome 14 (Orr *et al.* 2010). This dwarfism in Friesians has a phenotype quite similar to skeletal atavism in the Miniature horse (Tyson *et al.* 2004; Rafiti *et al.* 2013). The Friesian dwarfism has shown to be inherited as autosomal recessive, but as yet the genetic cause is unknown. With recent and past GWAS done regarding dwarfism, other diseases and stature, this process was used as the basis for this study conducted to discover the genetic basis for dwarfism in Miniature horses.

Aggrecan and Dwarfism

Gleghorn *et al.* (2005), was first to identify a mutation in *Aggrecan (ACAN)* that caused a form of human dwarfism. Spondyloepiphyseal Dysplasia (SED) type Kimberley (SEDK) is an autosomal dominant mild condition that expresses characteristics similar to (SED) however with severe premature arthritis. SED is a heterogeneous collection of disorders that are characterized by general shortening of the limbs and torso of individuals. SEDK was mapped to a locus in the human genome that contained *ACAN*, and discovered a single base insertion in the exon 12 variable repeat region of *ACAN* causing a premature stop codon and thus a truncated non-functional protein that is ~60% of normal size. Gleghorn proposed that the novel protein accumulated in the Endoplasmic Reticulum (ER) and not in the Extra-Cellular Matrix (ECM) as was seen in the chicken nanomelic mutation (Li *et al.* 1993; Vertel *et al*. 1993) and in the *CMD* mouse model (Watanabe *et al*. 1997) and thus caused a reduction in steady state levels of *Aggrecan (ACAN)* in the ECM. These observations in animal models suggest that even the reduction of normal levels of functional *ACAN* in the ECM is sufficient to cause a dwarf phenotype. Therefore, it was proposed that the SEDK mutation may result in lower mRNA amounts of functional *ACAN* thus causing the mild dominant heterozygous condition seen.

Thompson *et al.* 2008, describes the condition SEMD *Aggrecan* Type involving a recessive mutation in the C-type Lectin domain of *ACAN* in humans. Characteristics involve rhizomelia and mesomelia of the limbs with no bowing, and severe

brachydactyly in the hands and phalanges. Torsos were barrel chested with mild lumbar lordosis and shortened necks. Facial features included posteriorly rotated ears, prognathism that was slightly low set with severe midfacial hypoplasia and almost absent nasal cartilage and macrocephaly. The mutation was inherited from both parents who were heterozygotes for the (c.6799 G → A) p. D2267N mutation (aspartic acid → Asparagine) and the unaffected sibling was homozygous for the WT. An interesting side note from the data showed the two parents to be somewhat short (59" – 60") but proportionate and the homozygote WT sibling was 70" in height.

Stattin *et al*. 2010, describes an autosomal dominant familial dwarfism osteochondritis dissecans (OCD) which develops early onset osteoarthritis. This familial type affects numerous joints within the individual and is accompanied with a disproportionate short stature. A missense mutation was found in the G3 domain of the C-type lectin domain (CLD) of *ACAN* involving a valine to methionine. This mutated residue is located within the hydrophobic core and is physically directly below the ligand binding surface of the CLD. This missense mutation may alter the binding surface conformation. Loss of ligand binding interaction was shown using mammalian-expressed recombinant G3 fragments. This loss of ECM ligand binding and interaction leads to the disruption of the ECM framework, ultimately leading to unstable cartilage causing OCD and early osteoarthritis.

ACAN mutations are known to cause dwarfism in humans as well as animals. In avian species the disease nanomelia (nm) is an autosomal recessive homozygous embryonic lethal mutation. Embryos express a phenotype of hypoplastic limbs and a parrot shaped beak. The single based substitution causes a premature stop codon to occur in the CS-1 region. This makes a truncated *ACAN* protein that was shown neither to be processed post translation nor to be secreted into the ECM (Li *et al.* 1993).

Canvanagh *et al.* (2007) found two different mutations in *ACAN* inherited in Dexter cattle that cause the same phenotype. The Dexter Cow has a type of dwarfism inherited as autosomal co-dominant where the carriers show characteristic short legs

compared to non-carriers. This condition has been well documented anecdotally since the 19th century. Lethal homozygote affected calves are aborted in late gestation with micromelia of the limbs, cleft palate, extreme shortening of ribs and trunk with herniated abdomen. The head is abnormal with midface hypoplasia and substantial prognathia with protruding tongue. Two separate mutations were found in two different families. One family contained an insertion in exon 11 between bp 2266_2267 of GGCA, which introduced a frameshift and caused a stop codon in exon 11. The frameshift occurred at amino acid position 756 subsequently, the stop codon is amino acid position 914. This is substantially shorter than the normal *ACAN* protein of 2327 amino acids. The transition mutation in exon 1 of another family (-198C→ T) is predicted to introduce a new ATG start codon 199bp upstream of the normal start codon which produces a 91 amino acid protein bearing no resemblance to the *ACAN* protein due to the frameshift nature of the new start codon. These two variants account for all cases of Dexter chondrodysplasia tested and carriers.

Mice have been ideal models for genetic research in mammals. Cartilage Matrix Deficiency (CMD) mice were the first known genetic disorder of proteoglycans identified in a mammal species (Rittenhouse *et al.* 1978). It is an autosomal recessive disorder (Kimata *et al.* 1981) where homozygous CMD mice are born with severe dwarfism characterized by short limbs, tail, and snout with a cleft palate resulting in death shortly after birth. The presumed cause of death is the malformation of the tracheal cartilage. The mutation is a 7-bp deletion in exon 5 of *ACAN* causing an early termination codon in exon 6 (Watanabe *et al.* 1994). Heterozygous cmd mice are noted to have slight dwarfism and develop age-related hyperlordosis in the cervical spine and kyphosis in the thoraco-lumbar spine resulting in spastic abnormal gait causing marked decrease of movement with eventual death due to inability to eat. Expression data showed decreased *ACAN* mRNA in both heterozygotes (81% of normal) and homozygotes (41% of normal). Chondroitin sulfate levels were 87% of normal in the spine of heterozygotes from WT mice (Watanabe 1997).

Aggrecan Structure and Function

The investigation of dwarfism led us to a study of the gene for *ACAN*, as described in chapter 2. Mutations of *ACAN* have been found to cause dwarfism in several species (Li *et al.* 1993; Watanabe *et al.* 1994; Cavanagh *et al.* 2007; Tompson *et al.* 2009). Therefore, reviewing the genetics and biology is pertinent to this introduction.

Synovial joint structure and function are intricately designed to provide smooth gliding surfaces for joint movements with little wear and tear for most of the organism's lifespan. This integral design within structures of joints involves articular cartilage covering the joint's areas of wear. These areas are specifically composed of hyaline cartilage which is simply a thin, smooth and firm layer able to provide a cushioned surface to sustain weight bearing loads (Kiani *et al.* 2002). The cellular composition of cartilage is chondrocytes and the extra-cellular matrix (ECM) they produced. The biochemical makeup of cartilage and physical structure and function of joints are vitally reliant on the integrity of the ECM (Kiani *et al.* 2002). The matrix molecules in cartilage include type II collagen, hyaluronan (HA), proteoglycans, glycoproteins and a variety of elastic fibers. Proteoglycans are made up of a family of glycoconjugates with a central core protein. This core protein is then covalently bonded, in post-translational modification, to one or more glycosaminoglycan (GAG) side chains (Wight *et al.* 1992). A majority of all proteoglycans function as aggregates (Sajdera *et al.* 1969) created by the non-covalent bonding with hyaluronan and link protein (Hardingham *et al.* 1972; Hardingham *et al.* 1973; Heingard *et al.* 1974).

When proteoglycans are in high concentration, it creates a substantial osmotic swelling effect drawing water into the tissues where the proteoglycans are located. The most essential cartilage proteoglycan to the normal performance of articular cartilage is *ACAN*. Up to 10% of cartilage is made up of proteoglycans, and of those, *ACAN* is the main constituent. This is one of the large multi-modular aggregating chondroitin sulfate proteoglycans within the ECM (Zimmermann *et al.* 1989). The molecule is essential to the proper function of articular cartilage by supplying a hydrated gelatinous constitution formed by its interaction with hyaluronan and link protein. It is found in other connective

tissues in a lesser amount as well. This influx of water causes the *ACAN*-rich matrix to swell and expand. This matrix in the water swollen state is critical to the biomechanical properties of cartilage and the ability to allow joints to handle wear and weight bearing loads. The interaction of *ACAN* with collagen is another vital characteristic to the cartilage's functional ability of stiffness and resistance to deforming under stress. *ACAN* is significantly restricted within the matrix with regards to its mobility which gives it the ability to retain fluid or the redistribution of water (Mow *et al.* 1989). Figure 1.1 shows a schematic of the interactions of the proteins within the human ECM.

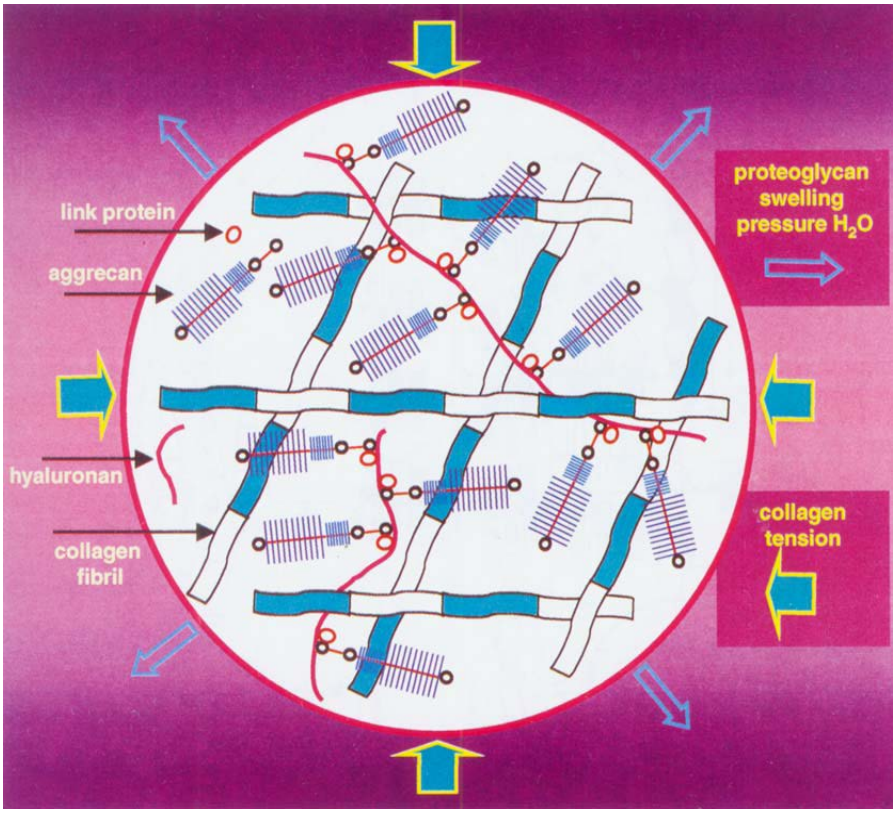


Figure 1.1. Combined properties and spatial interactions of collagens and *Aggrecan* in articular cartilage environment (Kiani *et al.* 2002).

This gel-like substance is due to the numerous negatively charged anionic molecules on the GAG side chains within the *ACAN* protein that carry Na+ as counter positive ions. This causes a significant disparity in ion concentrations between cartilage and the surrounding tissue resulting in diffusible free anions and cations to be out of balance. This osmotic pressure draws water into the cartilage tissue due to *ACAN*'s size and bonded structure with other proteoglycans.

The *ACAN* core protein is composed of three domains designated G1, G2 and G3, and has attached numerous keratan sulphate (KS) chains with N- and O-linked oligosaccharides and a large region of chondroitin sulphate chains (CS) between the G2 and G3 domains allowing glycosaminoglycan (GAG) side chain attachment post-translationally. The link protein is a small glycoprotein that helps stabilize the overall structure formation and has a similar structural design to G1. Whereby, the G1 globular domain of the protein located at the N-terminal end specifically binds to hyaluronan by a non-covalent bond with link protein (Heingard *et al.* 1974). This is part of the tertiary structure complex within the ECM. Situated between the G2 domain and G3 domain is a large sequence modified by KS side chains and CS chains. Each *ACAN* contains ~ 100 chondroitin sulfate chains, which are typically ~ 20 kDa each. There are fewer keratan sulfate chains (up to 60) and they are usually of smaller size (5-15 kDa). These chains combine to make up 90% of *ACAN*'s mass (Kiani *et al.* 2002). The G2 has tandem repeats which make it homologous to G1 and link protein. G2 is separated from G1 by an interlobular domain (IGD) and is involved in processing product (Fosang *et al.* 1989; Paulsson *et al.* 1987). Located on the C-terminus end within the G3 domain are two epidermal growth factor-like sections, EGF1 and EGF2. Other sites include a carbohydrate recognition domain (CRD), complement binding protein (CBP)-like sites, a folded immunoglobulin region and tandem repeats of proteoglycan (Perkins *et al.* 1989). Figure 1.2 shows the globular and domain structures of *ACAN*.

Sequence conservation is very high within the three globular domains (G1, G2, and G3) of known *ACAN* sequences across species, including humans. The regions of chondroitin sulfate chains and keratan sulfate chains are less conserved across species. There is a strong correlation between exon number sequence and structure domain layout of *ACAN*. This is most evident in that the chondroitin sulfate chain domain is specific to a single exon (Kiani *et al.* 2002). Exon 1 encodes a 5' end sequence that is untranslated and exon 2 contains a translation start codon site for protein synthesis. G1 domain is divided into loop sections of A, B, B' and correspond to subsequent exons 3, 4 - 5, and 6, respectively. Link protein exons 3, 4-5, 6 correspond to the G1 loop sections in sequence similarity confirming binding site affinities (Valhmu *et al.* 1995). The G1 domain is

comprised of three modules: an immunoglobulin fold module and two copies of an HA-binding link module, this is also referred to as the PTR or proteoglycan tandem repeat. The G1 domain is also comprised of two cysteine-rich motifs that form disulfide bonds. These disulfide bonds are involved in *ACAN* interactions with HA to form large complexes (Watanabe *et al.* 1998).

The rod shaped structure separating the G1 and G2 domains is the interglobulin domain (IGD). The IGD is encoded strictly by exon 7 of the gene *ACAN*. This structure and sequence is unique to the *ACAN* molecule. Interestingly, during pathological cartilage degradation, this is the area of proteolytic attack by Aggrecanase-1 at the Glu373-Ala374 bond. This appears to be the site of *ACAN* turnover physiologically (Tortorella *et al.* 1999). Matrix metalloproteinases (MMP) and Aggrecanase-1 are both believed to be involved in *ACAN* molecule turnover in both diseased and normal cartilage (Lark *et al.* 1997).

The G2 domain involves exons 8, 9, 10 of the gene. There are two tandem repeats of proteoglycan that are about 67% similar in sequence to the G1 PTR repeats, however, it shows no functional ability to bind with HA-binding ability with link protein possibly due to karatan sulfate side chains (Fosang *et al.* 1991). Interestingly, *ACAN* is the only proteoglycan family member that has this G2 sequence and structure domain. *ACAN* is also the only proteoglycan family member with the keratan sulfate and chondroitin chains that make this protein highly glycosylated. The G2 domain presently is thought to only function as a quality control domain in producing a mature working *ACAN* protein (Kiani *et al.* 2001).

The (KS) keratan sulfate domain, encoded by exon 11 in humans, follows the G2 domain. This sequence of amino acids makes up a group of tandem repeat chains which varies in number greatly across species. Rats, mice and chicken (Krueger *et al.* 1990) lack this repeat sequence region, however; bovine (Antonsson *et al.* 1989), human (Doege *et al.* 1991) and horse seen in this current study, contain these sequences in various amounts and in a strikingly similar repeat fashion. This hexamer sequence

contains proline-serine and proline-threonine repeats suggesting this region as possible candidate areas for KS chain substitution. Humans contain additional residue sequences of 66 highly conserved hexamer motifs repeated 11 times. This confirms the human sequence is more glycosylated than the rat or bovine which contain less repeats. The KS region is not unique within the *ACAN* core protein. KS attachments also occur within the CS chains by being oligosaccharide linked to threonine (Hopwood *et al.* 1974). A KS structural characteristic seen in differing cartilage tissue sources within an organism is most noticeably between non-bearing and load-bearing cartilage. KS chains in joints that are load-bearing contain 1-3 fuctose residues and 2-6 N-acetyl neuraminic acid residues unlike cartilage in the nose (Nieduszynski *et al.* 1990). Another unique characteristic to keratan sulfate chains within human meniscus is that, as humans age the KS chains and the 6-sulfated disaccharide of chondroitin sulfate chains increase in concentrations (McNicol *et al.* 1980). The complete function of the keratan sulfate chains are not known, however it is suggested that they may contribute to tissue development, processing and distribution. The one important characteristic of the KS chain is the binding of water molecules within *ACAN*, significantly enhancing the ECM's ability to handle load bearing stress (Kiani *et al.* 2001).

Chondroitin sulfate (CS) region is *ACAN*'s largest domain. It is solely encoded by exon 12 in humans with a product size of approximately 3.5 kb. The domain includes dipeptide repeats that contain serine-glycine residues and are separated by acidic and hydrophobic residues frequently (Krueger *et al.* 1990). CS chain recognition sites for attachment have been suggested to be S-G-X-G (Bourdon *et al.* 1987) or (D/E)-X-S-G (Krueger *et al.* 1990). There is some thought that these recognition sequence sites are dependent on chaperone proteins and the presence of specific enzymes that are necessary for post-translational modifications. The chief role of *ACAN* as a structural ECM proteoglycan is to hold water within the ECM. This is accomplished by the large number of negatively charged CS chains. The glycosaminoglycan (GAG) chains may play a vital role in *ACAN* processing because of their similar property characteristics of holding water within the ECM. This role in *ACAN* processing and secretion has been shown by the

addition of GAG chains to the CS region. Without these GAG modified CS sequences, CS containing constructs are not secreted (Kiani *et al.* 2001).

GAG chain modification is dependent on the G3 domain. This domain is highly complex involving alternative splicing of exons post-transcriptionally. In human *ACAN* it consists of two alternatively spliced EGF-like domains encoded by exons 13 and 14, a C-type Lectin-like domain also called LEC-like domain (also referred to as CRD module) encoded by exons 15, 16, and 17. An alternatively spliced CRP-like domain (also referred to as CBP module) is encoded by exon 18 (Doege *et al.* 1991; Fulop *et al.* 1993). In humans only about a quarter of the EGF-l like domain gets translated and a small portion contain the EGF-2 like domain or both (Fulop *et al.* 1993), due to variable alternative splicing. In rats, mice and dogs the EGF-1 domain is not translated due it being part of an intron, and in chickens some of the *ACAN* molecules have one or even none of the EGF domains encoded by exon 14. The C-type Lectin-like CRD are encoded by exons 14, 15, 16, 17 and the CBP/CRP domain in chickens encoded by exon 18 (Kiani *et al.* 2002). The function of the CRP/CBP domain is to bind fucose and galactose (Halberg *et al.* 1988) as well as the processing of *ACAN* proven by nanomelia in chickens and cmd mice. Without the G3 domain *ACAN* cannot be modified by GAG and this unmodified core protein cannot be secreted by the cell. It has been shown that in chicken *ACAN* the G3 domain's major functions involve the LEC-like/CRD domain encoded by exon 15. These animal models propose the attachment of GAG chains are modulated by G3 and this affects secretion from the cells (Kiani *et al.* 2001). G3 in overall functions of *ACAN* processing, is an important factor in the control of quality *ACAN* secretion and breakdown through degradation pathways dependent on ubiquitin-proteosome control. This role shows the significant role the G3 domain plays in GAG chain attachment and facilitating *ACAN* secretion. This is most notably controlled by the cysteine residues within the CRD/LEC-like and the CRP/CBP modules (Chen *et al.* 2002). The most interesting feature of the G3 inter-domain dependent regulation controls is that GAG modification is separate from product secretion in that the modification of the GAG matrix formation overrides its effect to the secretion of *ACAN* but not vice-versa (Kiani *et al.* 2002). The G3 domain as a whole could be involved in other aspects of *ACAN*

production, modification and degradation as seen with other proteoglycans like versican, however; no studies to date shown this to be the case.

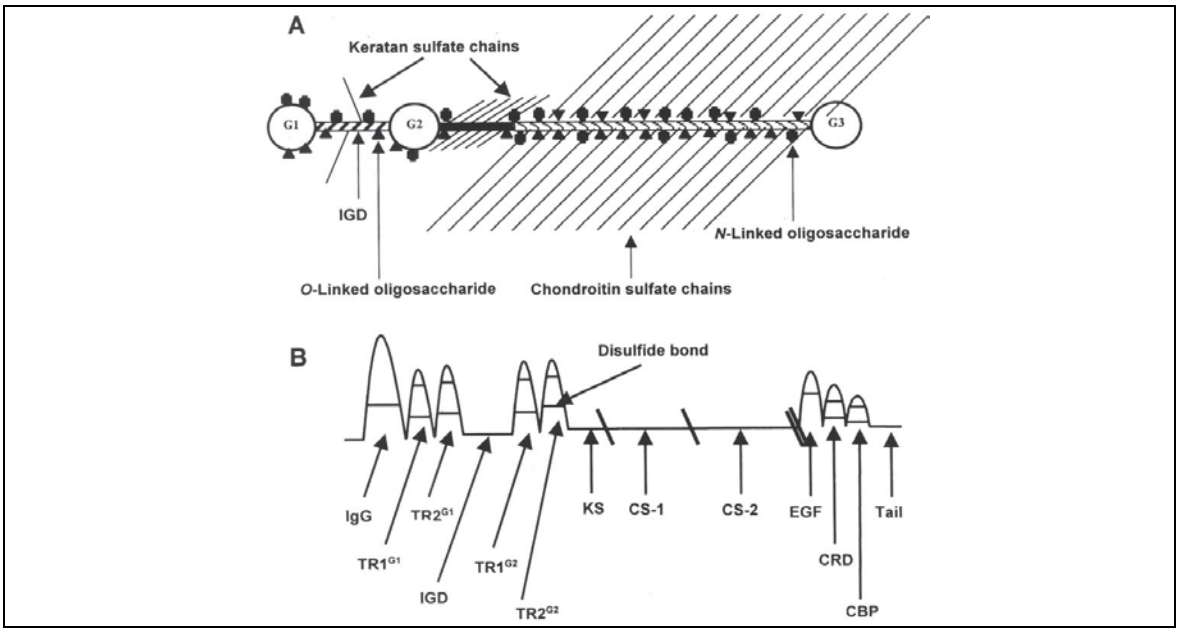


Figure 1.2. *Aggrecan* structure
(A) Globular protein and attached GAG chain structure. (B) Protein domain structure. Folded modules: IgG, immunoglobulin fold: TR, tandem repeats: EGF, epidermal growth factor-like module: CRD, carbohydrate recognition domain: CBP, complement binding protein-like module. Extended domains: IGD, interglobular domain: KS, keratan sulfate attachment domain: CS-1 and CS-2, chondroitin sulfate attachment domains (Kiani *et al.* 2002).

Purpose of Study

The purpose of the study described in this thesis was to uncover the cause of dwarfism among Miniature horses. At the beginning of the study it appeared that multiple genes with mutations might be responsible for the condition and candidate gene studies in families were conducted to uncover which of the more than 200 genes causing dwarfism among people were responsible for the condition in horses. However, the study revealed that most cases of dwarfism in Miniature horses were the consequence of one of at least 4 mutations in the gene, *Aggrecan* (*ACAN*).

Chapter 2

Disproportionate Dwarfism in Miniature Horses

Summary

This chapter focuses on describing the protocols utilized and results in the discovery of the mutations in *Aggrecan* (*ACAN*) which are responsible for the majority of dwarfisms within the Miniature horse breed. DNA samples were isolated from dwarfs and normal Miniature horses and select samples were sent to the Mayo Clinic for testing using the Illumina Equine SNP50 chip. The results were analyzed using PLINK which revealed a SNP with statistical significance near *ACAN* on chromosome 1 of the equine genome. Subsequent sequencing of *ACAN*indicated the existence of at least four separate recessive mutations that cause dwarfism. These mutations are independently segregating within the population worldwide. Novel mutation combinations occur causing combined expression in heterozygote dwarfs of two different mutations. Two of the mutations are deletions of a single base in exons 2 (*D1*) and 11(*D3*) causing a frameshift and presumed loss of function of the protein. The exon 2 (*D1*) deletion causes lethality in homozygous form and in combination with any of the other mutations. Another mutation is a deletion of 21 bases in exon 15 (*D4*) causing a loss of 7 amino acids in the resulting protein and loss of reading frame. The last mutation is a single base substitution in exon 6 (*D2*) causing changes in amino acid in the resulting protein. The exon 6 (*D2*) mutation predicts a decrease in function of the resulting protein due to the intricate binding interaction of link protein and G1 domain of *ACAN* to produce a functional ECM. During the sequencing process numerous non-causative SNPs were discovered within *ACAN* of the horse and are listed within this dissertation. Custom Taqman® Assay Probes were designed for mutations in exons 2, 6, and 11. A simple PCR of exon 15 and results run on a 2% agarose gel was the method used to test for the mutation in exon 15.Testing for dwarfisms *D1* (exon 2), *D2* (exon 6), and *D4*(exon 15) has been performed and completed on the entire sample population of Miniature horses and select large sized horse breeds. However, testing for *D3*(exon 11) dwarfism has not been completed.

Introduction

The main objectives of Miniature horse breeders are to produce small and proportionate horses. However, selection for small size could result in selection for dwarfism. With a foundation of immense genetic variation and no strict quality standards of bloodstock except height, some Miniature horse breeders utilized severely malformed small horses in breeding programs. These conformationally inferior horses exhibited consistent physical abnormalities typical of disproportionate dwarfisms as seen in other mammal species.

Other animal species with documented dwarfism include: dog, mouse, chicken, and cow. Defects in over 200 genes have been found to cause dwarfism in humans. Defects in genes for humans can involve genes responsible for pituitary gland hormones, thyroid gland hormones,hormone production and/or recognition, metabolism, cartilage development and bone growth plate development. During the course of the study it became apparent that the common forms of dwarfism in the Miniature horses fell into two groups. One group was characterized as Chondrodysplasia-like and were subject of this study. This type of dwarfism involves malformations of the cartilage and subsequent maturation to bone during fetal development and subsequent growth of the individual during sexual maturation. The other form of dwarfism seen in the Miniature horse population did not display Chondrodysplasia-like characteristics. However, these individuals did exhibit slightly enlarged sized head and shortened upper leg bones, but normal lower leg length and normal size of body and neck. This type of skeletal atavism is phenotypically similar to the dwarfism characterized in Friesians (Back *et al.* 2008). However, the cause of dwarfism among Friesian horses appears to affect only the long bones of the leg (skeletal atavism). These individuals were not subject to this study and appear to have a different genetic basis (Orr *et al*. 2010; Lindgren, personal communication 2013). The Miniature horse dwarfs displaying skeletal atavism are part of a study for this disease within the Miniature horse population (Rafiti *et al.* 2013). The dwarfism under study here causes disproportionate reduction of the entire individual. Disproportionate dwarfism involves the shortening and malformation of all the long

bones of the body and causes abnormal growth of the other bones within the body (i.e. skull). Skeletal atavism is a disease that involves only specific bones of the body being shortened and the rest of the bones are left relatively unaltered and normal.

Disproportionate dwarfism in Miniature horses, in decades past, has been a well-known unspoken issue. Its existence has been shown through anecdotal photo evidence and record keeping of breeders. These horses with conformation typical of disproportionate dwarfism have significant health, quality of life and soundness issues. Many Miniature horse registries try to avoid accepting minis affected by dwarfism for breeding stock registration (AMHA 2012; AMHR 2012). Unfortunately, these individuals are still used in breeding programs today. Dwarfism is a problem for the horse breeder who wishes to produce a healthy, diminutive horse which has proportions and conformation similar to that of large horse breeds. Therefore, this study was undertaken to uncover the gene(s) responsible for some types of dwarfism and develop diagnostic tests to guide breeders.

Materials and Methods

Horses used in Dwarfism Study

Two hundred AMHA registered Miniature horses from private farms were phenotyped for dwarfism and hair, blood or postmortem tissue collected for DNA isolation. Horses were selected based on a history of producing dwarfs or having a pedigree from known dwarf producers. Pedigree information, if available, was attained through the studbook of the AMHA registry. In addition, a total of 44 horses from the AMHR registry were tested but were of unknown phenotype for dwarfism. These samples were acquired from the University of Kentucky DNA typing Laboratory. Horses from other breeds which have not reported the occurrence of dwarf phenotypes were also tested which included 28 Thoroughbred, 22 American Standardbred, 4 American Saddlebred, 4 Tennessee Walking horses, 4 Arabian, 1 hackney pony and 1 Caspian horse.

Phenotyping

Miniature horses normally have a small stature. Dwarfism is distinguished by exhibition of several of the following characteristics: severely shortened stature, shortened limbs relative to overall body size, bowed forelegs, shortened neck, disproportionately large cranium, flat faces with large bulging eye sockets with prominent eyes, low nasal bridge, severe under bite, retruded muzzle, cleft palate and protruding tongue and a large abdominal hernia, or embryonic loss.

During the course of the study it became apparent that the common forms of dwarfism fell into two groups. One group was characterized as Chondrodysplasia-like and were subject of this study. This type of dwarfism involves malformations of the cartilage and subsequent bone during fetal development and growth of the individual. The other form of dwarfism seen in the Miniature horse population did not display Chondrodysplasia-like characteristics. However, these individuals did exhibit a slightly enlarged head and shortened upper leg bones, but normal lower leg length and normal size of body and neck.

DNA Isolation

DNA from blood or tissue was extracted using Puregene whole blood DNA extraction kits (Gentra Systems Inc., Minneapolis, MN) and Puregene tissue DNA extraction kits according to published protocols. Hair samples submitted were processed using 7-10 hair bulbs according to the method described by Locke *et al*. (2002). The hair bulbs were placed in 100 µl lysis solution containing 1X FastStart Taq Polymerase PCR buffer (Roche), 2.5 mM MgCl (Roche), 0.5% Tween 20 (JT Baker, Phillipsburg, NJ) and 0.01 mg proteinase K (Sigma-Aldrich, St. Louis, MO) and incubated at 60°C for 45 minutes followed by 95°C for 45 minutes to deactivate the proteinase K. Aliquots of the DNA samples were made for working dilutions at concentrations of approximately 150 ng/µl.

Illumina Equine SNP50 Testing

DNA samples at concentrations of 150 ng/µlfrom 46 horses were selected for testing with the Illumina Equine SNP50 chip. Of the 46 samples, 20 were from horses characterized as dwarves while the remainder were from Miniature horses with normal appearance. Among the 26 normal phenotype horses, all 26 were related to 6 of the affected horses as sibling or as parents. DNA samples were tested at the Mayo Clinic under the auspices of a Morris Animal Foundation Grant awarded to the University of Minnesota. The DNA samples were tested by the Mayo Clinic and results returned for analysis.

PLINK Analysis

PLINK (Purcell *et al.* 2007) was used for analysis of the data. The Illumina SNP50 chip assayed for 59,349 SNPs. The entire call rate was greater than 95% for all 46 horses so all 46 were retained for analyses. SNPs were eliminated from the analyses when the minor allele frequency (MAF) was less than 0.05 and the call rate for the SNP (GENO) was less than 90%. As a consequence, analyses included 40,368 SNPs. The data for dwarf horses was designated as cases and the data for normal horses was designated controls.Association study with a Monte Carlo based approach (5000 permutations) was run to compare differences in the occurrence of SNPs among the case and control horses. The permutation option (EMP2) corrected for the large number of comparisons being made. P-values less than 0.05 for EMP2 were considered significant and the region selected for further analyses. The SNPs from the region with statistical significance were further tested using the hap-phase option to identify haplotypes associated with the trait and to deduce the mode of inheritance.

GWAS

The only SNP that achieved statistical significance based on EMP2 score was BIEC2_38994 with a P- value of 0.003 found on ECA1. Below is Figure 2.1 identifying ECA 1 at SNP BIEC2_38994. See Appendix for supplementary Table 2.S1 SNPs with the highest EMP2 scores.

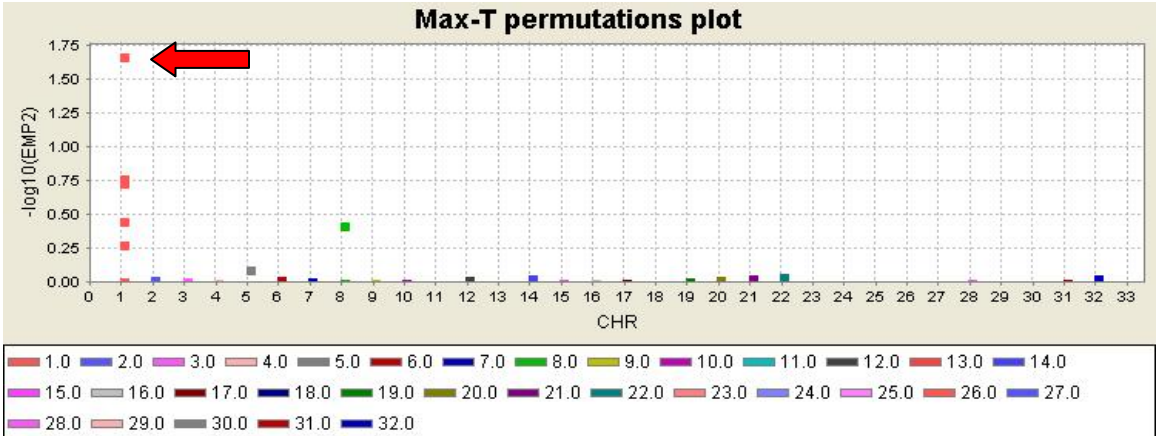


Figure 2.1. Graph displaying Max-T permutation quantification analysis showing statistical significance of SNP BIEC2_38994 on chromosome 1.

Haplotypes of individual dwarf samples within the GWAS study were compared using the Hap-Phase option within PLINK. These haplotypes associated with dwarfism showed at least 4 different haplotypes involved in the frequency differences between the controls and the affected horses. Furthermore, most of the dwarf individuals were heterozygous for these haplotypes. If a single recessive mutation were responsible for dwarfism, we would have expected to find homozygosity for a single haplotype among the affected horses. See Appendix for supplementary Table 2.S2 for haplotype association with dwarfs and controls.

The region possessing these statistically associated SNPs includes a gene-homologue known to cause dwarfism among human, mice, chicken and cattle, namely *ACAN*. In genome assembly Ecab 2.0, this gene was reported between 94,344,146 and 94,381,944; the associated SNP BIEC2_38994 is located at 94,215,030 on ECA 1. This is approximately 129,000 bp downstream of *ACAN*(XM_001917528.2) shown in Figure 2.2.

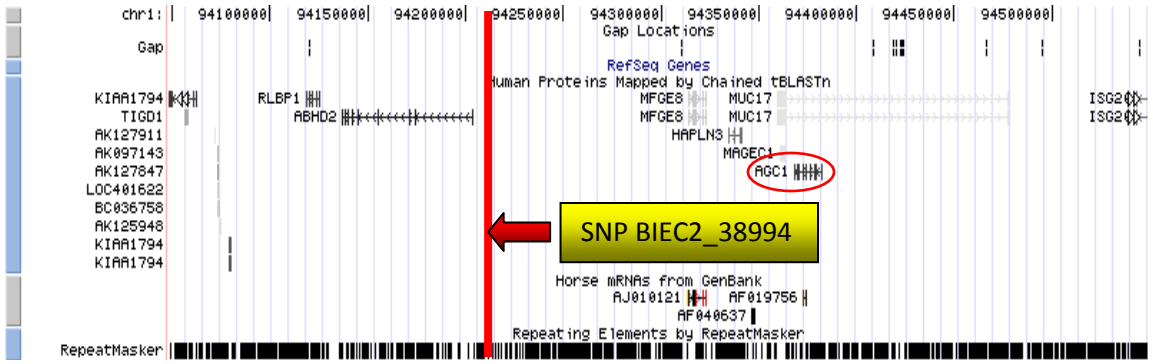


Figure 2.2. UCSC Genome Browser on Ecab 2.0 shows location of AGC1 (*ACAN*) in relation to SNP BIEC2_38994 129,000 bp downstream.

DNA Sequencing

A candidate gene, specifically *ACAN*, was identified near the region of highest statistical significance. Therefore, sequencing studies were begun to determine whether mutations in this gene might cause dwarfism among these horses. Evaluating predicted exon information from UCSC, Ensembl and NCBI, discrepancies were noticed in predictions of number and location of exons for equine *ACAN*. The various databases showed discrepancy with regard to exon 11 and 12 and an unexpressed yet *in silico* predicted exon in NCBI. Therefore to simplify, all exon descriptions will be based on NCBI count with regards to exon count shown in Figure 2.3.

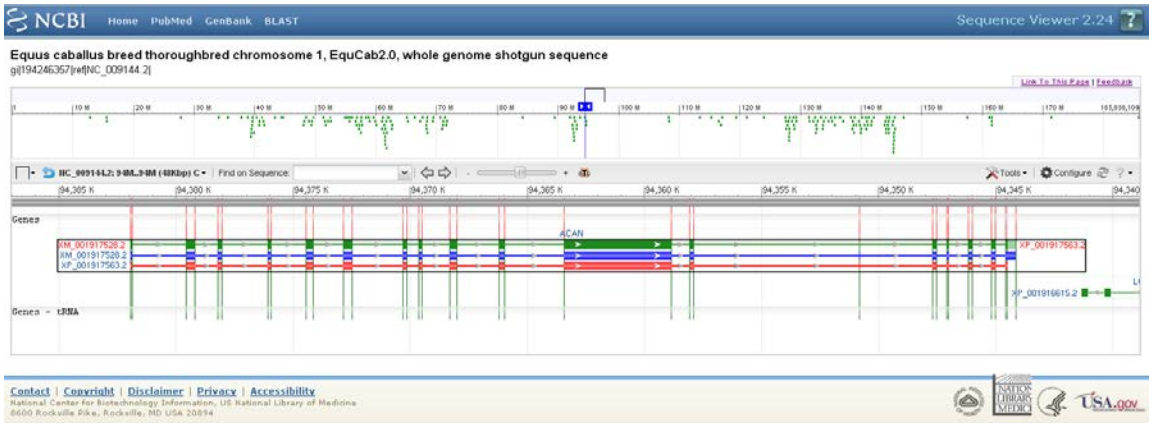


Figure 2.3. NCBI *In Silico* genome location, number and size prediction of exons for equine *ACAN*.

Once exon expression accuracy was determined specific exon primers were designed. Exons for *ACAN* were amplified by PCR (see Appendix for primers used for

exons in supplementary Table 2.S3). PCR Template for sequencing was amplified in 20 µl PCR reactions using 1X PCR buffer with 2.0 mM $MgCl_2$, 200 µM of each dNTP, 150 ng genomic DNA from hair, blood or tissue extractions, 0.2 µl FastStart Taq DNA polymerase (Perkin Elmer) and 50 nM of each primer. Each exon template PCR product was quantified on a 2% agarose gel run in 1X TBE solution and then amplified using the BigDye Terminator v1.1 cycle sequencing kit according to manufacturer's instructions (Applied Biosystems, Foster City, CA). Resulting sequence product was cleaned using Centri-Sep columns (Princeton Separations Inc., Adelphia, NJ), and read on an ABI 310 genetic analyzer (Applied Biosystems). At least six samples were initially sequenced for each exon: two dwarf samples, two parents of the dwarf samples and two non-dwarf producing samples. The non-producing samples were stallions with at least 100 registered foals with AMHA and no known dwarfs among the offspring. The results were analysed using Vector NTI Advance 10.3 software package using ContigExpress alignment program (Invitrogen Corporation, Carlsbad, CA). Exon sequencing identified SNPs and deletions described in this and the subsequent chapter.

Custom Taqman® Assay Probes

Three of the DNA variants identified associated with the occurrence of dwarfism, exon 2, exon 6, and exon 12 were chosen for custom Taqman® probe assays. To assay the presence of these variants among other normal and dwarf horses, including horses of other breeds, custom Taqman® SNP Genotyping Assays (Applied Biosystems) were designed for the three variants; exon 2 deletion c. 245(A/-) SNP and exon 6 c. 1270(G/A) c. 1290 (C/T) SNP; and exon 12 c. 6700 (C/-) deletion, in Filebuilder 3.1 software (Applied Biosystems) to test the population distribution of these three sets of independent variants. These assays were run on a 7500 HT Fast Real Time-PCR System (Applied Biosystems). The fourth variant, exon 15 with a 21 bp deletion, was simply amplified by PCR for the exon template and the PCR product was quantified on a 2% agarose gel to determine existence sequence size variant.

RESULTS

Genotype Variants

ACAN exons were sequenced and compared among dwarf and non-dwarf individuals. The sequence data for exonic regions was compared to the Reference sequence entry for equine *ACAN* (XM_001917528.2) using the reference *Equus caballus* 2.0 from NCBI whole genome shotgun sequence using sequence viewer 2.24. Initially, a SNP in exon 6 associated with homozygosity for dwarfism was found. When a significant number of the dwarf individuals appeared to be heterozygotes for this mutation,*ACAN* exons were sequenced on some of those dwarfs resulting in discovery of additional frameshift or non-synonymous mutations which may alter gene expression or function. Table 2.1 identifies 4 DNA variants identified as likely causes of dwarfism based on two criteria: 1) the DNA change was predicted to have a significant effect on the function of the protein, and 2) investigations of normal horses and horses with chondrodysplasia-like dwarfism showed a complete association between genotypes for these variants and theoccurrence of dwarfism.

Table 2.1. DNA Variants for Dwarfism. Position of mutation on NCBI *Equus caballus* build 2.0 with reference WT allele and mutant allele with resulting effect.

Allele	Effect	Position (build 2.0)	Reference allele	Mutant allele
Acan-N	normal			
Acan-D1	recessive dwarfism	94,379,389	A	-
Acan-D2	recessive dwarfism	94,372,649	G	A
Acan-D3	recessive dwarfism	94,358,998	C	-
Acan-D4	recessive dwarfism	94,346,122- 94,346,102	CGTGGTGATGATCTGGCACGA	-

Table 2.1(*continued*). DNA Variants for Dwarfism. Exon position of mutant allele cDNA base change and numbered position in gene and resulting numbered protein produced by mutant allele from WT.

Allele	Effect	Exon	cDNA XM_001917528.2	Protein
Acan-N	normal			
Acan-D1	recessive dwarfism	exon 2	c.245del A	p.K82fx
Acan-D2	recessive dwarfism	exon 6	c.1270G>A	p.V424M
Acan-D3	recessive dwarfism	exon 11	c.6700delC	p.P1875fx
Acan-D4	recessive dwarfism	exon 15	c.7299-7319del	p.F2433-O2440del

Allele designations for mutations involving *ACAN* were made as follows: *ACAN-N* wasused for the normal, non-dwarf form of *ACAN; ACAN- D1*, *D2*, *D3* and *D4* were

used for the four mutations identified and associated with dwarfism. Hereafter, the 5 alleles are referred to as *N, D1, D2, D3* and *D4*.

Four DNA variants were identified which appeared associated with unique dwarfism types seen in the breed. A single base deletion in NCBI exon 2 (*D1*), a single SNP substitution in NCBI exon 6 (*D2*), a single base deletion in NCBI exon 11 (*D3*) (UCSC exon 12) and a 21 base deletion in NCBI exon 15 (*D4*) (UCSC exon 15) were the four mutations found. In each case, the mutations would have resulted in a frame shift resulting in altered expression or loss of expression of *ACAN* or an amino acid substitution which may have altered function. In addition to the predicted effects of the mutation, we observed association of the mutations with the occurrence of the dwarf condition. Table 2.2 shows the genotypes with phenotype descriptions for the 5 *ACAN* alleles (*N, D1, D2, D3, and D4*).

Table 2.2. Genotype/phenotype descriptions. General descriptions of the five alleles of *ACAN* including novel combinations and resulting phenotype.

Genotype	Phenotype
N/-	normal, no dwarfism
D1/D1	Horses with this genotype were early gestational lethal dwarf
D1/D2	Horses with this genotype were early gestational lethal dwarf
D1/D3	Horses with this genotype were early gestational lethal dwarf
D1/D4	Horses with this genotype were early gestational lethal dwarf
D2/D2	Horses with this genotype were dwarf
D2/D3	Horses with this genotype were dwarf
D2/D4	Horses with this genotype were dwarf
D3/D3	No samples of this genotype, possibly due to sampling size
D3/D4	Horses with this genotype were dwarf
D4/D4	No samples of this genotype, possible early gestational lethal dwarf

Table 2.3 shows the genotype and associated phenotype count within the samples with regard to the presence or absence of dwarfism or fetal loss. The dwarfs with genotype *N/N* and *N/D2* were not the Chondrodysplasia type and may represent the form seen in Friesian horses with normal bodies and shortened legs referred to as skeletal atavism. The testing results of these dwarfs with skeletal atavism being heterozygous for one of the mutations within *ACAN* shows convincing evidence that skeletal atavism is not in linkage disequilibrium with any of the mutations found.Data shown in Table 2.3 illustrate the number of individual genotypes categorized within the population of

samples gathered. This table demonstrates the genotypes for normal and heterozygous normal carriers of the four dwarf alleles are all within the normal genotype group of the population sampled. No homozygous normal or heterozygous normal carrier genotyped individuals exist within any grouping of homozygous dwarf alleles nor in novel dwarf allele combinations.

Table 2.3. Genotype Association Studies.
* refers to individuals that display skeletal atavism type dwarfism
# refers to some individuals not typed yet for D3/D4 or N/D3 from the Sample population
** refers to no samples collected that are homozygous D3/D3 or D4/D4 from the sample population

		Physical Manifestations	
Genotypes	Term live dwarf	Nonviable/Aborted Dwarf	Normal
N/N	*1	0	66
N/D1	0	0	27
N/D2	*3	0	39
N/D3	#0	#0	#0
N/D4	0	0	2
D1/D1	0	1	0
D1/D2	0	9	0
D1/D3	0	2	0
D1/D4	0	1	0
D2/D2	17	1	0
D2/D3	11	1	0
D2/D4	5	0	0
D3/D3	**0	**0	0
D3/D4	#3	#0	0
D4/D4	**0	**0	0

Genotypes and Qualitative Assessment of Phenotypes

Physical characteristics of four unique and distinct dwarf phenotypes observed within the Miniature horse breed were categorized and characterized. Once proposed causative mutations were found, homozygous dwarf individuals were either physically examined or examined via multiple photos to assess common abnormal characteristics for each phenotype to match and confirm genotype. A population group of 200 AMHA Miniature horses were tested for evidence of heterozygosity of proposed mutations to confirm that normal carriers only possessed one copy of any given proposed mutation. If available, parents of dwarfs with homozygous results or novel combinations were tested

to confirm inheritance of specific mutations from each carrier parent to resulting dwarf foal. Miniature stallions with known offspring numbers statistically sufficient to qualify as a non-carrier were tested for lack of evidence of mutations for confirmation. In addition, a total of 44 Miniature horses from the AMHR registry were tested but were of unknown phenotype for dwarfism. The testing results of 44 random horses from the AMHR registry were of unknown phenotype for dwarfism. None of these tested homozygous for any of the proposed causative mutations or novel combinations thereof; however two tested as heterozygous carriers for the *D2* allele (*N/D2*). No other AMHR horses tested positive as carriers for any other dwarf allele. More testing needs to be performed in both AMHA and AMHR populations, however; from the limited data collected, AMHR horses have a much less incidence of carriers of dwarf alleles compared to the AMHA population.

A panel of other horse breeds was tested for the existence of any of the proposed causative mutations. Horses from other breeds which have not reported the occurrence of dwarf phenotypes included 28 Thoroughbred, 22 American Standardbred, 4 American Saddlebred, 4 Tennessee Walking horses, 4 Arabian, 1 hackney pony and 1 Caspian horse. These 64 large horse samples tested negative for dwarf alleles *D1*, *D2*, and *D4*, however; they have not been tested for allele *D3* as of this writing. Results are shown in the Appendix in Table 2.S4.

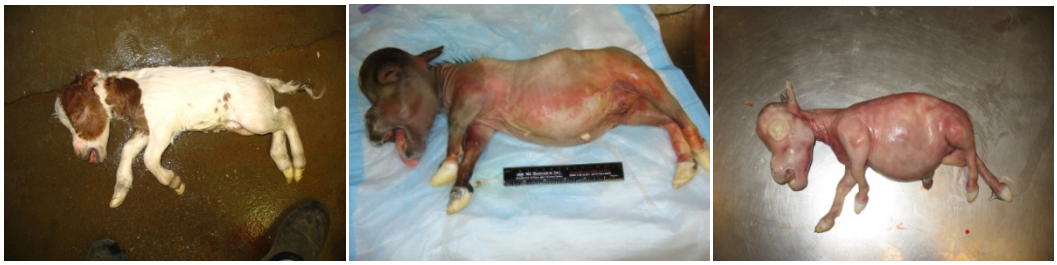
Genotypes with D1

Genotypes involving *D1* in combination with any other mutant allele appeared to be fetal lethals. Miniature dwarf genotype *D1/D1* is characterized as showing fetal abortion. Growth is extremely small stature, disproportionate dwarfism with a cranium large and domed for gestational age, extremely low nasal bridge, and extreme micrognathiawith severely retruded muzzle, cleft palate and protruding tongue. Characteristic ears set low and to the rear of cranium and large bulging eyes. Neck appears to be almost non-existent even in early gestation. Trunk involves a large abdominal hernia andextremely short vertebral column with ribs extremely short and thin. Limbs show severe micromeliaand appear to involve changes in limb bones that are

consistent with failure of endochondral ossification. All bones are poorly ossified for gestational stage. Figure 2.4 A, B, C show typical phenotype. When novel combinations of *D1* allele occur with other *ACAN* dwarf alleles, the resulting fetus is also lethal. Figure 2.5 A, B, C show novel heterozygous lethals.

A. **B.** **C.**

Figure 2.4 A, B, C. Fetus dwarfs homozygous for genotype *D1/D1*.



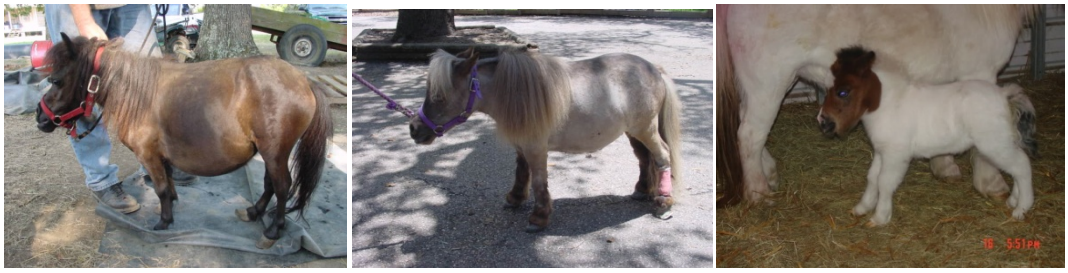
A. **B.** **C.**

Figure 2.5 A. Genotype *D1/D2*. **B.** Genotype *D1/D3*. **C.** Genotype *D1/D4*.

Genotypes with D2

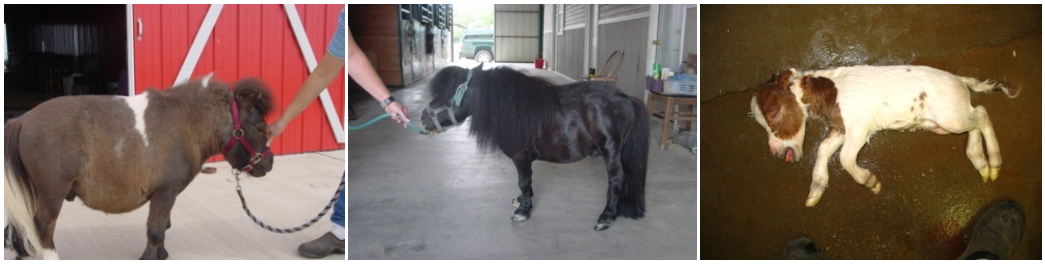
Horses inheriting the *D2* allele and *D2*, *D3* or *D4* were viable. Miniature dwarf genotype *D2/D2* is characterized with disproportionate short stature with short barrel shown in Figure 2.6 A, B, C. Craniofacial has a disproportionately large head with pronounced frontal bone domed prominence, enlarged eyes and orbit with a low and severely shortened nasal bone with retruded muzzle and under bite of varying severity. Airway obstruction develops and it is possible to be "central" in origin (due to foramen magnum compression) or "obstructive" in origin (due to narrowed nasal passages), necropsies were inconclusive. Symptoms of airway obstruction include snoring and laboured breathing. Limbs are short and thick comparatively to similar aged normal foal with various ranges of leg bowing due to joints. Joints are enlarged at birth with limited

joint mobility and flexation. Flexion contractures of hips, small pelvis size for overall size with rotation (goose rump). A pronounced shortened neck is noticeable at birth. The development of slight roach back and variable stiffness occurs with age. The barrel appears shortened due to slight curvature of spine. Motor milestones are often delayed secondary to joint deformities. Joint contractures and roach back progress with age. Progressive and painful joint enlargement can be accompanied by flexion contractures in all limb joints causing abnormal hoof growth very early. Muscle atrophy results from disuse. Novel combinations of *D2* allele with other *ACAN* dwarf alleles express both mutations' characteristics are shown in Figure 2.7 A, B, C.



A. **B.** **C.**

Figure 2.6 A, B, C. Dwarfs homozygous genotype *D2/D2*.



A. **B.** **C.**

Figure 2.7A. Genotype *D2/D3*. **B.** Genotype *D2/D4*. **C.** Genotype *D2/D1*

Genotypes with D3

Offspring with *D3* plus the *D2* (described above), *D3* and *D4*alleles were viable. The samples collected did not demonstrate any homozygous Miniature dwarf genotype *D3/D3*. This is possibly due to sample size of the population, however the number of samples found with this mutation was quite small and it is considered a rare mutation. The homozygote *D3*at this time is not considered a lethal due to novel combinations being viable, however; it is lethal with *D1*. A Dwarf with *D3/-* genotype is characterized

as disproportionate short stature with mature characteristics less severe than genotype *D2/D2*. Craniofacial structures are disproportionately large head with relative frontal bone slight domed prominence, enlarged eyes and orbit with a low and relatively shortened nasal bone, some show a retruded muzzle and generally have an under bite but not always. Limbs showrhizomelicshortening usually with bowed forelegs with enlarged joints and thick long bones at birth, limited joint flexion predominantly affecting the knee, stifle, pastern and hock joints initially. Progressive joint contractures of the pastern cause abnormal hoof growth with age. The development of a progressive roach back and hip rotation occur with a relatively shortened neck. Airway obstruction develops early and progresses with age. Muscle atrophy results from disuse. Novel combinations of *D3* allele with other *ACAN* dwarf mutations express both mutations' characteristics are shown in Figure 2.8 A, B, C.

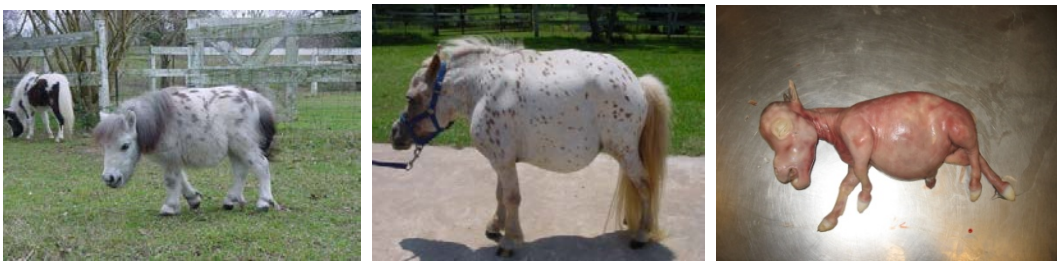
A. **B.** **C.**

Figure 2.8. A. Genotype *D3/D2* **B.** Genotype *D3/D4* **C.** Genotype *D3/D1*

Genotypes with D4

No samples of homozygous *D4* mutation were acquired. Only physical samples acquired were of novel combinations of other abnormal genotypes expressing both physical characteristics of each mutation. When *D4* allele was present with *D1* allele genotype, the results are lethal. Growth wasdisproportionate, exhibitingsevere short stature with a short barrel misshapen and more pronounced than *D2/D2* and *D3/D3*. Roaching of spine is more pronounced and severity progresses with age. Barrel appears even more shortened due to curvature of spine. Variable pain and stiffness of hip and limbs are progressive. Craniofacial structures are disproportionately large head with pronounced frontal bone domed prominence, enlarged eyes and orbit with a low and shortened nasal bone with some having a retruded muzzle and under bite with varying severity. Airway obstruction is evident early and progressive. Limbs are rhizomelic in

length with bowing and joints are enlarged at birth with limited joint mobility. Flexion contractures of hips, small pelvis with severe rotation (goose rump).Severely shortened neck is noticeable from birth that is almost nonexistent. Motor milestones are often delayed secondary to joint deformities. Joint contractures and roach back progress quickly with age. Progressive and painful joint enlargement can be accompanied by flexion contractures in all limb joints causing abnormal hoof growth very early. Muscle atrophy results from disuse some examples are shown in Figure 2.9 A, B, C.



A. **B.** **C.**

Figure 2.9 A. Genotype *D4/D2*. **B.**Genotype *D4/D3*. **C.**Genotype *D4/D1*

Discussion

Dwarfism within the Miniature horse breed has had various descriptions and terminology anecdotally used to characterize the various phenotypes observed. Many medical terms used have been incorrectly or only partially utilized in describing specific characteristics of an observed phenotype. When it was determined that *ACAN* was the gene involved in the dwarf types in this study, a concise yet encompassing descriptive term for these mutations was needed. Other dwarfisms involving *ACAN* of other mammal species were referenced when deciding to describe the complexity of this disease within the Miniature horse. Once it was discovered that multiple different mutations are independently segregating within the Miniature horse population and that these mutations produced specific distinct phenotypes in homozygotes, an inclusive term was necessary as a general descriptive term. Work in Dexter cattle dwarfism revealed multiple different mutations within *ACAN* that definitively showed causation of that disease (Cavanagh *et al.* 2007). The homozygous dwarf phenotype of Chondrodysplasia in the Dexter is similar to the Type *D1/D1* seen in the Miniature horseexpressing

identical gross anatomical deformities as shown if Figure 2.10 A and 2.10 B. In both the Dexter and the Miniature horse, the affected individuals of this phenotype are lethal and involve the same gene *ACAN*. Chondrodysplasia in general is a term used to describe abnormal growth of cartilage. It causes disproportionate dwarfism and occurs as an inherited trait. These mutations found within *ACAN* of the Miniature horse are Chondrodysplasia-like deformities of the cartilage causing subsequent abnormal growth and maturation of bone. Therefore these four dwarf types observed to occur in an inherited form in the Miniature horse are designated Chondrodysplasia-like dwarfisms.

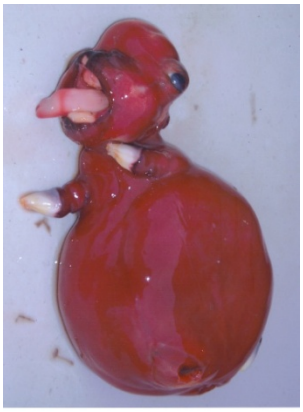


Figure 2.10.A. Miniature Lethal *D1/D1*

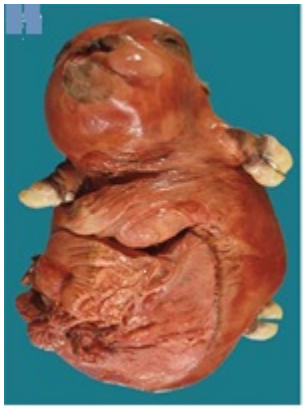


Figure 2.10.B. Dexter lethal chondrodysplasia dwarf (Cavanagh *et al*. 2007)

Skeletal atavism is another distinct type of dwarfism in the Miniature horse that had known to exist in Welsh and Shetlands previously (Tyson *et al.* 2004), see Figure 2.11A. However, pathology of skeletal atavism is much different than the Chondrodysplasia-like dwarfism in this study. Skeletal atavism is abnormal bone growth specifically restricted to the upper limb bones of the Miniature horse. This type is quite similar in phenotype to Osteochondrodysplasia seen in Friesians(Bach *et al.* 2008), see Figure 2.11B. It is noted in the results that Miniature dwarfs of this type (skeletal atavism) were tested for the presence of any of the causative mutations within *ACAN*. The dwarfs with genotype *N/N* and *N/D2* (Table 2.3) were of the skeletal atavism type. The testing results of these dwarfs with skeletal atavism being heterozygous for one of the mutations within *ACAN* shows convincing evidence that skeletal atavism is not in linkage disequilibrium with any of the mutations found.

Figure 2.11A. Miniature with Skeletal Atavism

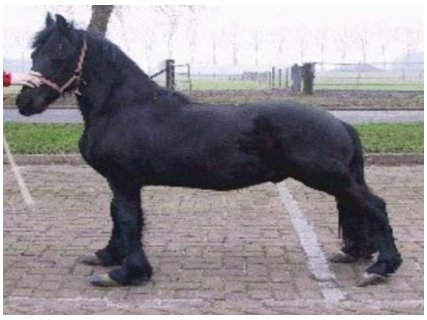


Figure 2.11B. Friesian dwarf exhibiting Osteochodrodysplasia. (Back *et al*. 2008)

Testing results showed that three skeletal atavism dwarfs were heterozygous carriers of *D2*. No skeletal atavism dwarfs were homozygous for any of the proposed causative mutations of Equine Chondrodysplasia-like dwarfism in the Miniature horses. This gives convincing evidence these two types of dwarfisms are caused by two separate and independently segregating genes in the population. The Miniature dwarfs displaying skeletal atavism are part of a study for this disease within the Miniature horse population (Rafiti *et al.* 2013).

Sequence databases used for reference of Equine *ACAN* revealed discrepancies with regards to predicted exon number and location. UCSC genome browser predicted an intron in exon 11 that included a major portion of the ~2400 bp repeat region, thus separating this sequence region into two exons (11 and 12) with the repeat region as the intron in between the two exons. NCBI sequence viewer however, predicted exon 11 to be the entirety of UCSC exon 11, including the intron repeat region and exon 12. NCBI also predicted another exon downstream of exon 11 as the next exon (NCBI exon 12). The UCSC genome browser did not predict NCBI exon 12 to exist as an exon and predicted this region to be an intron. Expression data showed that NCBI correctly predicted exon 11 to be one exon that included the repeat region (UCSC exon 11, intron region and UCSC exon 12). However, NCBI incorrectly predicted an exon 12 (exon does not exist in UCSC genome browser) that is not expressed. The location of this NCBI predicted exon in the UCSC genome browser is in intron 12. Therefore from this information, it is possible that the number of exons expressed is variable with splicing occurring during processing within the equine *ACAN*, as what has been shown to occur in humans and other mammal species (Valhmu *et al.* 1995). The current count of expressed

exons in the genome databases is to exon 11 using NCBI predicted equine *ACAN*. After which, the next exon expressed is UCSC exon 13 through exon 17. NCBI denotes the same as UCSC with regards to exons 13 through 17.

No studies in pathology have been conducted at this time. Gene expression data at the cellular level and protein function data of *ACAN* within the ECM involving these proposed mutations were not examined. Predicted loss of function is evident with deletions in exons 2, 12, and 15, for genotypes *D1*, *D3* and *D4*respectively, causing subsequent frameshifts resulting in novel protein products. Other published mammal mutations of *ACAN* involving deletions in various exons provide ample evidence of cause and effect (Watanabe *et al.*1997; Cavanagh *et al.*2007) as well as the sequence data from the Miniature horse samples with these mutations. No Miniature horse samples with normal phenotypes were homozygous for any of these deletions, nor were there any normal phenotypes with genotypes of novel deletion combinations.

The mutation in exon 6 involving a single base missense producing a change of amino acid valine to methionine causes the addition of a sulfur group. The compelling argument for this to be causative is that no samples in the Miniature horse population tested possessed this *D2* allele in homozygous form and a normal phenotype. All novel combinations of *D2* genotype with other dwarf genotypes (*D1, D2, and D4*) produced dwarf phenotypes. All parents of *D2/D2*genotype dwarf samples matched as heterozygous carriers with the causative missense sequence for this genotype. The published information regarding function of this exon is that link protein and its exons are complimentary to *ACAN* in structure and binding function at the exon sequence level. As mentioned, the G1 domain of *ACAN* is divided into loop sections of A, B, B' and correspond to subsequent exons 3 (A), 4 – 5 (B), and 6(B'), respectively. Link protein exons 3-6 correspond to the G1 loop sections in sequence similarity confirming binding site affinities (Valhmu *et al.* 1995). There is no information regarding mutations in *ACAN* G1 or Link protein in humans, inferring the importance of the aggregate formation of *ACAN* and the ECM in skeletal development (Aspberg *et al*. 2012). The existence of a valine to methionine missense mutation in the C-type lectin domain (CLD) of the G3

domain of *ACAN* causing a type of dwarfism in humans (Stattin *et al*. 2010), leads to the possibility of a similar situation with the valine to methionine mutation in exon 6 of the Miniature horse. This information provides evidence that this missense likely causes incomplete or loss of binding ability between*ACAN* and link protein to properly form the ECM. The sequence data of Miniature horse samples also shows high conservation of exons 3, 4, 5 and 6 with complete conservation in exons 3 or 4, one synonymous SNP in exon 5, and exon 6 containing two non-causative SNPs changing amino acid sequence and one synonymous SNP. This confirms known sequence data across species of high sequence conservation in this region (Valhmu *et al.* 1995).

Chapter 3

Polymorphisms of *ACAN* Not Found to Cause Dwarfism

Summary

Single nucleotide polymorphisms, referred to as SNPs, are the result of mutations that occur through genetic evolution of a species within its environment. These mutations sometimes produce a change in a resulting protein that makes it advantageous to the species. Other times the mutation may be deleterious and the resulting protein has a partial or complete loss of function. However, there are also times synonymous mutations occur resulting in no change of the resulting protein. These mutations may or may not alter the resulting phenotype of an individual. Polymorphisms allow a species to genetically maintain variability in its environment. This presumably gives the species a better chance of survival if that environment were to change.

During the course of investigating dwarfism in the Miniature horse and sequencing the gene responsible, *ACAN,* numerous SNPs were discovered that did not appear annotated in the reference sequence noted in this chapter. No SNPs causing a change in amino acid sequence exhibited linkage disequilibrium with any known dwarf phenotype investigated. However, some synonymous SNPs displayed linkage disequilibrium with specific dwarf phenotypes. Various numbers of DNA samples were sequenced for each exon to initially look for mutations causing dwarfism. When numerous SNPs were noted, certain specific individual samples were sequenced through most, if not all, of the exons. Some of these samples were dwarfs already determined homozygous for a specific mutation, and other samples were stallions that were considered statistically normal for all dwarf types due to the number of registered offspring and never producing a dwarf. However, not all samples acquired were sequenced entirely or at all and were only tested for genotype status of mutations found. Therefore, these are proposed non-causative SNPs of any known disease at this time within *ACAN* of the horse.

Introduction

During the sequencing of equine *ACAN*, the standard sequence used in build 2.0 as the reference sequence is from a thoroughbred, Twilight. When using the reference sequence, it was noted the existence of two previously annotated SNPs in the database, one in exon 6 (c.1249C→ G) and the other in exon NCBI 11 (c.6592 A→ G) which is referred to as exon 12 in UCSC. Other SNP variations throughout the gene were discovered within the sample group of Miniature horses, some were synonymous and some caused a non-pathogenic amino acid change. These SNPs were variations to the reference sequence and were confirmed through repeated sequencing and trace value call scores above 400 using Vector NTI ContigExpress program. There are quite possibly additional SNPs within this gene. This is due to our inability to sequence entirely through the ~2400 bp highly conserved repeat region of exon 11.

Materials and Methods

Horses Used in Aggrecan SNP Study

Two Hundred AMHA registered Miniature horses from private farms were phenotyped for the dwarfism study and hair, blood or postmortem tissue collected for DNA isolation. Horses were selected for the SNP study were ones with a history of producing dwarfs or had a pedigree from known dwarf producers. Pedigree information, if available, was attained through the studbook of the AMHA registry. Stallions that had numbers of offspring statistically significant to consider normal for all dwarf types were used to compare SNPs found that were potentially non-causative. Normal horses genotyped with one copy of a dwarf mutation were also used to compare SNPs. Homozygous dwarfs for types *D1* and *D2* were used as well for comparison of SNPs found, as well as specific novel combination dwarf samples with *D3/D4*, *D1/D3*, *D1/D4*, *D2/D3* and *D2/D4* genotypes.

DNA Sequencing

The candidate gene, *ACAN*, was the gene utilized in this study. During sequencing studies to determine whether mutations in this gene might cause dwarfism among these horses, numerous possibly non-causative SNPs were located. Exons for *ACAN* were amplified by PCR (see Appendix for primers used for exons in supplementary Table 2.S3). PCR template for sequencing was amplified in 20 µl PCR reactions using 1X PCR buffer with 2.0 mM $MgCl_2$, 200 µM of each dNTP, 1 µl genomic DNA from hair, blood or tissue extractions, 0.2 µl FastStart Taq DNA polymerase (Perkin Elmer) and 50 nM of each primer. Each exon template PCR product was quantified on a 2% agarose gel run in 1X TBE solution and then amplified using the BigDye Terminator v1.1 cycle sequencing kit according to manufacturer's instructions (Applied Biosystems, Foster City, CA). Resulting sequence product was cleaned using Centri-Sep columns (Princeton Separations Inc., Adelphia, NJ), and read on an ABI 310 genetic analyzer(Applied Biosystems). The sequence data for exonic regions was compared to the reference sequence entry for equine *ACAN* (XM_001917528.2) using the reference *Equuscaballus* 2.0 from NCBI whole genome shotgun sequence using sequence viewer 2.24. Exon sequenceswere aligned and analysed using Vector NTI Advance 10.3 software package using ContigExpress program (Invitrogen Corporation, Carlsbad, CA).

Genotype Variants

ACAN exons were sequenced and compared among dwarf and non-dwarf individuals. Initially, at least six samples were sequenced for each exon. Two dwarf samples, two parents of the dwarf samples and two non-dwarf producing samples. The non-producing samples were stallions with at least 100 registered foals with AMHA and producing no known dwarfs. Once the proposed non-causative SNPs were identified, other dwarf samples were sequenced that were known to be homozygous for mutations *D1* and *D2*, as well as normal carriers of those mutations, and novel dwarf combination genotypes. At least two carriers of each dwarf mutation were also sequenced when possible. This was to confirm the SNP was non-causative as well as determine if any were in linkage disequilibrium with any genotype, including normal.

Results

The SNPs identified were variations to the reference sequence and were confirmed through repeated sequencing and trace value call scores above 400 using Vector NTI Advance 10.3 software packageContigExpress alignment program (Invitrogen Corporation, Carlsbad, CA). Each SNP was categorized according to NCBI gene base pair number in Microsoft Excel spreadsheets with each sample sequenced and corresponding genotype for that sample. Table 3.1 lists 24 proposed SNP variants uncovered within *ACAN* that have been found to not cause a dwarf phenotype.

Table 3.1. Proposed non-causative SNPs found in *ACAN*. Shows position in genome, reference WT allele, mutant allele, exon location of SNP, cDNA base number of gene, amino acid number in gene and alteration.

Allele	Effect	Chr	Position	Reference	Mutant	Exon	cDNA	Protein
ACAN-N	normal	1	94,381,920	G	T	exon 1	c.25G>T	p.V9L
ACAN-N	normal	1	94,379,421	A	G	exon 2	c.213A>G	p.P71P
ACAN-N	normal	1	94,374,328	T	C	exon 5	c.867T>C	p.A289A
ACAN-N	normal	1	94,372,670	C	G	exon 6	c.1249C>G	p.P417A
ACAN-D2	dwarfism	1	94,372,629	C	T	exon 6	c.1290C>T	p.P430P
ACAN-N	normal	1	94,372,587	A	G	exon 6	c.1333A>G	p.I445V
ACAN-N	normal	1	94,370,258	G	C	exon 7	c.1513G>C	p.A505P
ACAN-N	normal	1	94,370,248	A	G	exon 7	c.1523A>G	p.A508P
ACAN-N	normal	1	94,368,175	G	C	exon 9	c.1888G>C	p.A630P
ACAN-N	normal	1	94,363,292	G	A	exon 11	c.2406G>A	p.T802T
ACAN-N	normal	1	94,363,085	G	A	exon 11	c.2613G>A	p.G871G
ACAN-N	normal	1	94,362,996	C	T	exon 11	c.2702C>T	p.A901A
ACAN-N	normal	1	94,362,939	C	A	exon 11	c.2759C>A	p.S920T
ACAN-N	normal	1	94,360,319	C	T	exon 11	c.5379C>T	p.S1793S
ACAN-N	normal	1	94,360,074	C	T	exon 11	c.5624C>T	p.A1875V
ACAN-N	normal	1	94,359,923	C	T	exon 11	c.5775C>T	p.Y1925Y
ACAN-N	normal	1	94,359,812	G	A	exon 11	c.5886G>A	p.S1962S
ACAN-N	normal	1	94,359,062	G	T	exon 11	c.6264G>T	p.G2088G
ACAN-N	normal	1	94,359,062	C	T	exon 11	c.6275C>T	p.A2092V
ACAN-N	normal	1	94,359,106	A	G	exon 11	c.6592A>G	p.R2198G
ACAN-N	normal	1	94,359,062	G	A	exon 11	c.6636G>A	p.S2212S
ACAN-N	normal	1	94,347,621	G	A	exon 13	c.7102G>A	p.A2368T
ACAN-N	normal	1	94,345,187	G	C	exon 16	c.7412G>C	p.C2471S
ACAN-N	normal	1	94,345,096	C	T	exon 16	c.7503C>T	p.T2501T

Table 3.2 through Table 3.5 list specific samples sequenced and the resulting genotype at SNP locations as described in the following subsections for each exon. Not all samples were sequenced for all exons therefore if a space is blank in Tables 3.2 – 3.5 it signifies that sample not having been sequenced for that specific exon.

Exon 1 Non-Causative SNP

In exon 1 a single SNP at c.25G→T involved a change of amino acid 9 from valine to leucine which showed ubiquity among all genotypes within the population. This occurrence is non-consequential due to this exon not being translated into a part of the functioning protein.

Exon 2 Non-Causative SNP

Exon 2 a synonymous SNP occurs at c.213A→G producing no change of amino acid. This SNP is located upstream of the pathogenic mutation in exon 2 at c.245 and shows no linkage to that deletion; however it appears to show linkage disequilibrium with samples without the genotype *D1*. Of the 59 samples sequenced for this exon, only 8 samples possessed at least one copy of this SNP. No sample with a genotype of *D1* possessed this base change. More samples of the Miniature population would need to be analyzed to verify if this SNP is in true linkage disequilibrium with the WT genotype.

Exon 5 Non-Causative SNP

Exon 5 a synonymous SNP occurs at c.867T→C producing no amino acid change. Of the 18 samples sequenced from the population there was no linkage to any of the genotypes listed previously for either base T or C.

Exon 6 Non-Causative SNPs

In exon 6 the previously annotated SNP at c.1249C→G produces the same base change within the Miniature horse population. Of the 77 samples sequenced, this SNP shows ubiquity among all genotypes within the sample population.The SNP in exon 6 at c.1290C→T is a synonymous SNP that among the 77 samples shows linkage disequilibrium with T segregating completely with *D2* genotype and the causative

mutation at c.1270G→A and wildtype C at c.1290C→T segregates completely with non-carriers of *D2* genotype. This SNP could be used with the causative mutation for confirmation of genotype. The SNP in exon 6 at c.1333A→G produces an amino acid change of Isoleucine to Valine at position 445. This SNP change to G shows linkage disequilibrium with all samples not possessing a genotype of *D3*. All samples with the genotype *D3* show linkage disequilibrium with the wildtype SNP A.

Exon 7 Non-Causative SNP

In exon 7, at base pair location c.1513G→C, change resulting in amino acid alanine to proline. Also, at base pair location c.1523A→G, change resulting in amino acid alanine to proline. This amino acid change possibly may cause a conformational change in the inter-globular domain (IGD) which is the region for aggrecanase - 1 cleavage activity (Sandy *et al.* 1992). The IGD also contains regions for proteolytic cleavage sitesvulnerable to a range of proteinases. One of which is matrix metalloproteinases (MMPs). The IGD also contains sites for serine proteinases some examples areleukocyte elastase and plasmin. And finally, the IGD contains sites for acid proteinasessuch as cathepsin B (cysteine protease) (Hardingham *et al.* 1995: Mort *et al.* 1997; Mort *et al.* 1998). The non-causative snips in exon 7 are located at residues Ala 505 to Pro 505 and Ala 508 to Pro 508. It is not known if these changes interfere with the two key cleavage sites identified inthe IGD domain which are between residues Asn 341 and Phe 342 and betweenresidues Glu 373 and Ala 374 (Singer *et al.* 1995; Fosang *et al.* 1996; Lark *et al.* 1997).

Exon 9 Non-Causative SNP

Exon 9 contains a single SNP at location c.1888G→C causing an amino acid change p.A630P alanine to proline. The introduction of Pro 630 may produce a conformational change in the tertiary structure formation of *ACAN* of the G2 domain in this region due to the change in amino acid from residue Ala 630.

Exon 11 Non-Causative SNPs

Exon 11 (NCBI) involves a rather large highly conserved repeat (~2400 bp) that was not completely sequenced. The first part of the exon prior to the repeat region, there were found three synonymous SNPs at c.2406G→A from 19 samples sequenced, c.2613G→A from 22 samples sequenced, c.2702C→T from 24 samples sequenced. The next SNP, at location c.2759C→A,results in amino acid change p.S920T (serine to threonine). This SNP does not show linkage disequilibrium with any genotype. SNP c.2759 is located in the beginning of the repeat region of exon 11. The repeat region encodes for the keratan sulfate (KS) chains of the *ACAN* protein and has shown variability in sequence length in humans (Doege *et al.* 1997). Towards the latter part of the repeat region, a SNP at location c.5379 C→T is synonymous. At location c.5624 C→T, a SNP produces an amino acid change p. A1875V (alanine to valine). This change shows no linkage to any genotype of the 20 samples sequenced. SNPs at location c.5775C→T, 31 samples genotyped; c.5886G→A, 23 samples genotyped; c.6264G→T, 25 samples genotyped; are all synonymous SNPs showed no linkage to a specific genotype. For SNP c.6275C→T, it produces and alanine to valine amino acid change. Linkage disequilibrium does not appear to exist with the 28 samples sequenced. The SNP at location c.6592 A→G shows no linkage disequilibrium within the 21 samples sequenced. This SNP of (p.R→G) arginine to glycine amino acid change has been previously noted on the UCSC genome browser as a known SNP. The SNP at location c.6636G→A is synonymous and more samples need to be sequenced to determine if this SNP is in linkage disequilibrium with a dwarf genotype/phenotype.

Exon 13 Non-Causative SNP

Exon 13 shows a SNP located at c.7102G→A results in an amino acid change of p.A2368T. Of the 16 samples, there was no linkage disequilibrium with any genotype.

Exon 16 Non-Causative SNPs

Exon 16 shows two SNPs at location c.7412G→C producing an amino acid change of p.C2471S and c.7503C→T that is synonymous.

Table 3.2. Proposed non-causative SNPs found in *ACAN* for exons 1-6. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 1	exon 2	exon 5	exon 6	exon 6	exon 6
Allele	sample	94,381,920	94,379,421	94,374,328	94,372,670	94,372,629	94,372,587
		c.25G>T	c.213A>G	c.867T>C	c.1249C>G	c.1290C>T	c.1333A>G
		p.V9L	p.P71P	p.A289A	p.P417A	p.P430P	p.I445V
	2698	G/G	A/A	T/T	G/G	C/C	G/G
D2/D2	2705	G/G	G/G	C/C	C/C	T/T	A/A
D2/D2	2700	G/G	G/G	C/C	C/C	T/T	A/A
D2/D2	2970	G/G	G/G				
N/D2	2971	G/G	G/G				
N/D1	2756	G/G	G/G	C/C	C/C	C/C	G/G
N/N	2758	G/G	G/G	C/C	G/G	C/C	G/G
D4/D	2676	G/T	G/G	C/T			
N/D4	2675	G/T	G/G				
N/D	2677	G/T	G/G				
N/D2	2972	G/T	G/G				
N/D1	2757	G/T	G/G		C/C	C/C	G/G
N/D1	2516	G/T	G/G	C/T	C/G	C/C	G/G
N/N	2726	G/T	G/G	T/T	G/G	C/C	G/G
N/D1	2737	G/T	G/G		C/G	C/C	G/G
D1/D	2683	T/T	G/G		C/G	C/C	G/G
N/D	2736	T/T	G/G		C/G	C/C	G/G
D1/D	2703	T/T			C/G	C/C	G/G
N/N	2727	T/T	G/G	C/T	C/G	C/C	G/G
N/D2	2723		G/G		C/G	C/T	A/G
N/D2	2847		G/G				
D2/D	2674		G/G		C/G	C/T	A/G
D2/D	2673		G/G		C/G	C/T	A/G
N/D2	2724		G/G		C/G	C/T	A/G
D2/D2	2748		G/G				
D2/D	2746		G/G		C/G	C/T	A/G

Table 3.2(continued). Proposed non-causative SNPs found in *ACAN* for exons 1-6. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 1	exon 2	exon 5	exon 6	exon 6	exon 6
Allele	sample	94,381,920	94,379,421	94,374,328	94,372,670	94,372,629	94,372,587
		c.25G>T	c.213A>G	c.867T>C	c.1249C>G	c.1290C>T	c.1333A>G
		p.V9L	p.P71P	p.A289A	p.P417A	p.P430P	p.I445V
D2/D2	2789		G/G		C/C	T/T	A/A
D4/D	2745		G/G	C/C	C/C	C/C	G/G
D1/D1	2738		G/G				
D2/D	2749		G/G		C/G	C/T	A/G
D2/D4	2744		G/G				
N/D2	2718		G/G				
N/D2	2714		G/G				
D2/D	2701		G/G		C/G	C/T	A/G
D1/D	2747		G/G		C/C	C/T	A/G
D1/D2	2849		G/G		C/C	C/T	A/G
D1/D	2683		G/G				
D1/D2	2704		G/G		C/C	C/T	A/G
N/D1	2513		G/G		C/C	C/C	G/G
N/D1	2515		G/G				
D2/D2	2715		G/G				
N/D2	2716		G/G				
D1/D2	2678		G/G		C/C	C/T	A/G
D1/D2	2668		G/G				
N/D2	2718		G/G				
D1/D4	2896		G/G		C/C	C/C	G/G
D1/D2	2719		G/G				
D2/D2	2706		G/G		C/C	T/T	A/A
N/D2	2865		G/G				
D1/D2	2711		G/G				
N/D2	2792		G/G				
N/D1	2866		G/G				
D2/D2	2709		G/G	C/C			
N/N	2823		G/A				

Table 3.2(continued). Proposed non-causative SNPs found in *ACAN* for exons 1-6. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 1	exon 2	exon 5	exon 6	exon 6	exon 6
Allele	sample	94,381,920	94,379,421	94,374,328	94,372,670	94,372,629	94,372,587
		c.25G>T	c.213A>G	c.867T>C	c.1249C>G	c.1290C>T	c.1333A>G
		p.V9L	p.P71P	p.A289A	p.P417A	p.P430P	p.I445V
D2/D	2906		G/A				
D3/D4	2741		G/A	C/T			
N/D3	2742		A/A	T/T	G/G	C/C	G/G
D2/D4	2694				C/C	C/T	A/G
D2/D4	2696				C/C	C/T	A/G
D2/D3	2702				C/C	C/T	A/G
D2/D3	2710				C/C	C/T	A/G
D1/D2	2719				C/C	C/T	A/G
D2/D4	2751				C/C	C/T	A/G
N/D2	2764				C/C	C/T	A/G
D2/D2	2752				C/C	T/T	A/A
N/D1	2721				C/C	C/C	
N/D1	2514				C/C	C/C	G/G
N/D2	2669				C/G	C/T	A/G
D2/D	2695				C/G	C/T	A/G
D2/D2	2707				C/G	C/T	A/G
D2/D2	2708				C/G	T/T	A/A
N/D2	2750				C/G	C/T	A/G
D2/D	2911				C/G	C/T	A/G
N/D2	2860				C/G	C/T	A/G
N/D1	2880				C/G	C/C	G/G
N/D2	2722				C/G	C/T	A/G
N/D2	2518				C/G	C/T	A/G
N/D1	2512				C/G	C/C	G/G
D2/D2	2699				C/C	T/T	A/A
N/D2	2524				C/C	C/T	A/G
N/D1	2514				C/C	C/C	G/G
N/D4	2672						
D2/D	2697						

Table 3.3. Proposed non-causative SNPs found in *ACAN* for exons 7-11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 7	exon 7	exon 9	exon 11	exon 11	exon 11
Alleles	sample	94,370,258	94,370,248	94,368,175	94,363,292	94,363,085	94,362,996
		c.1513G>C	c.1523A>G	c.1888G>C	c.2406G>A	c.2613G>A	c.2702C>T
		p.A505P	p.A508P	p.A630P	p.T802T	p.G871G	p.A901A
	2698	G/G	G/G	G/C			
D2/D2	2705	G/G	G/G	G/G		G/G	T/T
D2/D2	2700	G/G	G/G				
D2/D2	2970						
N/D2	2971						
N/D1	2756	G/G	G/G	G/G	G/A	G/G	T/T
N/N	2758	G/G	A/G	G/G			
D4/D	2676	G/C	A/G	G/C	G/A	G/A	C/T
N/D4	2675						
N/D	2677						
N/D2	2972						
N/D1	2757	G/G	A/G	G/G	G/A	G/G	T/T
N/D1	2516	G/G	G/G	G/G	G/A	G/A	C/T
N/N	2726	G/G	G/G	G/G			
N/D1	2737						
D1/D	2683						
N/D	2736						
D1/D	2703				G/A	G/A	C/T
N/N	2727	G/G	G/G				
N/D2	2723						
N/D2	2847						
D2/D	2674						
D2/D	2673						
N/D2	2724				G/G	G/A	C/T
D2/D2	2748						
D2/D	2746						

Table 3.3(continued). Proposed non-causative SNPs found in *ACAN* for exons 7-11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 7	exon 7	exon 9	exon 11	exon 11	exon 11
Alleles	sample	94,370,258	94,370,248	94,368,175	94,363,292	94,363,085	94,362,996
		c.1513G>C	c.1523A>G	c.1888G>C	c.2406G>A	c.2613G>A	c.2702C>T
		p.A505P	p.A508P	p.A630P	p.T802T	p.G871G	p.A901A
D2/D2	2789	G/G	G/G	G/G		G/G	T/T
D4/D	2745	G/G	A/G	G/C	G/A	G/G	C/T
D1/D1	2738						
D2/D	2749						
D2/D4	2744						
N/D2	2718						
N/D2	2714						
D2/D	2701	G/G	A/G	G/C	G/G	G/A	C/T
D1/D	2747						
D1/D2	2849						
D1/D	2683						
D1/D2	2704						
N/D1	2513						
N/D1	2515						
D2/D2	2715						
N/D2	2716						
D1/D2	2678						
D1/D2	2668						
N/D2	2718						
D1/D4	2896						
D1/D2	2719						
D2/D2	2706	G/G	A/G				
N/D2	2865	G/G	A/G	G/C			
D1/D2	2711						
N/D2	2792						
N/D1	2866	G/G	A/G	G/G			
D2/D2	2709	G/G	G/G	G/G			
N/N	2823	G/G	A/G	G/G			

Table 3.3(continued). Proposed non-causative SNPs found in *ACAN* for exons 7-11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 7	exon 7	exon 9	exon 11	exon 11	exon 11
Alleles	sample	94,370,258	94,370,248	94,368,175	94,363,292	94,363,085	94,362,996
		c.1513G>C	c.1523A>G	c.1888G>C	c.2406G>A	c.2613G>A	c.2702C>T
		p.A505P	p.A508P	p.A630P	p.T802T	p.G871G	p.A901A
D2/D	2906						
D3/D4	2741	G/G	G/G	G/C	G/G	G/A	C/T
N/D3	2742	G/C	G/G	G/C	G/A	G/G	C/T
D2/D4	2694						
D2/D4	2696						
D2/D3	2702						
D2/D3	2710						
D1/D2	2719						
D2/D4	2751						
N/D2	2764						
D2/D2	2752						
N/D1	2721						
N/D1	2514						
N/D2	2669						
D2/D	2695				G/G	G/A	C/T
D2/D2	2707				G/G	G/A	C/T
D2/D2	2708						
N/D2	2750						
	2911						
N/D2	2860						
N/D1	2880						
N/D2	2722						
N/D2	2518						
N/D1	2512						
D2/D2	2699						
N/D2	2524						
N/D1	2514						
N/D4	2672						
D2/D	2697						

Table 3.4. Proposed non-causative SNPs found in *ACAN* for exon 11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 11	exon 11	exon 11
Alleles	sample	94,362,939	94,360,319	94,360,074	94,359,923	94,359,812	94,359,062
		c.2759C>A	c.5379C>T	c.5624C>T	c.5775C>T	c.5886G>A	c.6264G>T
		p.S920T	p.S1793S	p.A1875V	p.Y1925Y	p.S1962S	p.G2088G
	2698		C/C	C/C	C/C	G/G	G/G
D2/D2	2705	C/A					T/T
D2/D2	2700						
D2/D2	2970						
N/D2	2971						
N/D1	2756	A/A	C/T	C/T	C/T	G/G	G/G
N/N	2758						
D4/D	2676			C/T	C/C	G/G	G/G
N/D4	2675		T/T		C/T	G/G	G/G
N/D	2677						
N/D2	2972						
N/D1	2757	A/A	C/T	C/T	C/T	G/G	G/G
N/D1	2516	A/A	C/T	C/T	C/T		
N/N	2726						
N/D1	2737				C/T	G/G	G/G
D1/D	2683						
N/D	2736					G/G	G/G
D1/D	2703		C/T	C/T	C/T		G/G
N/N	2727				C/T	G/G	
N/D2	2723						
N/D2	2847						
D2/D	2674				C/C	G/A	T/T
D2/D	2673		C/C		C/C	G/A	T/T
N/D2	2724	A/A					
D2/D2	2748		C/T				
D2/D	2746		C/T				

Table 3.4(continued). Proposed non-causative SNPs found in *ACAN* for exon 11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 11	exon 11	exon 11
Alleles	sample	94,362,939	94,360,319	94,360,074	94,359,923	94,359,812	94,359,062
		c.2759C>A	c.5379C>T	c.5624C>T	c.5775C>T	c.5886G>A	c.6264G>T
		p.S920T	p.S1793S	p.A1875V	p.Y1925Y	p.S1962S	p.G2088G
D2/D2	2789		C/C		C/T	A/A	T/T
D4/D	2745	C/A	C/T	C/T	C/T	A/A	T/T
D1/D1	2738						
D2/D	2749		C/T				
D2/D4	2744						
N/D2	2718						
N/D2	2714						
D2/D	2701	C/C	C/C	C/C	C/C	G/A	G/T
D1/D	2747		C/T				G/T
D1/D2	2849						
D1/D	2683						
D1/D2	2704						
N/D1	2513						
N/D1	2515						
D2/D2	2715						
N/D2	2716						
D1/D2	2678						
D1/D2	2668						
N/D2	2718						
D1/D4	2896						
D1/D2	2719						
D2/D2	2706						
N/D2	2865						
D1/D2	2711						
N/D2	2792						
N/D1	2866						
D2/D2	2709						
N/N	2823						

Table 3.4(continued). Proposed non-causative SNPs found in *ACAN* for exon 11. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 11	exon 11	exon 11
Alleles	sample	94,362,939	94,360,319	94,360,074	94,359,923	94,359,812	94,359,062
		c.2759C>A	c.5379C>T	c.5624C>T	c.5775C>T	c.5886G>A	c.6264G>T
		p.S920T	p.S1793S	p.A1875V	p.Y1925Y	p.S1962S	p.G2088G
D2/D	2906				C/C	G/A	T/T
D3/D4	2741	C/C		C/T	C/T		
N/D3	2742	C/A		C/T	C/T		
D2/D4	2694				C/C	G/A	T/T
D2/D4	2696						
D2/D3	2702						
D2/D3	2710						
D1/D2	2719						
D2/D4	2751		C/T				
N/D2	2764						
D2/D2	2752						
N/D1	2721						
N/D1	2514						
N/D2	2669						
D2/D	2695	C/C	C/C	C/C	C/C		
D2/D2	2707	C/C	C/C	C/C	C/C	G/A	G/T
D2/D2	2708						
N/D2	2750		C/T				
	2911						
N/D2	2860						
N/D1	2880						
N/D2	2722						
N/D2	2518						
N/D1	2512						
D2/D2	2699						
N/D2	2524						
N/D1	2514						
N/D4	2670	C/C					
D2/D	2697	N/A	N/A	N/A	N/A	N/A	G/T

Table 3.5. Proposed non-causative SNPs found in *ACAN* for exons 11-16. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 13	exon 16	exon 16
Alleles	sample	94,359,062	94,359,106	94,359,062	94,347,621	94,345,187	94,345,096
		c.6275C>T	c.6592A>G	c.6636G>A	c.7102G>A	c.7412G>C	c.7503C>T
		p.A2092V	p.R2198G	p.S2212S	p.A2368T	p.C2471S	p.T2501T
	2698	C/C	G/G	A/A	A/A		
D2/D2	2705	T/T	G/G	G/G			
D2/D2	2700				G/G	C/C	C/C
D2/D2	2970						
N/D2	2971						
N/D1	2756	C/C			G/G		
N/N	2758				G/G		
D4/D	2676	C/C	G/G	G/A	G/G	G/G	C/T
N/D4	2675	C/C	G/G	G/G			
N/D	2677		G/G	G/A			
N/D2	2972						
N/D1	2757	C/C	G/G	G/G	G/G		
N/D1	2516				G/G	G/G	C/T
N/N	2726				G/G		
N/D1	2737	C/C	G/G	G/A			
D1/D	2683						
N/D	2736	C/C	G/G	G/A			
D1/D	2703	C/C	G/G	G/A			
N/N	2727				G/G	G/G	C/T
N/D2	2723						
N/D2	2847						
D2/D	2674	C/T	G/G	G/A			
D2/D	2673	C/T	G/G	G/A			
N/D2	2724						
D2/D2	2748						
D2/D	2746						

Table 3.5(continued). Proposed non-causative SNPs found in *ACAN* for exons 11-16. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 13	exon 16	exon 16
Alleles	sample	94,359,062	94,359,106	94,359,062	94,347,621	94,345,187	94,345,096
		c.6275C>T	c.6592A>G	c.6636G>A	c.7102G>A	c.7412G>C	c.7503C>T
		p.A2092V	p.R2198G	p.S2212S	p.A2368T	p.C2471S	p.T2501T
D2/D2	2789	T/T	G/G	G/G			
D4/D	2745	C/T	G/G		G/G	G/G	C/T
D1/D1	2738					G/G	T/T
D2/D	2749						
D2/D4	2744						
N/D2	2718						
N/D2	2714						
D2/D	2701	C/T	G/G	G/A	G/G		
D1/D	2747	C/T	G/G	G/A			
D1/D2	2849						
D1/D	2683		G/G	G/A			
D1/D2	2704						
N/D1	2513						
N/D1	2515						
D2/D2	2715						
N/D2	2716						
D1/D2	2678						
D1/D2	2668						
N/D2	2718						
D1/D4	2896						
D1/D2	2719						
D2/D2	2706						
N/D2	2865						
D1/D2	2711						
N/D2	2792						
N/D1	2866						
D2/D2	2709				G/G		
N/N	2823						

Table 3.5(continued). Proposed non-causative SNPs found in *ACAN* for exons 11-16. First column genotype of sample; N=normal, D1=dwarf type D1, D2= dwarf type D2, D3=dwarf type D3, D4=dwarf type D4: Second column = sample accession #: Columns of exon SNP genotype shows - position in genome, cDNA base number of gene, amino acid number in gene and alteration, sample genotype at exon location of SNP. Blank spaces indicate no sequence or genotype known for sample.

ACAN		exon 11	exon 11	exon 11	exon 13	exon 16	exon 16
Alleles	sample	94,359,062	94,359,106	94,359,062	94,347,621	94,345,187	94,345,096
		c.6275C>T	c.6592A>G	c.6636G>A	c.7102G>A	c.7412G>C	c.7503C>T
		p.A2092V	p.R2198G	p.S2212S	p.A2368T	p.C2471S	p.T2501T
D2/D	2906	C/T	G/G	G/A			
D3/D4	2741				A/A		
N/D3	2742				G/G		
D2/D4	2694	C/T	G/G	G/G			
D2/D4	2696						
D2/D3	2702						
D2/D3	2710		G/G	G/G			
D1/D2	2719						
D2/D4	2751						
N/D2	2764						
D2/D2	2752						
N/D1	2721						
N/D1	2514						
N/D2	2669						
D2/D	2695				G/G		
D2/D2	2707	C/T	G/G	G/A	G/G		
D2/D2	2708						
N/D2	2750						
	2911						
N/D2	2860						
N/D1	2880						
N/D2	2722						
N/D2	2518						
N/D1	2512						
D2/D2	2699						
N/D2	2524						
N/D1	2514						
N/D4	2670						
D2/D	2697	C/T	G/G	G/A			

Discussion

Since different horses were used at times when sequencing individual exons, this data is a composite of partial gene sequence on the individuals used in this project. The SNPs annotated in this study that cause an amino acid change at the protein level have the potential to have other effects within the Miniature horse that we are not aware of at this time. Considering Miniature horses are the smallest of the equines and GWAS studies have implicated *ACAN*as a contributor to stature in organisms, the diminutive size could possibly be a result from a unique mutation in *ACAN* that does not result in a diseased animal. There are possibly other dwarf types that exist in Miniature horses that were not seen and acquired for the dwarfism study. If other dwarf types are found to exist, some of these SNPs that appear non-causative at this time should be considered for investigation.

There are quite possibly additional SNPs within this gene. This study only considered the sequences of Miniature horses to the reference sequence of a Thoroughbred, Twilight. With the existence of numerous other breeds of equines throughout the world, it is likely more variability exists in the sequence of *ACAN*. With better understanding of the exons expressed in *ACAN* in articular cartilage and other tissues, we will have a more accurate information to determine the correct exon number and splicing information.

The inability to sequence entirely through the ~2400 bp highly conserved repeat region of exon 11 restricted our efforts to better understand the highly conserved sequence of the KS domain in *ACAN*. This region of *ACAN* has interesting and unique properties in structure and function for *ACAN* and more importantly the ECM and has great potential in all areas of research.

Chapter 4

Reflections on Project Complexity, Technology, Remaining Questions

Summary of Aggrecan and Dwarfism study in Miniature horses

Thesis Goals

The primary works of this thesis were focused towards the discovery of the genetic cause of a common type of dwarfism in the Miniature horse with an ultimate goal of providing a genetic test for the breeders of Miniature horses to assist them in identifying carriers and better managing those horses with the genetic defect. More specifically, the objectives were: (1) collect, identify and categorize the different dwarf phenotype samples and related unaffected individuals that exist in the breed, (2) from the dwarf samples gathered throughout the United States, determine the most common type and provide documentation and physical description of that phenotype, (3) to utilize GWAS techniques to find the genetic cause of that type, (4) design a genetic test that can be employed by genetic testing labs for the breeders of Miniature horses.

Project Complexity

Utilizing the GWAS technique was crucial in this project. With the advent of the equine genome being sequenced, this new technology afforded the equine genetics research community the ability to find a genetic cause of a disease in a breeding population that would otherwise not have been possible. Family studies of dwarfism within the Miniature horse breed would have been difficult. This difficulty is in large part due to the inaccuracies with pedigree information of the Miniature horse breed, confidentiality concerns, and sensitivity of breeders.The origins of these mutations will not be completely understood. This is due to the breeding practices in the foundation of the breed, and the use of numerous different pony breeds and various other larger

horse breeds to ultimately produce the Miniature horse of today. These mutations show evidence to be exclusively within the Miniature horse breed. These dwarf types have not been documented to exist in any other pony or large horse breed. Miniature horses have other dwarf types not involving *Acan* and have a type similar to dwarfism in the Friesian mentioned earlier. This provides strong support that these mutations are relatively new within the equine genome. More specifically, it is probable these mutations arose within the smaller pony breeds in the United States given the anecdotal history of how many of the first Miniature horse breeders were well-known pony breeders. Other anecdotal evidence mentions that some of the first small ponies in America, referred to by the Shetland pony breeders as "midget ponies", were later some of the first registered Miniature horses.

This thesis began a project that has become more intricate and complicated than was originally designed. This is due mainly to the most common phenotype of dwarfism in the Miniature horse being a compound heterozygote of two different recessive mutations of the same gene independently segregating in the population and ultimately leading to the discovery of two other mutations with less allelic frequency in the sample group. This thesis began a process that will take much more time and effort in order to properly annotate and categorize these different mutations that have been discovered and determine how they affect the Miniature horse.

Remaining Questions - Possible Future Projects Moving Forward

Some possible projects that could emanate from this original work could involve various specialties of equine research. The following is a list of possible projects that could be beneficial to the equine research community and breeders of Miniature horses.

Quantify exon 11 deletion in the Miniature population

This mutation exists in a small portion of compound heterozygote dwarf samples within this project. Designing a reliable Taqman® SNP Genotyping Assay (Applied Biosystems) for this mutation would need to be accomplished. Testing of dwarf samples within this study that are of unknown status with regards to the mutations found and /or are heterozygous for a known mutation would need to be performed. Also, testing of non-dwarf individuals in the control population of this study group would need to be performed as well. However, testing other dwarfs of unknown cause that do not exhibit the genetic mutations found in this thesis would need to be performed to discover status, i.e. Miniature horse dwarfs with skeletal atavism. To further determine population penetrance of this mutation within the breed, a large random group of individuals in the population could be tested. This large group could be utilized to determine the penetrance of the other mutations found in *Acan* that cause dwarfism as well.

Pathological Examination of Articular Cartilage and ECM of 4 Mutations of Aggrecan Causing Dwarfism

This project would entail work in Equine Pathology and/or Musculoskeletal Sciences. *Acan* is an integral part of cartilage and allows the ECM to function properly. *Acan* is the reason weight bearing cartilage can retain water enabling joints to handle the stresses experienced on a daily basis. Four different independently segregating mutations within this gene cause dwarfism in the Miniature horse. This project's goal would be to determine through examination of the cellular structures of articular cartilage how each of these mutations affect the horse in homozygous form and in compound heterozygous form to produce the resulting phenotypes seen in the breed. Collecting cartilage tissue and bone samples of each donated dwarf would be done

after DNA testing provided genotype confirmation of mutations the dwarf carried. Physical records with picture documentation would be performed of phenotype with detailed gross anatomical descriptions of each homozygous mutation and novel combinations. Microscopic examinations would be performed on specific cartilage samples of each type of dwarf to provide explanation of resulting cartilage and bone malformations due to the mutations.

Use of Vector / Plasmid to Sequence the CS Region of Equine Aggrecan

A large portion of the CS region of Equine *Acan*was not able to be sequenced by typical short run sequencing utilized in this project by the ABI 310 genetic analyzer. Due to the high conservation of sequence and unique repeat pattern of the CS region in exon 11 of *Acan*, it is not possible with current sequencing technology to read through this region in a single long read and produce consist and reliable results. Therefore, a project utilizing unique restriction enzyme splice sites that sections this ~2.4 – 3.5 kb DNA sequence into sizes that could be inserted in plasmids / vectors and then be reliably sequenced by a short read sequencer repeatedly could be designed. First would be to determine the existence of length variability within this region by simply designing reliable primers and PCR protocol that produced consistent bands of reliable size. In humans, this region has been shown to be highly variable in the general population (Doege *et al.* 1997). This would be a key to know in the horse to determine possible splice site variability thus possible fragment size variability with digestion and transfection ultimately causing sequence read overlapping inconsistencies. If there is variability in sequence length of the CS region, the PCR protocol could be utilized to determine allelic variability in the size of this repeat region in the equine species as a whole or even a specific breed. If Allelic variability exists within the equine, the samples that are determined to be homozygous for an allelic size

could then be utilized in the vector /plasmid transfection protocol for sequencing that allelic size. This region would also be an area to consider for any unknown dwarf types seen in the Miniature horse that do not express the four mutations described here.

Final Thoughts

Evaluating the Miniature horse's foundation provides a general concept of the genetic diversity that has ultimately produced this breed. It is necessary to understand the foundation and breeding practices of this breed in order to better comprehend the inherited genetic problems that permeate it. The goal of the Miniature horse is to breed the smallest most correct and proportionate horse in miniature. However, small size was the overriding breeding criteria used in the mid to later part of the 20th century, during the greatest popularity of the breed and population boom. With this variation in genetics, the desire to produce the smallest Miniature horses and having no strict quality standards of bloodstock led to the use of severely malformed small Miniature horses being used in breeding programs. A majority of these conformationally inferior Miniature horses showed consistent physical abnormalities typical of disproportionate dwarfisms as seen in other mammal species. Disproportionate dwarfism in Miniature horses, in the decades past, has been a well-known unspoken problem within the breed and has known to exist through anecdotal evidence and record keeping of breeders. This problem has now become a more openly discussed concern within the Miniature horse industry throughout the world. Dwarf Miniature horses, while often setting world records for small size are not considered to have desirable traits. These horses possessing incorrect conformation typical of disproportionate dwarfism have significant health, quality of life and soundness issues. Therefore, many Miniature horse registries try to avoid accepting Miniature horses affected by dwarfism for

breeding stock during the registration process (AMHA 2012; AMHR 2012). Unfortunately, Miniature horses which are not dwarfs and are small in size may also express extremely inferior conformational characteristics. These individuals may be confused with dwarfism and therefore compound the problem of identification of actual dwarfism and identification of other possible inferior genetic issues within the breed. These other inferior individuals, along with dwarfs, are still used in some breeding programs today to produce the smallest horse possible. These practices may reveal the possibility of the existence of other mutations that cause a type of dwarfism within *Acan* that was not found in this project or even a mutation of a different genetic cause.

APPENDIX

Supplementary Data, Tables and Figures

Table 2.S1. Shows the 10 most significant associations found for SNPs and the occurrence of dwarfism. The only SNP that achieved statistical significance based on EMP2 score was BIEC2_38994 found on ECA1. This observation was supported by 6 of the 10 strongest associations occurring with SNPs on ECA1.

CHR	SNP	Chi-Square	EMP2
1	BIEC2_38994	23.54	0.01898
1	BIEC2_40049	18.92	0.1558
1	BIEC2_38884	18.83	0.1628
1	BIEC2_38959	17.1	0.3516
1	BIEC2_38970	17.1	0.3516
8	BIEC2_1027066	16.89	0.3956
1	BIEC2_39609	15.75	0.5315
5	BIEC2_901067	14.1	0.7892
31	BIEC2_841695	13.63	0.8631
22	BIEC2_577120	13.28	0.8911

Table 2.S2
Haplotypes with SNP BIEC2_38994 shows 5 Haplotypes among dwarfs.

Genotype	Dwarf	Control
1/2	1	7
1/3	6	1
2/2	0	8
2/3	4	5
2/4	0	1
2/5	1	0
3/3	5	0
3/5	2	0
4/5	0	1
	19	23

Table 2.S3. Primers used for ACAN sequencing exons 1-17 Lower case sequences indicate intron. Upper case indicate exon.Exon numbering is according to UCSC Genome browser prediction onEcab build 2.0.

Exon #	Forward Primer Sequence	Reverse Primer Sequence
Exon 1	gtgacctttgccctcactgt	gcacccagaatccagtcttc
Exon 2	tgggtggtcctctctagcac	tttccacaggtgaagcaaca
Exon 3	gtgcctgacctgctctatcg	agctcagtgctggtcaacg
Exon 4	cctgagtgtcacatcccactt	gagtggtagtggggtgaagg
Exon 5	gcctgctttgtccttcacag	aaaacagccccctattccac
Exon 6	gggctgagccgctaaagtt	aggccaagttccttccactt
Exon 7	gtctctccttctcgccctct	aagcctgacccttgagactg
Exon 8	agaacaggccctcattctgc	aggtaatgccctctcctcgt
Exon 9	gtgccacctgcctctgtc	cagaagtgggttctggagga
Exon 10	aggaggaaccttcaccacct	cgcccaagccatacgaac
Exon 11a	ggagcagtttctaatcccaca	ACTGAGGTCCTCTGCTCCAG
Exon 11b	TATCTCTGCAGTGGGCTCAG	AGGCAGTGGGCTCTAGATGA
Exon 11c	GAGTGGAGGACCTTGGTGAA	gaggcagtgggctctaaatg
Exon 12a	tgcctccggagtagaggac	TGACGACTTCCACCAATGTC
Exon 12b	GGCAAGCTCCTGAAGCAAGT	TCCAGATGTTGTCCCACTGA
Exon 12c	GTGGACTGTCCTCTGGACAAC	CCCTTCTCCTGCTTCTTGG
Exon 12d	TCCTCTGGAGCTGAGACTGG	ATCAGGGGACCCAGAAGC
Exon 12e	CCCAGCTTGTTGAGTCCAGT	ctgaactacccaaacccctttt
Exon 13	cttcaagcccctgacctgt	ctgctgacttctggcaagtg
Exon 14	cccaaacccacatcttctct	atgccgtccgaatgtatctc
Exon 15	ctgccctctgctcacctct	ggagccgaagtcttgattct
Exon 16	cacaggagccctttctgaag	cagggaggaggaggtgct
Exon 17	ccgagggctcactaggattt	gacgaagtgtcggtgatctg

Table 2.S4. Testing results of samples using Taqman® RT-PCR for exons 2 and 6. As well as PCR results of exon 15 run on 2% agarose gel. Taqman® exon 6; CC = Normal genotype, CT= Carrier genotype *D2*: Taqman® Exon 2: AA = Normal genotype, A- = Carrier genotype *D1*: Exon 15; CC = Normal genotype, C- = Carrier genotype.

Acc. #	Breed	Taqman exon 6	Taqman exon 2	Exon 15 Del
13307	AMHR	CT	AA	CC
13308	AMHR	CC	AA	CC
13309	AMHR	CC	AA	CC
13310	AMHR	CC	AA	CC
13311	AMHR	CC	AA	CC
13312	AMHR	CC	AA	CC
13313	AMHR	CC	AA	CC
13314	AMHR	CC	AA	CC
13315	AMHR	CC	AA	CC
13316	AMHR	CC	AA	CC
13317	AMHR	CC	AA	CC
13318	AMHR	CC	AA	CC
13319	AMHR	CC	AA	CC
13320	AMHR	CC	AA	CC
13321	AMHR	CC	AA	CC
13322	AMHR	CC	AA	CC
13323	AMHR	CC	AA	CC
13324	AMHR	CC	AA	CC
13325	AMHR	CC	AA	CC
13326	AMHR	CC	AA	CC
13327	AMHR	CC	AA	CC
13328	AMHR	CC	AA	CC
13329	AMHR	CC	AA	CC
13330	AMHR	CC	AA	CC
13462	AMHR	CC	AA	CC
13463	AMHR	CC	AA	CC
13464	AMHR	CC	AA	CC
13465	AMHR	CC	AA	CC
13466	AMHR	CC	AA	CC
13467	AMHR	CC	AA	CC
13468	AMHR	CC	AA	CC
13469	AMHR	CC	AA	CC
13470	AMHR	CC	AA	CC
13471	AMHR	CC	AA	CC
13472	AMHR	CC	AA	CC
13473	AMHR	CC	AA	CC
13474	AMHR	CC	AA	CC
13475	AMHR	CT	AA	CC
13476	AMHR	CC	AA	CC

Table 2.S4(continued). Testing results of samples using Taqman® RT-PCR for exons 2 and 6. As well as PCR results of exon 15 run on 2% agarose gel. Taqman® exon 6; CC = Normal genotype, CT= Carrier genotype *D2*: Taqman® Exon 2: AA = Normal genotype, A- = Carrier genotype *D1*: Exon 15; CC = Normal genotype, C- = Carrier gentoype

Acc. #	Breed	Taqman exon 6	Taqman exon 2	Exon 15 Del
H102526	AMHR	CC	AA	CC
H102527	AMHR	CC	AA	CC
H102528	AMHR	CC	AA	CC
2936	AMHR	CC	AA	CC
2798	Hackney	CC	AA	CC
2813	Caspian	CC	AA	CC
1835	ARA	CC	AA	CC
1836	ARA	CC	AA	CC
1827	ARA	CC	AA	CC
1828	ARA	CC	AA	CC
Q1-17	SAB	CC	AA	CC
R1-20	SAB	CC	AA	CC
R1-12	SAB	CC	AA	CC
R1-14	SAB	CC	AA	CC
C-180	TW	CC	AA	CC
C-110	TW	CC	AA	CC
C-230	TW	CC	AA	CC
C-120	TW	CC	AA	CC

Table 2.S4(continued). Testing results of samples using Taqman® RT-PCR for exons 2 and 6. As well as PCR results of exon 15 run on 2% agarose gel. Taqman® exon 6; CC = Normal genotype, CT= Carrier genotype *D2*: Taqman® Exon 2: AA = Normal genotype, A- = Carrier genotype *D1*: Exon 15; CC = Normal genotype, C- = Carrier gentoype

Acc. #	Breed	Taqman exon 6	Taqman exon 2	Exon 15 Del
Standard	Twilight TB	CC	AA	CC
2433	TB	CC	AA	CC
1197	TB	CC	AA	CC
2449	TB	CC	AA	CC
307	TB	CC	AA	CC
1178	TB	CC	AA	CC
431	TB	CC	AA	CC
421	TB	CC	AA	CC
1179	TB	CC	AA	CC
463	TB	CC	AA	CC
1273	TB	CC	AA	CC
1261	TB	CC	AA	CC
315	TB	CC	AA	CC
415	TB	CC	AA	CC
429	TB	CC	AA	CC
1195	TB	CC	AA	CC
457	TB	CC	AA	CC
1232	TB	CC	AA	CC
2269	TB	CC	AA	CC
1345	TB	CC	AA	CC
467	TB	CC	AA	CC
471	TB	CC	AA	CC
472	TB	CC	AA	CC
475	TB	CC	AA	CC
603	TB	CC	AA	CC
605	TB	CC	AA	CC
622	TB	CC	AA	CC
1242	TB	CC	AA	CC

Table 2.S4(continued). Testing results of samples using Taqman® RT-PCR for exons 2 and 6. As well as PCR results of exon 15 run on 2% agarose gel. Taqman® exon 6; CC = Normal genotype, CT= Carrier genotype *D2*: Taqman® Exon 2: AA = Normal genotype, A- = Carrier genotype *D1*: Exon 15; CC = Normal genotype, C- = Carrier gentoype

Acc. #	Breed	Taqman exon 6	Taqman exon 2	Exon 15 Del
923	ST	CC	AA	CC
924	ST	CC	AA	CC
925	ST	CC	AA	CC
926	ST	CC	AA	CC
927	ST	CC	AA	CC
928	ST	CC	AA	CC
929	ST	CC	AA	CC
930	ST	CC	AA	CC
931	ST	CC	AA	CC
932	ST	CC	AA	CC
933	ST	CC	AA	CC
934	ST	CC	AA	CC
935	ST	CC	AA	CC
936	ST	CC	AA	CC
937	ST	CC	AA	CC
938	ST	CC	AA	CC
939	ST	CC	AA	CC
940	ST	CC	AA	CC
942	ST	CC	AA	CC
943	ST	CC	AA	CC
944	ST	CC	AA	CC
945	ST	CC	AA	CC

REFERENCES

Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., & Hayward, C. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832-838.

AMHA. American Miniature Horse Association (2012) *Official Rule Book of American Miniature Horse Association, 25th Edition* (2012). Alvarado, TX: American Miniature Horse Association.

AMHR. American Miniature Horse Registry (2012). *The American Miniature Horse Registry Official Rule Book*,(2012). Morton IL: American Miniature Horse Registry.

ASPC. American Shetland Pony Club (2012).*American Shetland Pony Club Official Rule Book*, 2012. Morton, IL: American Shetland Pony Club.

Antonsson, P., Heinegård, D., & Oldberg, A. (1989). The keratan sulfate-enriched region of bovine cartilage proteoglycan consists of a consecutively repeated hexapeptide motif. *Journal of Biological Chemistry*, *264*(27), 16170-16173.

Aspberg, A. (2012). The different roles of aggrecan interaction domains. *Journal of Histochemistry & Cytochemistry*, *60*(12), 987-996.

Back, W., LUGT, J., Nikkels, P. G. J., BELT, A., KOLK, J., & Stout, T. A. E. (2008). Phenotypic diagnosis of dwarfism in six Friesian horses. *Equine Veterinary Journal*, *40*(3), 282-287.

Bellus GA, McIntosh I, Smith EA, Ayslworth AS, Kaitila I, Horton WA, Greenhaw FA, Hecht JT, Francomano CA. (1995). A recurrent mutation in the tyrosine kinase domain of fibroblast growth factor receptor 3 causes hypochondroplasia. *Nature Genetics*, 10:357–359.

Blum J W , Elsasser T H , Greger D L , Wittenberg S , de Vries F , Distl O. (2007). Insulin-like growth factor type-1 receptor down-regulation associated with dwarfism in Holstein calves. *Domestic Animal Endocrinology*, 33(3):245-68.

BMHS. British Miniature Horse Society (2012). *British Miniature Horse Society Hand Book*, (2012). Stretcholt, Bridgewater, Somerset, United Kingdom: The British Miniature Horse Society.

Bochukova, E., Schoenmakers, N., Agostini, M., Schoenmakers, E., Rajanayagam, O., Keogh, J. M., & Chatterjee, K. (2012). A mutation in the thyroid hormone receptor alpha gene. *New England Journal of Medicine*, *366*(3), 243-249.

Bottner A, Keller E, Kratzsch J, Stobbe H, Weigel JF, Keller A, Hirsch W, Kiess W, Blum WF, Pfaffle RW. PROP1 mutations cause progressive deterioration of anterior

pituitary function including adrenal insufficiency: a longitudinal analysis. *Journal of Clinical Endocrinology & Metabolism*. 2004; 89:5256–65.

Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B., & Ruoslahti, E. (1987). Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proceedings of the National Academy of Sciences*, *84*(10), 3194-3198.

Cavanagh, J.A., Tammen, I., Windsor, P.A., Bateman, J.F., Savarirayan, R., Nicholas, F.W., and Raadsma, H.W. (2007). Bulldog dwarfism in Dexter cattle is caused by mutations in ACAN. *Mammalian Genome,* 18, 808–814.

Chen, L., Wu, Y., Lee, V., Kiani, C., Adams, M. E., Yao, Y., & Yang, B. B. (2002). The folded modules of *aggrecan* G3 domain exert two separable functions in glycosaminoglycan modification and product secretion. *Journal of Biological Chemistry*, *277*(4), 2657-2665.

Clutton-Brock, J. (1999). *A natural history of domesticated mammals*. Cambridge University Press.

Dartmoor Pony Society (2013). *Dartmoor Pony Stud Book*, (2013) Swn-Yr-Afon, Cwmgwili, Llanelli: Dartmoor Pony Society.

Doege, K. J., Sasaki, M., Kimura, T., & Yamada, Y. (1991). Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, *aggrecan*. Human-specific repeats, and additional alternatively spliced forms. *Journal of Biological Chemistry*, *266*(2), 894-902.

Doege, K. J., Coulter, S. N., Meek, L. M., Maslen, K., & Wood, J. G. (1997). A Human-specific Polymorphism in the Coding Region of the Aggrecan Gene VARIABLE NUMBER OF TANDEM REPEATS PRODUCE A RANGE OF CORE PROTEIN SIZES IN THE GENERAL POPULATION. *Journal of Biological Chemistry*, *272*(21), 13974-13979.

Dutson, Judith (2005). *Storey's Illustrated Guide to 96 Horse Breeds of North America*. Storey Publishing. 321-22.

Exmoor Pony Society (2013). *Exmoor Pony Stud Book*, (2013). Woodmans, Brtithem Bottom, Cullompton, Devon, United Kingdom: Exmoor Pony Society.

Fosang, A. J., & Hardingham, T. E. (1989). Isolation of the N-terminal globular protein domains from cartilage proteoglycans. Identification of G2 domain and its lack of interaction with hyaluronate and link protein. *Biochemical Journal*, *261*, 801-809.

Fosang, A. J., & Hardingham, T. E. (1991). 1-C-6 epitope in cartilage proteoglycan G2 domain is masked by keratan sulphate. *Biochemical Journal*, *273*, 369-373.

Fosang, A. J., Last, K., & Maciewicz, R. A. (1996). Aggrecan is degraded by matrix metalloproteinases in human arthritis. Evidence that matrix metalloproteinase and aggrecanase activities can be independent. *Journal of Clinical Investigation*, *98*(10), 2292.

Fülöp, C., Walcz, E., Valyon, M., & Glant, T. T. (1993). Expression of alternatively spliced epidermal growth factor-like domains in *aggrecan*s of different species. Evidence for a novel module. *Journal of Biological Chemistry*, *268*(23), 17377-17383.

Gleghorn, L., Ramesar, R., Beighton, P., & Wallis, G. (2005). A Mutation in the Variable Repeat Region of the Aggrecan Gene (AGC1) Causes a Form of Spondyloepiphyseal Dysplasia Associated with Severe, Premature Osteoarthritis. *The American Journal of Human Genetics*, *77*(3), 484-490.

Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V., Zusmanovich, P., & Stefansson, K. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, *40*(5), 609-615.

Halberg, D. F., Proulx, G., Doege, K., Yamada, Y., & Drickamer, K. (1988). A segment of the cartilage proteoglycan core protein has lectin-like activity. *Journal of Biological Chemistry*, *263*(19), 9486-9490.

Hardingham, T. E., & Muir, H. (1972). The specific interaction of hyaluronic acid with cartilage proteoglycans. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *279*(2), 401-405.

Hardingham, T. E., & Muir, H. (1973). Binding of oligosaccharides of hyaluronic acid to proteoglycans. *Biochemical Journal*, *135*(4), 905.

Hardingham, T. E., & Fosang, A. J. (1995). The structure of aggrecan and its turnover in cartilage. *The Journal of rheumatology. Supplement*, *43*, 86.

Heinegård, D., & Hascall, V. C. (1974). Aggregation of cartilage proteoglycans III. Characteristics of the proteins isolated from trypsin digests of aggregates. *Journal of Biological Chemistry*, *249*(13), 4250-4256.

Hendriks-Stegman, B., Augustijn, K., Bakker, B., Holthuizen, P., Van Der Vliet, P. & Jansen, M. (2001) Combined pituitary hormone deficiency caused by compound heterozygosity for two novel mutations in the POU domain of the *PIT1*/*POU1F1*. *Journal of Clinical Endocrinology and Metabolism*, 86, 1545 – 1550.

Hopwood, J. J., & Robinson, H. C. (1974). The alkali-labile linkage between keratan sulphate and protein. *Biochemical Journal*, *141*, 57-69.

International Museum of the Horse, Legacy of the Horse Exhibit. (2013). Kentucky Horse Park, Lexington, KY.

Irie, Y., Tatsumi, K., Ogawa, M., Kamijo, T., Preeyasombat, C., Suprasongsin, C. & Amino, N. (1995) A novel E250X mutation of the *PIT1* gene in a patient with combined pituitary hormone deficiency. *Endocrinology Journal,* 42, 351 – 354.

Kiani C, Lee V, Cao L, Chen L, Wu Y, Zhang Y, Adams ME, Yang BB. (2001). Roles of *aggrecan* domains in biosynthesis, modification by glycosaminoglycans and product secretion. *Biochemical Journal*, *354*, 199-207.

Kiani C, Chen L, Wu YJ, Yee AJ, Yang BB. 2002. *Structure and function of aggrecan. Cell Research,* 12: 19–32.

Kim, J. J., Lee, H. I., Park, T., Kim, K., Lee, J. E., Cho, N. H., & Lee, J. K. (2009). Identification of 15 loci influencing height in a Korean population. *Journal of Human Genetics*, *55*(1), 27-31.

Kimata, K., Barrach, H. J., Brown, K. S., & Pennypacker, J. P. (1981). Absence of proteoglycan core protein in cartilage from the cmd/cmd (cartilage matrix deficiency) mouse. *Journal of Biological Chemistry*, *256*(13), 6961-6968.

Krueger, R. C., Fields, T. A., Hildreth, J., & Schwartz, N. B. (1990). Chick cartilage chondroitin sulfate proteoglycan core protein. I. Generation and characterization of peptides and specificity for glycosaminoglycan attachment. *Journal of Biological Chemistry*, *265*(20), 12075-12087.

Lark, M. W., Bayne, E. K., Flanagan, J., Harper, C. F., Hoerrner, L. A., Hutchinson, N. I., & Lohmander, L. S. (1997). *Aggrecan* degradation in human cartilage. Evidence for both matrix metalloproteinase and *aggrecan*ase activity in normal, osteoarthritic, and rheumatoid joints. *Journal of Clinical Investigation*, *100*(1), 93.

Lettre, G., Jackson, A. U., Gieger, C., Schumacher, F. R., Berndt, S. I., Sanna, S., & Hirschhorn, J. N. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. *Nature Genetics*, *40*(5), 584-591.

Li H, Schwartz NB, Vertel BM (1993) cDNA cloning of chick cartilage chondroitin sulfate (*aggrecan*) core protein and identification of a stop codon in the *aggrecan* gene associated with the chondrodystrophy, nanomelia. *Journal of Biological Chemistry,* 268: 23504–23511

Makvandi-Nejad, S., Hoffman, G. E., Allen, J. J., Chu, E., Gu, E., Chandler, A. M., & Sutter, N. B. (2012). Four loci explain 83% of size variation in the horse. *PLoS One*, *7*(7), e39929.

McNICOL, D., & Roughley, P. J. (1980). Extraction and characterization of proteoglycan from human meniscus. *Biochemical Journal*, *185*, 705-713.

Moritomo, Y., Ishibashi, T., & Miyamoto, H. (1992). Morphological changes of epiphyseal plate in the long bone of chondrodysplastic dwarfism in Japanese brown cattle. *The Journal of Veterinary Medical Science/The Japanese Society of Veterinary Science*, *54*(3), 453.

Mort, J. S., & Buttle, D. J. (1997). Cathepsin B. *The international journal of biochemistry & cell biology*, *29*(5), 715-720.

MORT, J., MAGNY, M., & LEE, E. (1998). Cathepsin B: an alternative protease for the generation of an aggrecan 'metalloproteinase'cleavage neoepitope. *Biochemical Journal*, *335*, 491-494.

Mow, V. C., Zhu, W., Michael Lai, W., Hardingham, T. E., Hughes, C., & Muir, H. (1989). The influence of link protein stabilization on the viscometric properties of proteoglycan aggregate solutions. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *992*(2), 201-208.

Naviaux, Barbara. *Miniature Horses - Their Care, Breeding And Coat Colors*. 1999. Pp. 2-9.

Nieduszynski, I. A., Huckerby, T. N., Dickenson, J. M., Brown, G. M., Tai, G. H., Morris, H. G., & Eady, S. (1990). There are two major types of skeletal keratan sulphates. *Biochemical Journal*, *271*, 243-245.

Orr, N., Back, W., Gu, J., Leegwater, P., Govindarajan, P., Conroy, J. & Brama, P. A. J. (2010). Genome-wide SNP association–based localization of a dwarfism gene in Friesian dwarf horses. *Animal Genetics*, *41*(s2), 2-7.

Pannier, S., Martinovic, J., Heuertz, S., Delezoide, A. L., Munnich, A., Schibler, L., & Legeai-Mallet, L. (2009). Thanatophoric dysplasia caused by double missense FGFR3 mutations. *American Journal of Medical Genetics Part A*, *149*(6), 1296-1301.

Paulsson, M., Morgelin, M., Wiedemann, H., Beardmore-Gray, M., Dunham, D., Hardingham, T., & Engel, J. (1987). Extended and globular protein domains in cartilage proteoglycans. *Biochemical Journal*, *245*, 763-772.

Perkins, S. J., Nealis, A. S., Dudhia, J., & Hardingham, T. E. (1989). Immunoglobulin fold and tandem repeat structures in proteoglycan N-terminal domains and link protein. *Journal of Molecular Biology*, *206*(4), 737-748.

Pryce, J. E., Hayes, B. J., Bolormaa, S., & Goddard, M. E. (2011). Polymorphic regions affecting human height also control stature in cattle. *Genetics*, *187*(3), 981-984.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics, 81*(3), 559-575.

Rafati Nima, Mikko Sofia, Andersson Lisa S., Thyreen Gunilla, Dalin Göran, Eberth John E., Bailey Ernie, Andersson Leif, Lindgren Gabriella, Rubin Carl-Johan. (2013) Identification of the locus causing skeletal atavism in Shetland ponies. *Proceedings of the International Havemeyer Horse Genome Workshop*, Azores, Portugal; July 10-12, 2013.

Rittenhouse, E., Dunn, L. C., Cookingham, J., Calo, C., Spiegelman, M., Dooher, G. B. & Bennett, D. (1978) *Journal of Embryology and Experimental Morphology*, 43, 71–84.

Sajdera, S. W., & Hascall, V. C. (1969). Protein polysaccharide Complex from Bovine Nasal Cartilage A COMPARISON OF LOW AND HIGH SHEAR EXTRACTION PROCEDURES. *Journal of Biological Chemistry*, *244*(1), 77-87.

Sandy, J. D., Flannery, C. R., Neame, P. J., & Lohmander, L. S. (1992). The structure of aggrecan fragments in human synovial fluid. Evidence for the involvement in osteoarthritis of a novel proteinase which cleaves the Glu 373-Ala 374 bond of the interglobular domain. *Journal of Clinical Investigation*, *89*(5), 1512.

Shiang R, Thompson LM, Zhu YZ, Church DM, Fielder TJ, Bocian M, Winokur ST, Wasmuth JJ. (1994). Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell*, *78*(2), 335-342.

Shetland Pony Stud-Book Society (2013). *Articles of Association of the Shetland Pony Stud-Book Society, 2010*. Perth PH2 8EH Scotland: Shetland Pony Stud-Book Society.

Singer, I. I., Kawka, D. W., Bayne, E. K., Donatelli, S. A., Weidner, J. R., Williams, H. R., & Glant, T. T. (1995). VDIPEN, a metalloproteinase-generated neoepitope, is induced and immunolocalized in articular cartilage during inflammatory arthritis. *Journal of Clinical Investigation*, *95*(5), 2178.

South Africa Miniature Horse Breeders' Society (2013). *Articles of Association of the Miniature Horse Breeders' Society of South Africa* BLOEMFONTEIN, SA: South Africa Miniature Horse Breeders' Society.

Stattin, E. L., Wiklund, F., Lindblom, K., Önnerfjord, P., Jonsson, B. A., Tegner, Y., & Aspberg, A. (2010). A missense mutation in the aggrecan C-type lectin domain disrupts extracellular matrix interactions and causes dominant familial osteochondritis dissecans. *The American Journal of Human Genetics*, *86*(2), 126-137.

Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, *et al*. (2007). A single IGF1 allele is a major determinant of small size in dogs. *Science*, 316: 112–115.

Takeda H, Takami M, Oguni T, Tsuji T, Yoneda K, Sato H, Ihara N, Itoh T, Kata SR, Mishina Y, Womack JE, Moritomo Y, Sugimoto Y, Kunieda T. (2002) Positional cloning of the gene LIMBIN responsible for bovine chondrodysplastic dwarfism. *Proceedings of the National Academy of Science, 99*, 10549–10554.

Tompson, S. W., Merriman, B., Funari, V. A., Fresquet, M., Lachman, R. S., Rimoin, D. L., & Krakow, D. (2009). A recessive skeletal dysplasia, SEMD *aggrecan* type, results from a missense mutation affecting the C-type lectin domain of *aggrecan*. *The American Journal of Human Genetics*, 84(1), 72-79.

Tortorella MD, Burn TC, Pratta MA, Abbaszade I, Hollis JM, Liu R, Rosenfeld SA, Copeland RA, Decicco CP, Wynn R, Rockwell A, Yang F, Duke JL, Solomon K, George H, Bruckner R, Nagase H, Itoh Y, Ellis DM, Ross H, Wiswall BH, Murphy K, Hillman MC, Jr., Hollis GF, Arner EC, *et al*. (1999). Purification and cloning of *aggrecan*ase-1: a member of the ADAMTS family of proteins. *Science*, *284*(5420), 1664-1666.

Tyson, R., Graham, J. P., Colahan, P. T., & Berry, C. R. (2004). Skeletal atavism in a miniature horse. *Veterinary Radiology & Ultrasound*, *45*(4), 315-317.

van Mullem, A., van Heerebeek, R., Chrysis, D., Visser, E., Medici, M., Andrikoula, M., Tsatsoulis, A., Peeters, R., Visser, T. J. (2012) Clinical phenotype and mutant TR-alpha-1. (Letter) *New England Journal of Medicine*, 366: 1451-1453.

Valhmu, W. B., Palmer, G. D., Rivers, P. A., Ebara, S., Cheng, J. F., Fischer, S., & Ratcliffe, A. (1995). Structure of the human *aggrecan* gene: exon-intron organization and association with the protein domains. *Biochemical Journal*, *309*, 535-542.

Vertel, B. M., Walters, L. M., Grier, B., Maine, N., & Goetinck, P. F. (1993). Nanomelic chondrocytes synthesize, but fail to translocate, a truncated *aggrecan* precursor. *Journal of Cell Science, 104*(3), 939-948.

Watanabe, H., Yamada, Y., & Kimata, K. (1998). Roles of aggrecan, a large chondroitin sulfate proteoglycan, in cartilage structure and function. *Journal of biochemistry, 124*(4), 687-693.

Watanabe, H., Nakata, K., Kimata, K., Nakanishi, I., & Yamada, Y. (1997). Dwarfism and age-associated spinal degeneration of heterozygote cmd mice defective in aggrecan. *Proceedings of the National Academy of Sciences*, *94*(13), 6943-6947.

Watanabe, H., Kimata, K., Line, S., Strong, D., Gao, L. Y., Kozak, C. A., & Yamada, Y. (1994). Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the *aggrecan* gene. *Nature Genetics*, *7*(2), 154-157.

Weedon, M. N., Allen, H., Lindgren, C. M., Wallace, C., Evans, D. M., Mangino, M., & Frayling, T. M. (2008). Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics*, *40*(5), 575-583.

Wight, T. N., Kinsella, M. G., & Qwarnström, E. E. (1992). The role of proteoglycans in cell adhesion, migration and proliferation. *Current Opinion in Cell Biology*, *4*(5), 793-801.

Zimmermann, D. R., &Ruoslahti, E. (1989). Multiple domains of the large fibroblast proteoglycan, versican. *The EMBO Journal*, *8*(10), 2975.

V ITA

John Edmund Eberth

Education

University of Kentucky 1996-2013
Lexington, Kentucky, USA
Master of Science in Veterinary Science
College of Agriculture, Department of Veterinary Science
Maxwell H. Gluck Equine Research Center

Western Kentucky University 1995-1996
Bowling Green, Kentucky, USA
College of Agriculture, Department of Animal Science

DePauw University 1990-1995
Greencastle, Indiana, USA
Bachelor of Arts in Biology

Professional Positions

Graduate Research Assistant, University of Kentucky – August 1996 to Present

Licensed AMHA Judge, AMHA – February 2008 to Present

President, Arion Management Inc. – January 1998 to December 2012

Equine Trainer, Little King Farm Inc. – May 1989 to August 1996

Veterinary Technician, Snodgrass Veterinary Clinic – August 1995 to June 1996

Student Intern, U.C. Davis Veterinary Genetics Testing Lab – December 1994 – February 1995

Student Intern, Equine Services PSC. – December 1993 – February 1994

Student Intern, Louisville Zoo – December 1992 – February 1993

Awards and Honors

Chairman Equine Genetics Committee**,** AMHA – February 2006 to February 2011

Science Award for High School Senior Achievement – May 1990

Publications

JE. Eberth, T. Swerczak, E. Bailey. (2009) Investigation of Dwarfism Among Miniature Horses using the Illumina Horse SNP50 Bead Chip. *Journal of Equine Veterinary Science*. **29**(5). P.315.

John E. Eberth. (2011) It's All in the Genes. *International Show Horse*. Blue Ribbon Edition. **2.** P. 162.

Publications Submitted and in Preparation

Patent application in preparation for identification of dwarfism mutations in *ACAN*.

Abstracts

Sofia, M., L. Andersson, Thyreen Gunilla, Dalin Göran, **Eberth John E**., Bailey E., Andersson Leif, Lindgren Gabriella, Rubin Carl-Johan. (2013) Identification of the locus causing skeletal atavism in Shetland ponies. Proceedings of the International Havemeyer Horse Genome Workshop, Azores, Portuga; July 10-12, 2013. p.54.

Eberth J. & Bailey E. (2010) Genetics of dwarfism in Miniature Horses. ISAG, Edinburgh, Scotland. July 25-30, 2010.

Bailey, E., Easley, J, **Eberth, J.** (2013) GWAS for congenital brachygnathia of Thoroughbred Horses. Proceedings of the International Havemeyer Horse Genome Workshop, Azores, Portugal; July 10-12, 2013. p.136.

Presentations at Scientific Meetings

JE Eberth. (May 2009) *Investigation of Dwarfism Among Miniature Horses using the Illumina Horse SNP50 Bead Chip.* 2009 Equine Science Society Symposium, Keystone, CO.

Oral Presentations

JE Eberth. (July 2013) *Chondrodysplasia-like Dwarfism in the Miniature Horse*. Master's Thesis Defense Seminar. University of Kentucky, Lexington, KY.

JE Eberth. (May 2009) *Investigation of Dwarfism Among Miniature Horses using the Illumina Horse SNP50 Bead Chip.* 2009 Equine Science Society Symposium, Keystone, CO.

JE Eberth. (April 2008) *Preliminary Investigation of Candidate Genes for Type 1 Dwarfism in Miniature Horses.* Veterinary Science Graduate Seminar. University of Kentucky, Lexington, KY.

Outreach Presentations

JE Eberth. (January 2013) *Genetics of Dwarfism in the Miniature Horse*. Miniature Horse Club of the Czech Republic. Prague, Czech Republic.

JE Eberth. (January 2013) *Equine Coat Colors*. Miniature Horse Club of the Czech Republic. Prague, Czech Republic.

JE Eberth. (January 2013) *Genetics of Dwarfism in the Miniature Horse*. Miniature Horse Club of the Netherlands. Antwerp, Flanders, Belgium.

JE Eberth. (January 2013) *Equine Coat Colors*. Miniature Horse Club of the Netherlands. Antwerp, Flanders, Belgium.

JE Eberth. (June 2009) *Investigation of Dwarfism Among Miniature Horses using the Illumina Horse SNP50 Bead Chip.* American Miniature Horse Association Board of Directors Meeting 2009. Fort Worth, TX.

JE Eberth. (February 2009) *Genetics of the Equine Oral Cavity*. International Association of Equine Dentistry. 2009 Annual Convention. Orlando, FL.

JE Eberth. (February 2008) *Dwarfism in Miniature Horses*. Miniature Horse Club of Ontario. Toronto, Ontario. Canada.

JE Eberth. (February 2007) *Dwarfism in Miniature Horses*. American Miniature Horse Association 2007 National Convention. Ft Lauderdale, FL.

John Edmund Eberth

John Edmund Eberth

September 11, 2013

Date