MOLECULAR GENETIC STUDIES OF CANINE INHERITED DISEASES INCLUDING SAMS, NEURONAL CEROID LIPOFUSCINOSIS AND DILATED CARDIOMYOPATHY

_____

A Dissertation presented to the Faculty of the Graduate School

University of Missouri

_____

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy in Genetics

_____

by

Douglas H. Gilliam, Jr.

Dr. Gary S. Johnson, Dissertation Supervisor

May 2016

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

MOLECULAR GENETIC STUDIES OF CANINE INHERITED DISEASES INCLUDING SAMS, NEURONAL CEROID LIPOFUSCINOSIS AND DILATED CARDIOMYOPATHY

Presented by Douglas Howard Gilliam, Jr.

a candidate for the degree of Doctor of Philosophy in Genetics

and hereby certify that in their opinion it is worthy of acceptance

_____

Dr. Gary S. Johnson

_____

Dr. James Amos-Landgraf

_____

Dr. Dennis Lubahn

_____

Dr. Dennis O'Brien

_____

Dr. Bimal Ray

_____

ACKNOWLEDGEMENTS

I would like to take this chance to give my unending appreciation to everyone that had a hand in my reaching this goal. I would like to thank Dr. Gary Johnson who believed in me when others didn't and always had a kind word and an open door when they were needed. I would also like to thank the rest of my committee members, Dr. James Amos-Landgraf, Dr. Dennis O'Brien, Dr. Dennis Lubahn and Dr. Bimal Ray; your knowledge and assistance with this work and others would made this possible. I deeply thank you for all you've done and it was a fantastic experience working with all of you.

Thanks also to Dr. Joan Coates, Dr. Stacey Leach and Dr. Martin Katz for assistance with several projects, including those found within this dissertation. Your knowledge and expertise have been of unmeasurable help and I deeply appreciate it. For help both scientifically, personally and professionally, I would like to thank Dr. Elizabeth Bryda; though you may not realize it, your encouragement came at a pivotal time and the fact that I made it here is in part because of it. I sincerely thank you for your guidance.

I would also like to thank all of my labmates past and present, including Dr. Tendai Mutangadura, Dr. Fabiana Farias, Dr. Rong Zeng, Dr. Juyuan Guo, Ana Kolicheski, Jessie Devore and Liz Hansen. A graduate student, like anybody else, is not an island and I thank you for your daily support, assistance and everything you do to make this lab work as a functional unit. In addition to this, I would like to thank the Genetics Area Program and Robert Sanders for miscellaneous help throughout my graduate career.

Last, though certainly not least, I would like to thank my friends and family, particularly my wonderful girlfriend Barbara Trachsel. Words simply can't express how deeply grateful that you're all a part of my life. I thank you for the ceaseless love and support.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

ABSTRACT

The genome of Canis *lupus familiaris*, the domestic dog, is an ideal tool for the study of inherited diseases. Its genome is uniquely suited for the mapping of genes that cause disease, and its reference genome has been published since 2008. Dogs and humans share many of the same diseases, making them an ideal model for the study of comparative genetics. The identification of disease-causing genes in dogs has relevance to human health. We here describe the identification of causal mutations for three different canine diseases in orthologs of genes with orthologous human diseases. We used Next Generation Sequencing in order to generate Whole Genome Sequences for 145 dogs from 69 different breeds with various inherited canine diseases, most of which were suspected to be inherited recessively. We here report the discovery of the causes of; Spinocerebellar Ataxia with Myokymia and Seizures in Russell-Group Terriers, Neuronal Ceroid Lipofuscinosis in Golden Retrievers and Dilated Cardiomyopathy in Standard Schnauzers. While all three were identified via whole genome sequencing, different methodologies and techniques were used to discover and validate the findings. We discuss the importance of these studies and suggest future studies for the projects as well as possible other means of discovering potential canine disease models via Whole Genome Sequencing. Summaries of these three findings are provided in the next three paragraphs.

Juvenile-onset spinocerebellar ataxia has been recognized in Jack Russell Terriers and related Russell group terriers (RGTs) for over 40 years. Ataxia occurs with varying combinations of myokymia, seizures and other signs, thus more than one form of the

disease has been suspected. Our objective was to identify the mutation causing the spinocerebellar ataxia associated with myokymia and/or seizures and distinguish the phenotype from other ataxias seen in the RGTs. We collected DNA samples from 16 RGTs with signs of spinocerebellar ataxia beginning from 2-to-12 months of age, 640 control RGTs, and 383 dogs from 144 other breeds along with the medical records of affected dogs in order to elucidate the genetic cause of this disease. Whole-genome sequencing was performed on one RGT with ataxia and myokymia. Unique, homozygous variants were identified in this dog by comparing its sequence to whole-genome sequences from 81 other canids. We found a missense mutation in the gene coding for the inward rectifying potassium channel Kir4.1 (*KCNJ10:c.627C>G*) that was significantly associated with the disease. Dogs homozygous for the mutant allele all showed spinocerebellar ataxia with varying combinations of myokymia, seizures and other signs. The identification of a mutation in *KCNJ10* in dogs with spinocerebellar ataxia with myokymia and/or seizures (SAMS) clarifies the multiple forms of ataxia seen in these breeds and provides a DNA test to identify carriers. Understanding the role the Kir4.1 channel plays in extracellular potassium buffering by astrocytes could shed light on other conditions characterized by excessive neuronal membrane excitability such as other forms of ataxia, epilepsy and myokymia.

We studied a recessive, progressive neurodegenerative disease occurring in Golden Retriever siblings with an onset of signs at 15 months of age. As the disease progressed these signs included ataxia, anxiety, pacing and circling, tremors, aggression, visual impairment and localized and generalized seizures. A whole genome sequence, generated with DNA from one affected dog, contained a plausibly causal homozygous

mutation: *CLN5:c.934_935delAG*. This mutation was predicted to produce a frameshift and premature termination codon and encode a protein variant, CLN5:p.E312Vfs*6, which would lack 39 C-terminal amino acids. Eighteen DNA samples from the Golden Retriever family members were genotyped at *CLN5:c.934_935delAG*. Three clinically affected dogs were homozygous for the deletion allele; whereas, the clinically normal family members were either heterozygotes (n = 11) or homozygous for the reference allele (n = 4). Among archived Golden Retrievers DNA samples with incomplete clinical records that were also genotyped at the *CLN5:c.934_935delAG* variant, 1053 of 1062 were homozygous for the reference allele, 8 were heterozygotes and one was a deletion-allele homozygote. When contacted, the owner of this homozygote indicated that that their dog had been euthanized because of a neurologic disease that progressed similarly to that of the affected Golden Retriever siblings. We have collected and stored semen from a heterozygous Golden Retriever, thereby preserving an opportunity for us or others to establish a colony of *CLN5*-deficient dogs.

Young-adult onset dilated cardiomyopathy (DCM) segregates in the Standard Schnauzer dog breed in a pattern consistent with an autosomal recessive mode of inheritance. To identify the molecular genetic cause of Standard Schnauzer DCM, DNA from an affected dog was used to generate a whole genome sequence with 31-fold average coverage. Among the sequence variants in this whole genome sequence was a 22-bp deletion and frameshift in *RBM20*, the canine ortholog of a gene previously associated with human DCM. The *RBM20* deletion allele was homozygous in the whole genome sequence of the affected Schnauzer, but absent from 101 whole genome sequences of normal canids or dogs with other diseases. An additional 753 Standard Schnauzers, including 21 with

DCM, were genotyped for the *RBM20* deletion. In this cohort, all 20 of the deletion-allele homozygotes had DCM, only one of the 183 heterozygous dogs had DCM, and all of the reference-allele homozygotes were DCM free. *RBM20* deficiency is known to alter exon-splicing patterns and produced aberrant titin isoforms in a rat model and in human DCM patients. To determine if the Standard Schnauzers with DCM had similar exon-splicing abnormalities, RNA prepared from their left ventricular walls was compared by RT-PCR to similarly prepared RNA from normal adult dogs. Titin transcripts with extensive exon skipping were detected in the normal RNA but not in the RNA from the dogs with DCM. Conversely, titin transcripts with retained exons were detected in the RNA from dogs with DCM but not in the normal-dog RNA. Thus, we have identified a canine model for human *RBM20*-associated DCM.

While all three were identified via whole genome sequencing, different methodologies and techniques were used to discover and validate the findings. We discuss the importance of these studies and suggest future studies for the projects as well as possible other means of discovering potential canine disease models via Whole Genome Sequencing.

CHAPTER I

Introduction

Sequence analysis of mitochondrial DNA from dogs, wild canids and wolves suggests that Canis *lupus familiaris* (the domesticated dog) were derived from wolves[1]. To be precise, they diverged from the gray wolf around 15,000 years ago, though it may have been as much as 100,000 years ago.[2,3] It is likely that this occurred due to humans and wolves sharing common geographical locations.[1,4] Domestication was a mutually beneficial relationship; proximity to humans provided dogs with safe environments and a readily available food source and dogs provided humans with protection and assistance with hunting.

Modern dogs exhibit an incredibly wide range of morphological traits. Through strict selective breeding, humans selected for traits both beneficial to human needs (such as herding, hunting and protection) and for physical characteristics such as size and coat color. To date, selective breeding has produced >400 modern dog breeds which show an astounding range of behavioral, morphologic and physiologic diversity. Many purebred dog breeds are founded by a small number of dogs which best embody the physical and behavioral traits desired for their respective breed, which leads to a low amount of genetic diversity. Population bottleneck events such as World War I and World War II also served to lower many breed's genetic diversity.[5] Development of specific breeds with closed pedigrees can involve inbreeding, which can potentially lead to expression of recessive disease traits along with the desired selected trait(s). It has been found that 50 of the most popular dog breeds harbor at least one heritable physical defect.[6]

Dogs are a popular pet that exhibits large ranges of behavioral, morphological and physiologic phenotypes. They can serve as invaluable animal models for human disease, as many inherited diseases occur in both human and dog populations. According to the Online Mendelian Inheritance in Animals (OMIA) database (http://omia.angis.org.au/home/) there are 653 canine disorders or traits, 263 of which are Mendelian disorders. They estimate that 357 of these canine disorders have potential to serve as useful human disease models. The many inherited canine inherited disorders mimic diseases in a wide range of human disorder categories; including cardiovascular, inflammatory and immune-system, metabolic, neurological and ophthalmologic diseases. Identifying the causal mutations for canine inherited diseases not only provides a diagnostic DNA tool to allow for breeding away from undesirable traits, but could also allow for evaluation of therapeutic intervention and studying of the disease pathogenesis.[7]

The molecular genetic methods for the identification of genetic mutations responsible for canine inherited diseases have advanced rapidly in recent years from candidate gene analysis to linkage analysis; genome-wide association studies (GWAS) and next-generation whole genome sequencing (WGS). Prior to 1997, canine genetics was largely reliant on candidate gene analysis; requiring a previous understanding of the underlying physiology or biochemical pathways involved, or recognition of similarity of disease of known cause that occurs in another species.[8] Candidate gene analysis does not require large families with known phenotypes as does linkage analysis, nor large numbers of cases and controls as does GWAS.

An alternative to candidate gene analysis is linkage analysis, which is used to identify the chromosomal region harboring the disease-causing mutation. DNA markers, distributed evenly across the genome, are used to genotype the members of a pedigree with known phenotype (affected and normal). This functions on the principal that one or more genetic markers will be

located on the same chromosomal region as the causal mutation such that recombination events between markers and the causal mutation are incredibly rare. These markers and their alleles are inherited in a pattern similar to that of the inheritance of the mutant allele responsible for disease phenotype. This method can be further followed up by fine-mapping in which you genotype samples with additional genetic markers from the chromosomal region of interest. This allows the narrowing of chromosomal region containing potential candidate genes.

A third method for the study of canine inherited diseases is the genome-wide association study (GWAS) which identifies the disease chromosomal loci by comparing marker allele frequencies in samples from a cohort of affected phenotype dogs to the marker allele frequencies in samples from a cohort of normal phenotype dogs. The technique makes use of a dense array of SNP markers, distributed across the entire genome, in order to identify the chromosomal region harboring the disease-causing mutation. In 2005, the first successful GWAS using human patients was published; in the past decade, more than 2493 reports on human GWASs covering a 17 categories of disease with a wide range of disease phenotypes with $p \leq 5x10^{-8}$ (http://www.genome.gov/gwastudies/). In 2005, things took a leap forward with the publishing of the first high quality canine reference sequence with 7.8-fold average of the Boxer genome.[9] This advance not only facilitated canine genetic studies; it also enabled the comparison of a new genetic system model to human genetics and genomics.

Inclusion of the genome sequences from an additional 10 dog breeds provided a source of >2.5 million SNPs. The development of distinct breeds and the isolation of distinct genetic pools has altered the genomic structure of the dog to the extent of Linkage Disequilibrium (LD). Within breeds, the LD extends several megabases. Across breeds, LD extends over tens of kilobases, 100 times greater than human LD.[9] As a result, dog GWASs don't need the high density of SNPs

required for human GWASs. Several SNP array platforms have been available over the years, including the 22,363 (Illumina), 49,663 (Affymetrix), 127,000 (Affymetrix) and the 170,000 (Illumina) SNP genotype chips. GWAS studies do not need samples from members of large, multi-generation families, which are required for linkage analysis studies. Compared with the other form of chromosomal mapping, linkage analysis, GWAS studies are successful at finding common variants in populations, but not always the variants that are rare in subpopulations. They have the potential for falsely positive results and need quality control to be useful.[10] A further issue to GWAS studies is the large haplotype blocks and extensive LD found within dog breeds can prove problematic and require analysis of many candidate genes within the disease locus. One method to overcome these shortcomings is that of mRNA sequencing, known as RNAseq or Whole Exome Sequencing (WES).[11] This technique can be used to find causal variants associated with inherited diseases; in 2012 Forman et al. used the technique to find the causal variant for canine cerebellar cortical degeneration in Beagles, an 8 bp deletion in the SPTBN2 gene encoding β-III spectrin.[12] In 2014 Roche Nimblegen released a canine-specific target enrichment kit for use in WES, though its impact on the finding of canine causal mutations has yet to be determined fully.[13]

One more powerful tool exists for the study of canine inherited diseases, next generation Whole-Genome Sequencing (WGS). Heretofore, the industry standard for sequencing accuracy and read length was automated Sanger-sequencing. With the advent of new technologies and sequencing chemistries, however, Next-generation WGS can be used as a rapid and inexpensive means of generating gigabases of DNA sequence from a sample.[14] The use of WGS has led to identification of mutations that cause rare and de novo human diseases which couldn't have been found using traditional methods of mapping[15,16] though limitations exist. They include a higher

rate of systemic errors with respect to that of Sanger-sequencing, difficulty in aligning short reads, variation of sequence coverage depth within unique and repetitive elements, and inability to resolve sequences of repetitive elements, duplicate sequences and indels.[17] It is expected that these issues will be resolved with improvements in technology of the sequencing platforms as well as library chemistries within the next years.

Researchers at the University of Missouri-Columbia have collected >150,000 canine DNA samples representing over 200 different dog breeds. Of these samples, we have over 50 inherited disease phenotypes representing dogs with neurological, ophthalmological, immunological and metabolic disorders. Many of these samples have been contributed by collaborating veterinarians at the University of Missouri-Columbia as well as others around the world. By using advanced genetic and genomic techniques mentioned previously (candidate gene analysis, linkage analysis, GWAS, WGS), we have identified the probable causal mutations for many canine inherited disorders including degenerative myelopathy, dilated cardiomyopathy, gangliosidosis, L-2-hydroxyglutaric aciduria, neuronal ceroid lipofuscinosis, neonatal encephalopathy with seizures, neonatal cerebellar ataxia, primary lens luxation and Spinocerebellar ataxia.[18–30] In the following chapters I will describe the discovery of three genetic mutations that are the probable cause for three canine inherited diseases by using WGS.

CHAPTER II

SAMS

A *KCNJ10* Missense Mutation in Russell Group Terriers with Spinocerebellar Ataxia with Myokymia and/or Seizures

Doug Gilliam, Dennis P. O'Brien, Joan R. Coates, Gary S. Johnson, Gayle C. Johnson, Tendai Mhlanga-Mutangadura, Liz Hansen, Jeremy F. Taylor, and Robert D. Schnabel

Abstract

Juvenile-onset spinocerebellar ataxia has been recognized in Jack Russell Terriers and related Russell group terriers (RGTs) for over 40 years. Ataxia occurs with varying combinations of myokymia, seizures and other signs, thus more than one form of the disease has been suspected. Our objective was to identify the mutation causing the spinocerebellar ataxia associated with myokymia and/or seizures and distinguish the phenotype from other ataxias seen in the RGTs. We collected DNA samples from 16 RGTs with signs of spinocerebellar ataxia beginning from 2-to-12 months of age, 640 control RGTs, and 383 dogs from 144 other breeds along with the medical records of affected dogs in order to elucidate the genetic cause of this disease. Whole-genome sequencing was performed on one RGT with ataxia and myokymia. Unique, homozygous variants were identified in this dog by comparing its sequence to whole-genome sequences from 81 other canids. We found a missense mutation in the gene coding for the inward rectifying potassium channel Kir4.1 (*KCNJ10:c.627C>G*) that was significantly associated with the disease. Dogs homozygous for the mutant allele all showed spinocerebellar ataxia with varying combinations of myokymia, seizures and other signs. The identification of a mutation in *KCNJ10* in dogs with spinocerebellar ataxia with myokymia and/or seizures (SAMS) clarifies

the multiple forms of ataxia seen in these breeds and provides a DNA test to identify carriers. Understanding the role the Kir4.1 channel plays in extracellular potassium buffering by astrocytes could shed light on other conditions characterized by excessive neuronal membrane excitability such as other forms of ataxia, epilepsy and myokymia.

Introduction

Jack Russell Terriers, Parson Russell Terriers, and Russell Terriers are separately registered breeds believed to have descended from the stock of Parson Jack Russell, a 19th century British dog breeder. In this communication we refer to members of these closely related breeds as the Russell group terriers (RGTs). Heritable ataxia has been recognized in RGTs for over 40 years.[31] The wide differences in ages at onset, clinical signs, and histopathologic changes reported for ataxic RGTs[31–36] suggest that more than one heritable ataxia is segregating in these dogs. We have described a cerebellar ataxia that becomes apparent at <2 weeks of age when the RGT puppies begin to walk.[32] We refer to this early onset cerebellar ataxia as primary granular cell degeneration because *post mortem* examination of affected puppies revealed the depletion of cerebellar granular layer cells. Several investigators have described RGTs with a distinct spinocerebellar ataxia with an onset at two to ten months of age and preservation of the granular cells. Many of these RGTs have exhibited myokymia and/or seizures.[33,35–37] Because myokymia is an unusual clinical sign exhibited by human patients with voltage-gated potassium channel gene mutations,[38,39] Van Poucke et al.[40] resequenced part, or all, of four voltage-gated potassium channel genes (*KCNA1*, *KCNA2*, *KCNA6*, and *KCNQ2*) with DNA from RGTs that exhibited myokymia but failed to identify a causal mutation. A later-onset spinocerebellar ataxia has been reported in four RGTs, with an age of onset between 9 and 36 months; no episodes of myokymia

or seizures have been reported in this form of ataxia.[34] Foreman *et al.*[41] identified a missense mutation in *CAPN1* that was strongly associated with a spinocerebellar ataxia phenotype beginning at 6-12 months of age that is sometimes referred to as "late-onset ataxia", though the relationship to the cases reported by Simpson et al.(34) is not clear. No other clinical signs were reported in these cases.[41] The *CAPN1* mutation was discovered with a genome-wide association study followed by the massively-parallel sequencing of affected and control RGT DNA preparations enriched for the disease-associated chromosomal interval by hybridization capture.

We also made use of massively-parallel sequencing, but without prior association mapping, to investigate RGTs with spinocerebellar ataxia with myokymia and/or seizures. In this study, we resequenced the genome of a single affected RGT and compared the detected variants to variants detected in 81 normal canids or dogs with other diseases that had also been resequenced. Among the sequence variants that were unique to the affected RGT was a homozygous transversion that predicts a missense mutation in *KCNJ10*, a gene that codes for a voltage-gated potassium channel Kir4.1. The *KCNJ10* mutation was found to be strongly associated with RGT spinocerebellar ataxia which enabled us to describe the spectrum of disease signs associated with the genetically defined RGT ataxia.

Materials and Methods

The University of Missouri Animal DNA Repository contains samples from >100,000 individual dogs including 2,868 from RGTs. At least 38 of the RGTs represented in the collection had a clinical history of cerebellar ataxia. The onset of ataxia was less than two weeks of age for 21 of these RGTs, strongly suggesting that these dogs had primary granular cell degeneration. The

samples from these dogs were excluded from this study as was a sample from an ataxic RGT diagnosed with canine distemper. Included in this study were 16 samples from ataxic RGTs with an onset of ataxia between 2 and 12 months of age, 640 samples from control RGTs randomly selected from among the RGTs without a known clinical history of ataxia, and 383 samples from control dogs from 144 non-RGT breeds. All animal studies were conducted with approval of the University of Missouri, Animal Care and Use Committee and the informed consent of the owner.

The 16 cases were collected by JRC between 1996 and 2013 based on owner or clinician reports of spinocerebellar ataxia. When available, clinical examination and videotapes of the affected dogs were reviewed, and necropsies were performed when permitted. Other signs such as myokymia, neuromyotonia, excessive facial rubbing, hearing loss, or seizures were also recorded. When available, the final outcome for the dogs was assessed by phone calls with the owners and referring veterinarians.

A DNA sample from an RGT with spinocerebellar ataxia and myokymia and an extensive clinical evaluation was selected for whole-genome resequencing. Two DNA libraries (fragment sizes of approximately 300 bp and 400 bp) were prepared for paired-end sequencing with a commercial kit (Illumina TruSeq DNA sample preparation kit) and each library sequenced on a single flow-cell lane with a massively parallel DNA sequencer (Illumina HiSeq 2000) at the University of Missouri, DNA Core Facility. Initial quality control on the sequence data involved removing exact duplicate sequence read pairs using a commercial software (NextGENe® v2.3.2, SOFTGENETICS®) and adapter trimming using custom Perl scripts. Unique adapter-trimmed reads were error corrected using MaSuRCA v1.9.5 software.[42] Error corrected reads were aligned to the reference CanFam3.1 genome assembly and variants called using a commercial software (NextGENe® v2.3.2, SOFTGENETICS®). Variant calls were further processed using

custom Perl scripts to filter likely false positives and then the variant calls were uploaded to a custom PostgreSQL database which contained the variant calls for an additional 81 canid samples. Using custom SQL scripts we identified variants that fit an autosomal recessive mode of inheritance such that the case is homozygous for an allele not observed (in either heterozygous or homozygous form) in any of the remaining 81 canids used as controls. This candidate variant list was further filtered to include only variants predicted to alter the amino acid sequence of protein coding genes. The sequence data from the 81 canid genomes used as controls were generated in a similar manner. The controls represented dogs for which the breed and disease phenotype status were known and included: 37 genomes from our group; three wild canid genomes provided by University of California, Los Angeles; 28 genomes provided by the Institute for Translational Genomic Research; 10 genomes provided by North Carolina State University; 3 genomes provided by the University of Pennsylvania. A detailed description of the data processing pipeline is in preparation and will be published elsewhere.

An apparent *KCNJ10:c.627C>G* sequence variant (GenBank accession XM_545752.3) was verified by PCR amplification with primers *5'-GCCAACATGCGGAAGAGCCT-3'* and *5'-TCGAAGTCACCCTCGCCACT-3'* and automated Sanger sequencing. The effect of the mutation on protein function was estimated using a tool for the prediction of the impact of an amino acid substitution on the structure and function of the orthologous human protein (PolyPhen-2) available at http://genetics.bwh.harvard.edu/pph2/.[43] A TaqMan allelic discrimination assay[44] was used to genotype DNA samples from individual dogs at *KCNJ10:c.627C>G.* For this assay, the PCR primer sequences were *5'-CGCCAACATGCGGAAGAG-3'* and *5'-GGTGGGTCTGAAGCAGCTT-3'* and the competing probes were *5'-VIC-CCTCCTCATCGGCTGC-MGB-3'* (reference allele) and *5'-FAM-*

CTCCTCATGGGCTGC-MGB-3' (mutant allele). A second TaqMan allelic discrimination assay was used to genotype DNA samples at *CAPN1:c.344C>T* (GenBank accession XM_540866.3). For this assay, the PCR primer sequences were *5'-GAGGGTGAGGGAGGCAAT-3'* and *5'-GGGTCCCTCCCATCCCA-3'* and the competing probes were *5'-VIC-AGAAGCCAACAGTCCC-MGB-3'* (reference allele) and *5'-FAM-AGAAGCCAATAGTCCC-MBG 3'* (mutant allele). PCR amplification with primers 5'-AGGTTCTCTCTTTAGGAGCA-3' and 5'-CACCACATTCAAAGGGCTA-3' and automated Sanger sequencing, was used to genotype another apparent sequence variant, *KCNH7:c.C1180A,* p.394R>S (GenBank accession XM_005640254.1).

Electromyograms (EMG), motor nerve conduction measurements, and brainstem auditory evoked potentials were evaluated in two RGTs with myokymia according to previously described methods.[45,46] One-sided EMG recordings were made from the axial and appendicular muscles of the thoracic and pelvic limbs, as well as cranial and axial muscles. A commercially available electrophysiologic unit was used for all electrodiagnostic recordings.   All electrophysicologic recordings were performed under general anesthesia.

The brains of 3 RGTs were examined by magnetic resonance imaging (MRI) under general anesthesia. MRI was performed with a 1.5-Tesla unit or a 3-Tesla unit. Pulse sequences were selected to obtain 3D T1-weighted sequences or standard T1- and T2-weighted sequences in three planes. Complete necropsies were performed on 9 ataxic RGTs. The thoracic spinal cords were dissected, fixed and embedded in paraffin. Sections of these samples were stained with hematoxylin and eosin plus either luxol fast blue-periodic acid Schiff or luxol fast blue-cresyl violet to detect neuronal fiber and myelin loss. In addition, spinal cord sections were immune-stained for neurofilament medium and for fibrillary acid protein to detect astrogliosis.

Results:

Resequencing of the individual ataxic RGT case yielded an average of 33.9X coverage depth of the CanFam3.1 reference genome with 99.4% of the mappable genome covered by at least one sequence read. A total of 7,440,532 putative variants between the case and the reference genome were detected. Of these, 20,652 were predicted to alter the amino acid sequence of proteins with only 23 of these homozygous in the case and the alternate allele not being present in any of the 81 control dog genomes. Two of the 23 sequence variants that remained after data processing were in potassium channel genes. One was a *KCNJ10:c.627C>G* transversion (Figure 1A) that predicts a KCNJ10:p.I209M amino acid substitution. The presence of this missense mutation in the affected RGT was confirmed by PCR amplification and automated Sanger sequencing (Figure 1B). PolyPhen-2 predicted that this was "probably damaging" with a score of 0.979 on a 0-1 scale with 1 being most damaging.[43] The isoleucine at position 209 and the surrounding amino acids are conserved in all available orthologous mammalian sequences and in most vertebrate sequences, however the isoleucine is replaced by leucine in some fish species. Fourteen of the 16 samples from ataxic RGTs were homozygous for the *KCNJ10:c.627G* allele. The other two ataxic RGTs were *c.627C* homozygous siblings. The 640 randomly selected RGT samples were either heterozygous at *KCNJ10:c.627C>G* (n = 63) or homozygous for the reference *c.627C* allele (n = 577). Thus, among RGTs, there is a significant association between the ataxia phenotype and homozygosity for the *KCNJ10:c.627G* allele (p = 8.27 x 10^{-12}; Fisher's Exact Test). All genotyped DNA samples from 383 individual dogs from 144 non-RGT breeds were *KCNJ10:c.627C* homozygotes.

The other candidate sequence variant that remained after data processing was a *KCNH7:c.1180C>A* transversion that predicts a p.394R>S substitution (GenBank accession

XM_005640254.1). The *KCNH7:c.1180A* allele was only detected in the RGT DNA used for whole genome sequencing.   The other 15 ataxic RGTs were homozygous for the *KCNH7:c.1180C allele.*   The *CAPN1:c.344T* allele, previously associated with a late-onset ataxia of RGTs [41], was not detected in any of the 16 ataxic RGT samples or in any of the 44 genotyped control RGT samples.

Because 14 of the ataxic RGTs in this study were homozygous for the *KCNJ10:c.627G* allele, it is now possible to describe the associated disease phenotype in this genetically defined subpopulation of RGTs. Three of the dogs were female, one dog was a neutered male and 10 dogs were intact males. All 14 affected dogs showed marked cerebellar ataxia typical of spinocerebellar ataxia. The median age at onset of cerebellar ataxia in 12 dogs was 3 months (range, 2 to 6 months). The age of onset could not be ascertained for 2 dogs. Two dogs showed only signs of cerebellar ataxia without myokymia or seizure. Ten dogs showed muscle fasciculation characteristic of myokymia. The median age at onset for myokymia was 6 months (range, 3 to 8 months). Age of onset for myokymia could not be determined in 1 dog. Episodes of neuromyotonia were documented in 5 dogs. One of the dogs with myokymia also had a history of seizures while two dogs had seizures but no reported myokymia. The median age at onset for seizure was 2 months (range, 2 to 4 months). Excessive facial rubbing was not well documented in the histories, but it was noted in two dogs. Hearing loss was not evident in any of the dogs. These results are summarized in Table 1.

Brain MRIs in three *KCNJ10:c.627G* homozygous dogs were interpreted as normal. In two dogs that underwent electrodiagnostic testing, EMG showed myokymia in the form of semi-rhythmic bursts of doublet, triplet or multiplet discharges of a single motor unit (Figure 2) a third dog with no clinical evidence of myokymia showed no abnormalities on EMG. Nerve conduction

velocities and brainstem auditory evoked potentials were within normal limits. The diagnosis of spinocerebellar ataxia was confirmed by histopathologic examination of spinal cord sections in 9 RGTs, which revealed a bilateral myelopathy with loss of axons and myelin and astrogliosis in the dorsal and lateral portions of the lateral funiculus and in the ventral funiculus (Figure 3).

Discussion

Cerebellar ataxia was included in the clinical histories of 38 RGTs represented in our DNA collection. For 21 of these RGTs the ataxia became apparent before they were 2 weeks of age. It is likely that these puppies had the previously described primary granular cell degeneration,[32], but necropsies were not performed on all cases. The mutation responsible for primary granular cell degeneration is currently unknown. In the absence of a diagnostic DNA test, we cannot exclude the possibility that other diseases such as Dandy-Walker malformation caused the early-onset ataxia in some of these RGTs. Samples from these RGTs with early-onset ataxia were not included in this study. Another of the ataxic RGTs was diagnosed with canine distemper encephalitis and was not further studied. The remaining 16 RGTs were genotyped for the *CAPN1:c.344C>T* mutation reported to cause later-onset spinocerebellar ataxia in European RGTs.[41] All 16 were homozygous for the reference *CAPN1:c.344C* allele, as were all 44 of the genotyped control RGTs. This suggests that *CAPN1*-associated spinocerebellar ataxia may be more common in Europe than in North American where most of our samples originated.

Of the 7,440,532 sequence variants in the RGT whole genome sequence, only 23 were homozygous, absent from the 81 control whole genome sequences and predicted to alter the primary structure of the encoded protein. Among these 23, missense mutations in *KCNJ10* and *KCNH7* were considered most likely to be causal because they encode potassium channel genes

and mutations in other potassium channel genes are known to cause myokymia in human patients.[38,39] The *KCNH7* variant appears to be an incidental finding since it occurs in only one of the ataxic RGTs. Fourteen of the 16 ataxic RGTs in our study were homozygous for the *G* allele of the *KCNJ10:c.627C>G* transversion; whereas, none of the 640 randomly selected control RGTs was a *c.627G* homozygote.

The *KCNJ10:c.627C>G* mutation predicts a p.I209M amino acid substitution in Kir4.1, an inward rectifying potassium channel that is expressed in glia, stria vascularis of the inner ear, renal distal convoluted tubules, and gastric parietal cells.[47] In glia, Kir4.1 is responsible for the highly negative resting membrane potential and the large K^+ conductance at rest.[48] In contrast to most potassium channels, inward rectifying channels primarily conduct K^+ into cells.[49] Because Kir4.1 is one of the "weak" inward rectifiers which does not inactivate at depolarized potentials, it can also conduct K^+ out of astrocytes. These properties allow astrocytes to spatially buffer the increased extracellular K^+ concentrations produced by neuronal activity (Figure 4). Interference with this buffering capacity can lead to increased extracellular K^+ and neuronal membrane depolarization. Glutamate transporters located in astrocytes also regulate neuronal excitability by removing glutamate at the synapse. Glutamate transport is strongly facilitated by the negative membrane potential, and impaired by decreased Kir4.1 function.[48] Kir4.1 is also important in the development of oligodendroglia and the maintenance of myelin although the mechanisms are unclear.[48,50]

Mutations in human *KCNJ10* have been identified in patients with two autosomal recessive hereditary disease syndromes with overlapping signs: SeSAME (seizures, sensorineural deafness, ataxia, mental retardation, and electrolyte imbalances)[51] and EAST (epilepsy, ataxia, sensorineural deafness, and tubulopathy).[52] *Kcnj10* knock-out mice have severe ataxia, hind limb

weakness, and die prematurely, while mice with a conditional knock-out of *Kcnj10* in glia have stress-induced seizures and severe ataxia.[48,50] Several of the signs exhibited by patients with SeSAME and EAST and by the knockout mouse models are shared with the *KCNJ10:c.627G* homozygous RGTs (Table 2). The similarities between the clinical signs exhibited by the affected RGTs, the human patients with the *KCNJ10*-deficiency syndromes, and the *Kcnj10* knockout mouse models together with the highly significant association between the RGT spinocerebellar disease phenotype and homozygosity for the *KCNJ10:c.627G* allele strongly suggest that the *KCNJ10:c.627C>G* transversion is the mutation responsible for the autosomal recessive spinocerebellar ataxia in the 14 ataxic RGTs.

Thus the hereditary ataxias in RGTs can be clearly divided into at least three groups: 1) neonatal, granular cell degeneration,[32] 2) "late-onset" (6-12 months) spinocerebellar ataxia without other clinical signs associated with *CAPN1:c.344T*[41] and 3) spinocerebellar ataxia with myokymia and/or seizures associated with *KCNJ10:c.627G* . Further studies will be necessary to determine what histologic changes are associated with the *CAPN1:c.344T*-associated ataxia and whether cases with an onset of ataxia over 12 months of age[34] are also associated with that mutation. Since our inclusion criteria focused on RGTs between 2 and 12 months of age, we cannot rule out the possibility that earlier or later onset ataxia, other breeds, or other clinical signs may also be associated with the *KCNJ10:cC>G*. mutation. Neither the *CAPN1:c.344T* allele nor the *KCNJ10:c.627G* allele were detected in two RGT siblings with a history of only spinocerebellar ataxia indicating that there may be additional genetic factors responsible for ataxia in RGTs. No further information was available on these dogs, however, so acquired disease could not be excluded.

The identification of homozygosity at *KCNJ10:c.627G* as a probable cause of ataxia in RGTs allowed us to describe the disease phenotype in a genetically defined cohort. Episodes of myokymia were observed in 64% (9/14) of the *KCNJ10:c.627G* homozygous RGTs. Myokymia and neuromyotonia are common clinical signs in previous descriptions of spinocerebellar ataxia in RGTs.[33,35,53] The EMG recordings we generated from two *KCNJ10:c.627G* homozygous RGTs are similar to those from other studies of RGTs[33,35,37,53] and confirm that the episodic subcutaneous muscle rippling is indeed myokymia. Myokymia and neuromyotonia have not been reported in RGTs with primary granular cell degeneration[32] or in association with the *CAPN1* missense mutation.[41] In addition, we have not found reports of myokymia or neuromyotonia among the clinical signs of SeSAME/EAST patients or in descriptions of the *Kcnj10* knockout mouse phenotypes.[48,50–52] Mutations in human *KCNA1* and *KCNQ2* have been associated with myokymia along with episodic ataxia (*KCNA1*) or benign neonatal seizures (*KCNQ2*).[38,39] No unique homozygous coding variants were identified in these genes in our RGT whole genome sequence. It is likely that the *KCNJ10* mutation causes both the ataxia and excessive membrane excitability resulting in repetitive firing of motor neurons and myokymia.

Three of the 14 *KCNJ10:c.627G* homozygous RGTs (21%) had seizures which are also characteristic of SeSAME/EAST in people. This is lower than the frequency of seizures (37%) reported in a European cohort of ataxic RGTs.[36] During severe episodes of myokymia (neuromyotonia), the dogs hold their limbs in rigid extension and are unable to rise from lateral recumbency. These episodes could be confused with a generalized seizure, but they last for up to several hours and the dog remains conscious during the episode. In contrast, the seizure episodes observed by the author (DPO) were brief, generalized, tonic-clonic seizures with loss of

consciousness and autonomic signs. As in the mice with the conditional *Kcnj10* knockout in glia,[48] the seizures in that dog appeared to be elicited by stress.

Varying degrees of hearing impairment have been reported in SeSAME/EAST patients, but the exact cause has not been determined.[51,52] *Kcnj10* is expressed in the stria vascularis of the cochlea where it is thought to play a role in the maintenance of the endonuclear potential.[54] *Kcnj10* knockout mice have cochlear degeneration and sensorineural deafness.[55] Deafness has not been reported in RGTs with ataxia and hearing impairment was not included in the clinical histories of any of the 14 *KCNJ10:c.627G* homozygotes in this study. Others have recorded abnormal brainstem auditory-evoked potentials from ataxic RGTs with ataxia and myokymia.[35,36] Nonetheless, the presence of waveforms indicates that these dogs have at least some cochlear function and the abnormal wave forms more likely correlate with histologic changes in central auditory pathways.[31,36]

The excessive face rubbing reported in some RGTs with ataxia may represent a sensory paresthesia.[33,37] Paresthesias were not reported in SeSAME/EAST.[51,52] An itching sensation has been reported in humans with myokymia[56] and scratching in dogs with myokymia from other causes,[37] but excessive facial rubbing occurred in one *KCNJ10:c.627G* homozygote that did not show myokymia. In mice, inactivation of Kir4.1 in the trigeminal ganglia leads to pain behaviors suggesting that the Kir4.1 deficiency may also affect sensory functions.[57]

KCNJ10 is also expressed in the distal convoluted tubules of the human kidney where it is essential for the maintenance of the resting electrical potential.[58] Loss of Kir4.1 function in SeSAME/EAST syndromes leads to hypokalemic metabolic alkalosis and hypomagnesemia.[51,52] Altered electroretinograms from EAST patients and nullizygous *Kcnj10* knockout mice are

attributed to impaired Kir4.1 function in retinal Muller cells.[59,60] Future examinations of the renal and retinal functions of RGTs that are homozygous for *KCNJ10:c.627G* may reveal similar disease signs.

Kcnj10 knockout mice show hypomyelination, vacuolation, and axonal swellings in the spinal cord and cerebellum, predominantly in the white matter adjacent to the gray matter.[48,50] We and others have found similar, but less severe, changes in RGT with ataxia.[33,36,37] No *post-mortem* findings have been reported in people with SeSAME/EAST. Some affected children show hyperintensity of cerebellar roof nuclei and mild atrophy of the cerebellum, corpus callosum, brainstem and spinal cord on MRI.[61] Such changes were not apparent in the RGTs imaged in this study nor in many of the human cases.[52]

Conclusions

To distinguish the autosomal recessive spinocerebellar ataxia investigated in this report from the other ataxias that occur in RGTs, we propose the acronym "SAMS" for Spinocerebellar Ataxia with Myokymia and/or Seizures. The identification of the *KCNJ10:c.627C>G* mutation as the probable cause of SAMS confirms the suspicion that multiple forms of hereditary ataxia segregate in these breeds[36,41], and it will provide the basis for DNA tests that can facilitate the diagnosis of this disease in RGTs. In addition, RGT breeders can utilize DNA tests for the *KCNJ10:c.627G* and the *CAPN1:c.344T* alleles to identify carriers of either alleles and avoid producing dogs with spinocerebellar ataxia associated with either of these mutations. As discussed above, however, two RGT siblings in which the client reported signs of spinocerebellar ataxia did not have either of the identified mutant alleles. Further research is

needed to clarify whether additional genetic causes of spinocerebellar ataxia segregate in these breeds and to identify the mutation responsible for the neonatal, primary granule cell degeneration. Understanding how mutations in *KCNJ10* affect the function of the Kir4.1 inward rectifying potassium channel and the buffering of extracellular potassium by astrocytes may shed light on excessive neuronal membrane excitability in other forms of ataxia, epilepsy or myokymia.

Others have used whole genome sequencing to identify causal *CUBN* mutations in dogs with Imerslund-Grasbeck syndrome.[62,63] *KCNJ10:c.627G* may be the first reported disease-causing canine allele that was identified by whole genome sequencing even though the identity of the causal gene was not apparent before sequence generation was undertaken. The strategy of filtering sequence variants to identify only coding variants that are homozygous in the sequenced dog and are absent from the whole genome sequences of canids that do not have the targeted disease has facilitated identification of at least 7 other apparently causal mutations. Manuscripts describing these mutations are at various stages of completion. Nonetheless, this strategy will only work for rare recessive diseases caused by mutations that are predicted to alter the primary structure of the encoded protein and that occur in association with correctly annotated exons. We expect that improved genome annotation and the development of algorithms designed identify other types of causal sequence variants will soon make whole genome sequencing an even more effective strategy identify the causes of heritable canine diseases.
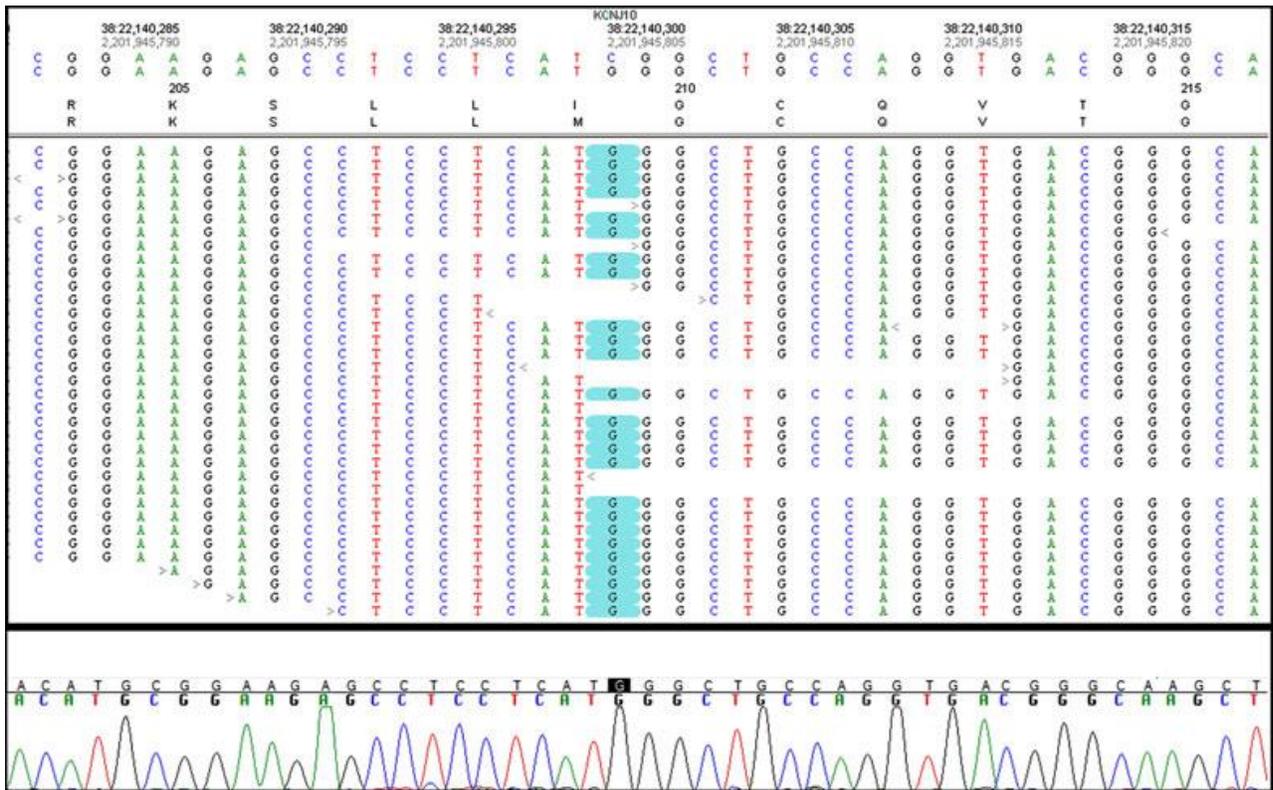
Figure 1: (A) Sequence reads from the RGT whole genome sequence aligned to a segment of canine chromosome 38 showing the homozygous *C>G* transversion in *KCNJ10*. (B) Automated Sanger sequence confirms the *C>G* transversion in DNA from the ataxic RGT.
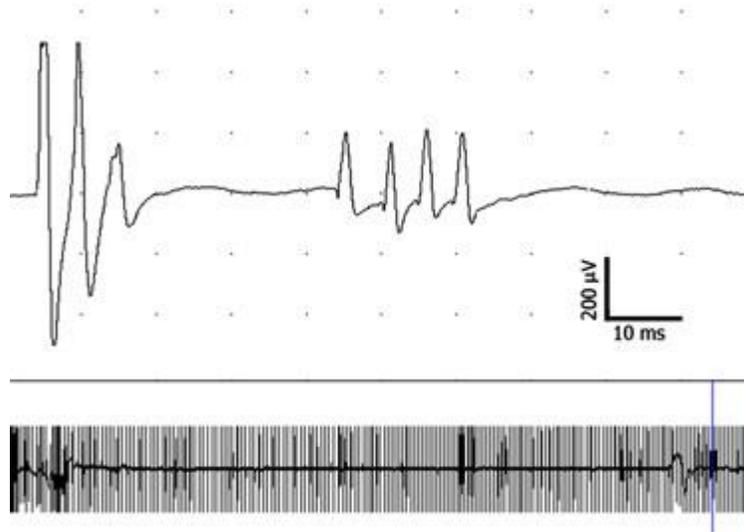
Figure 2: EMG from a RGT with myokymia. (A) Two waveforms consisting of trains of 3-4 motor unit action potentials. (B) 30 second sweep of the same recording showing the regularity of the firing pattern.
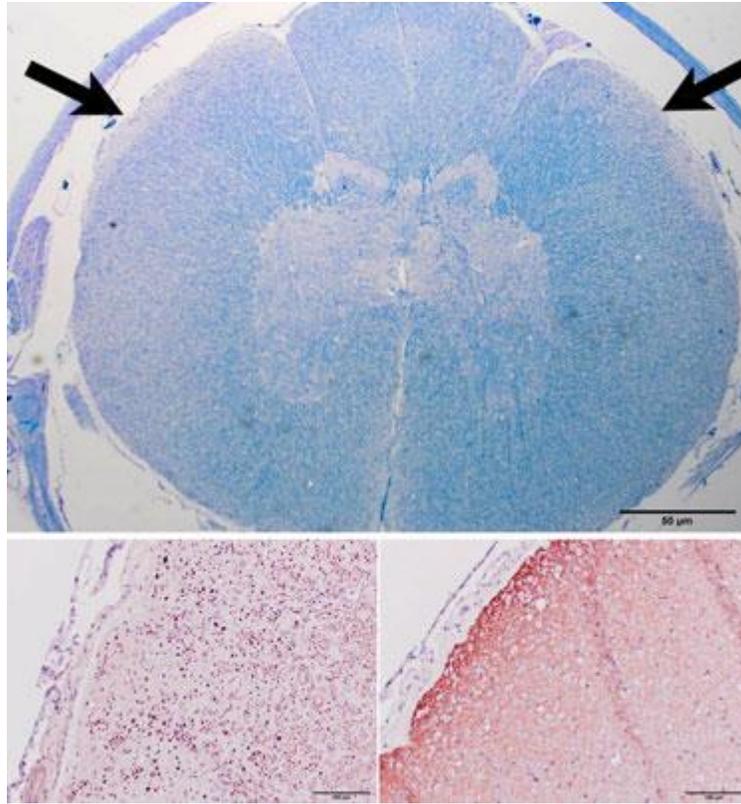
Figure 3: Histopathology of the spinal cord of an affected RGT. (A) Pallor of the dorsal portion of the lateral white matter. Luxol fast blue stain. (B) Immunohistochemical staining for neurofilament medium in an area of axonal loss. (C) Immunohistochemical staining for glial fibrillary acidic protein showing regional gliosis.
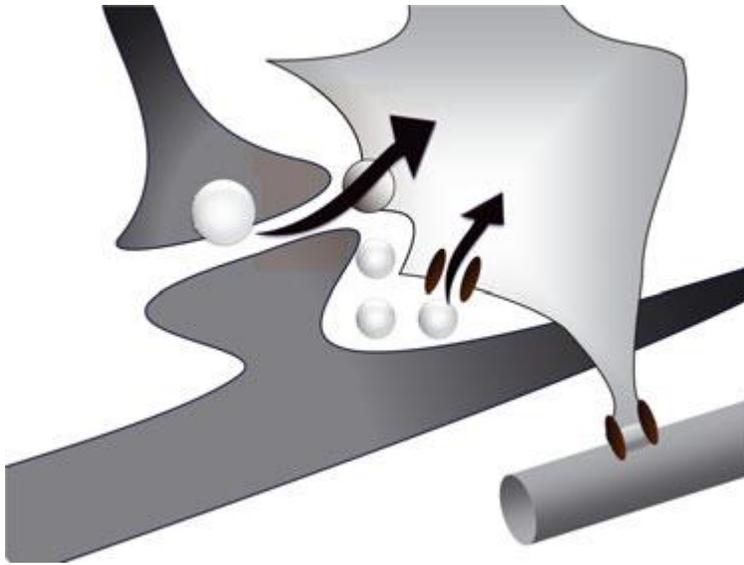
Figure 4: The Kir4.1 channel permits buffering of extracellular potassium by astrocytes. It also contributes to their highly negative resting membrane potential. This negative resting potential is essential for the uptake of glutamate from the synapse through glutamate transporters in astrocytes.

Table 1: Summary of clinical spectrum of SAMS in Russell Group Terriers homozygous for *KCNJ10:c.627G*

Case	Sex	Age at Death	Age at Onset of Signs			Diagnostic tests*
			Cerebellar ataxia	Myokymia	Seizure	
1	F	Unknown	Unknown	none	none	C,H
2	F	4 months	2.25 months	none	none	C,N,M,H
3	M	23 months	3 months	3 months	none	C
4	MC	23 months	2.75 months	4 months	none	C,V,N,M,H
5	M	2 months	4 months	4 months	4 months	C,H
6	M	14 months	3.5 months	3.5 months	none	C,H, V, N
7	F	14 months	2 months	8 months	none	C,H, V
8	M	17 months	Unknown	Unknown	none	C,H
9	M	10 months	6 months	6 months	none	C,H, V
10	M	13 months	4 months	6 months	none	C,V,H, N
11	M	11 months	2 months	none	2 months	C,V,N,M,E,H
12	M	13 months	4 months	none	4 months	C
13	M	Alive (10 months)	2-3 months	7 months	none	C,V,N,E
14	M	Alive (10 months)	2-3 months	7 months	none	C

Unknown: the clinical signs were present but the age at onset could not be ascertained

*C = client description, V = video examined by a neurologist, N = neurologic exam by a neurologist, M = magnetic resonance imaging, E = electrodiagnostic tests, H =histopathology.

Table 2: Signs associated with *KCNJ10* mutations in different species

	Cerebellar ataxia	Seizures	Myokymia	Deafness	Mental Retardation	Electrolyte Imbalances	Weakness /paralysis	Premature Lethality
SeSAME Patients (51)	+	+	-	+	+	+	-	-
EAST Patients (52)	+	+	-	+	-	+	-	-
Knockout Mice (50,55)	+	-	-	+	NE	+	+	+
Conditional Knockout mice (48)	+	+	-	NE	NE	NE	+	+
Spinocerebellar ataxia with myokymia and seizures in RGTs	+	+	+	-	-	NE	-	*

NE = not evaluated.

* Most dogs were euthanized due to poor quality of life

CHAPTER III

Neuronal Ceroid Lipofuscinosis

Golden Retriever dogs with neuronal ceroid lipofuscinosis have a two-base-pair deletion and frameshift in *CLN5*

Doug Gilliam, Ana Kolicheski, Gary S. Johnson, Tendai Mhlanga-Mutangadura, Jeremy F. Taylor, Robert D. Schnabel, Martin Katz

Abstract

We studied a recessive, progressive neurodegenerative disease occuring in Golden Retriever siblings with an onset of signs at 15 months of age. As the disease progressed these signs included ataxia, anxiety, pacing and circling, tremors, aggression, visual impairment and localized and generalized seizures. A whole genome sequence, generated with DNA from one affected dog, contained a plausibly causal homozygous mutation: *CLN5:c.934_935delAG*. This mutation was predicted to produce a frameshift and premature termination codon and encode a protein variant, CLN5:p.E312Vfs*6, which would lack 39 C-terminal amino acids. Eighteen DNA samples from the Golden Retriever family members were genotyped at *CLN5:c.934_935delAG*. Three clinically affected dogs were homozygous for the deletion allele; whereas, the clinically normal family members were either heterozygotes (n = 11) or homozygous for the reference allele (n = 4). Among archived Golden Retrievers DNA samples with incomplete clinical records that were also genotyped at the *CLN5:c.934_935delAG* variant, 1053 of 1062 were homozygous for the reference allele, 8 were heterozygotes and one was a deletion-allele homozygote. When contacted, the owner of this homozygote indicated that that

their dog had been euthanized because of a neurologic disease that progressed similarly to that of the affected Golden Retriever siblings.   We have collected and stored semen from a heterozygous Golden Retriever, thereby preserving an opportunity for us or others to establish a colony of *CLN5*-deficient dogs.

Introduction

A subset of the lysosomal storage diseases has been classified as neuronal ceroid lipofuscinoses (NCLs) because patients with these disorders exhibited abnormal accumulations of ceroid or lipofuscin-like autofluorescent materials within the lysosomes of their brains, retinas, and other tissues.[64] Among the various subtypes of NCL, the accumulated materials have characteristic ultrastructural appearances often referred to as GRODs (granular osmiophilic deposits) or as fingerprint, curvilinear, or rectilinear profiles.[65] The clinical manifestations of the NCLs include progressive visual impairment, declining cognitive and motor functions, seizures, generalized brain atrophy and premature death.[66] NCL-causing mutations were first identified in *PPT1* and *CLN3* and reported in 1995.[67,68] Since then, numerous NCL patients have been found to harbor mutations not only in *PPT1* and *CLN3* but also in any of 11 additional NCL-associated genes (http://www.ucl.ac.uk/ncl/mutation.shtml).[69] A minority of NCL cases have had adult onsets and the disorder in these adult-onset cases is sometimes referred to as Kufs disease.[70] The vast majority of NCL cases had infantile or juvenile onsets. Most or all of the childhood-onset NCLs are recessive traits. Estimates of the prevalences of the individual NCLs have varied greatly among different populations,[71] however taken together the infantile- and juvenile-onset NCLs comprise one of the most prevalent heritable neurodegenerative diseases of children.

In addition to the human patients, NCLs have been described in other mammalian species as well as in birds. One or more cases of NCL have been reported in cats, cattle, dogs, ducks, ferrets, goats, horses, monkeys, parrots, and pigs.[72] Among the animal species, NCLs have most frequently been reported in dogs. Previous studies have identified 10 different mutations in the canine orthologs of 8 of the 13 genes associated with human NCL as shown in Table 1.[22,23,25,26,73–81] Prior to the discovery of the causative mutations, NCL was common in three dog breeds: the Tibetan Terrier, the Border Collie, and the American Bulldog.[82–85] The identification of the mutations responsible for the NCL in each of these breeds[23,75,77,78] has led to the development of DNA tests that revealed the identity of asymptomatic heterozygous mutation carriers and, thereby, have assisted dog breeders in their efforts to reduce the incidence of NCL in these breeds.[83,84] The NCLs in most of the other dog breeds appear to be rare.[22,25,26,73,74,76] DNA-based screening to identify mutation carriers would not be cost-effective in these breeds except in breeding lines where dogs with the disease have already been identified. Nonetheless, the identification of rare NCL-causing mutations can still be important when it leads to the establishment of canine models for NCL patients with mutations in the orthologous human gene. These canine models can be used to study disease mechanisms or to evaluate therapeutic interventions. For instance, a research colony founded from dogs with NCL due to a rare *TPP1* mutation[74] provided subjects for preclinical trials that established the efficacy of enzyme replacement therapy for *TPP1*-deficient dogs.[86–88] This helped pave the way for an ongoing clinical trial with enzyme replacement therapy for *TPP1*-deficient children that appears to be slowing their disease progression (https://www.clinicaltrials.gov/ct2/show/NCT01907087).

We have examined related Golden Retrievers with an apparently rare progressive neurological disease suspected to be an NCL based on clinical signs. Postmortem evaluations of brain and

retinal tissues confirmed that the disorder was indeed a form of NCL. In addition, we generated a whole genome sequence using DNA from one of the affected dogs which led to the identification of the mutation responsible for the disorder.

Materials and Methods

Subject dogs

All studies were conducted with approval of the University of Missouri, Animal Care and Use Committee and with informed consent of the dogs' owners. We were contacted by a breeder of Golden Retrievers who had produced two litters of 9 puppies each from the same healthy parents. Both of the parents had the same healthy mother but each had a different healthy father (Fig. 1). At the time we were contacted, 4 of the dogs from the first litter had exhibited progressive neurological signs characteristic of NCL. Two of these dogs were euthanized at approximately 30 to 31 months of age due to the severity of their neurological signs. We were unable to obtain postmortem tissue samples from either of these dogs. Two additional dogs from the same litter exhibited similar progressive neurological signs and we were able to obtain tissue samples after the female proband was euthanized at 31 months and her female sibling was euthanized at 34 months of age.

The neurological signs in the affected dogs from the first litter became apparent at approximately 13 months of age, followed by slow increases in severity until the dogs were euthanized. The signs in the affected dogs from the first litter became apparent at approximately 13 months of age, followed by slow increases in severity until the dogs were euthanized. The first sign recognized in the affected dogs was a loss of coordination causing the

dogs to bump into objects and have difficulty climbing stairs. The impairments in coordination were particularly evident when the dogs were excited. The loss of coordination progressed with age and eventually became apparent even when the dogs were calm. Also among the early neurological signs were anxiety and agitation, accompanied by long periods of constant pacing and circling starting at approximately 15 months of age. At about the same age the affected dogs lost the ability to recognize or respond to previously learned commands and behaviors and began to exhibit compulsive behavior such fixation on a toy. Mild seizure-like activity such as snapping at the air (fly biting) and gum smacking began as early as 15 months of age. Seizure activity became progressively more severe and began to involve the whole body after about 18 months of age with some episodes lasting for up to an hour. In addition, trance-like behavior started at 18 months of age. Visual impairment in both bright and dim light was apparent by 18 months of age and became quite profound. In the later stages of the disease the affected dogs became aggressive toward people. Only one of the 9 dogs in the second litter had begun to exhibit the neurological signs described above when his health was last ascertained at 17 months of age. The remaining dogs in this litter appeared to be neurologically normal at this age.

We obtained blood anti-coagulated with EDTA from 18 members of the Golden Retriever family including the common grandmother, the maternal grandfather, both parents, the clinically affected proband, her clinically affected female littermate, three clinically normal littermates of the proband and all 9 siblings from the second litter. DNA was isolated from these blood samples as previously described.[22] In addition to the DNA from the Golden Retriever family, we genotyped DNA at a causal mutation candidate locus (see below) from 1062 Golden Retrievers randomly selected without regard to phenotypic information from our archived DNA samples (n = 530) and from the CHIC DNA bank (n = 532; http://www.caninehealthinfo.org/dnabank.html).

In addition, we genotyped archived DNA samples from 143 dogs representing 99 breeds other than Golden Retriever.

We obtained blood anti-coagulated with EDTA from 18 members of the Golden Retriever family including the common grandmother, the maternal grandfather, both parents, the clinically affected proband, her clinically affected female littermate, three clinically normal littermates of the proband and all 9 siblings from the second litter. DNA was isolated from these blood samples as previously described.[22] In addition to the DNA from the Golden Retriever family, we genotyped DNA at a causal mutation candidate locus (see below) from 1062 Golden Retrievers randomly selected without regard to phenotypic information from our archived DNA samples (n = 530) and from the CHIC DNA bank (n = 532; http://www.caninehealthinfo.org/dnabank.html). In addition, we genotyped archived DNA samples from 143 dogs representing 99 breeds other than Golden Retriever.

Fluorescence and electron microscopy and immunohistopathology

An identifying characteristic of the NCLs is the abnormal accumulation of autofluorescent storage material in the brain and retina with ultrastructural features that distinguish this material from normal age pigment (lipofuscin).[64,65] To determine if the progressive neurological signs exhibited by the affected dogs were a result of NCL, samples of the cerebral cortex (parietal lobe), cerebellum and retina from the proband and her affected littermate were examined with fluorescence and electron microscopy. Since in the NCLs the accumulation of autofluorescent storage material is not restricted to nervous tissue, samples of heart ventricular wall were also examined for the presence of autofluorescent storage material. For fluorescence microscopy, samples were either fixed in 10% buffered formalin or immune-fixative (3.5% formaldehyde,

0.05% glutaraldehyde, 120mM sodium cacodylate, 1mM $CaCl_2$, pH 7.4). Samples were processed and embedded in Tissue-Tek OCT Compound (Sakura Finetek, Torrance, CA) and frozen on dry ice. Sections of the tissues were cut with a cryostat at a thickness of 8 □m and mounted on Superfrost Plus slides (Fisher Scientific, Fairlawn, NJ) in 170mM sodium cacodylate. The sections were examined and photographed using fluorescence microscopy as previously described.[86] For electron microscopic examination, the tissue samples were fixed in 2.5% glutaraldehyde in 100mM sodium cacodylate, pH 7.4. The samples were post-fixed with osmium tetroxide and embedded in epoxy resin. Sections of the embedded tissue were cut on an ultramicrotome at thicknesses of 0.4 to 0.8 □m, mounted on glass slides and stained with toluidine blue. Areas of interest were identified by microscopic examination of these sections, and the blocks were trimmed to remove tissue from outside the areas of interest. Sections were then cut from the trimmed blocks at thicknesses of 70 to 90 nm. These sections were mounted on 200 mesh copper thin-barred grids, stained with uranyl acetate and lead citrate, and examined and photographed using a JEOL 1400 transmission electron microscope equipped with a Gatan digital camera.

As with other neurodegenerative diseases,[89] NCLs are characterized by astrogliosis as demonstrated by increased amounts of glial fibrillary acid protein (GFAP) in astrocytes. To determine whether the disease in the affected dogs was accompanied by an elevation in astrocytes containing high levels of GFAP, formalin-fixed samples of cerebellum from the affected dogs and from an unrelated normal dog were processed and embedded in paraffin. Sections of these tissues were immunostained for GFAP and examined with light microscopy as previously described.[90]

Molecular genetic analyses

The Illumina TruSeq DNA PCR-Free Sample Preparation Kit was used with DNA from a single NCL-affected Golden Retriever (the proband) to prepare two libraries with approximate fragment lengths of 350 bp and 550 bp for paired-end sequencing. The two libraries were sequenced in separate flow-cell lanes on an Illumina HiSeq 2500 sequencer at the University of Missouri DNA Core Facility. The adapter sequences were trimmed from the sequence reads with custom Perl scripts and the adapter-trimmed reads were error corrected using MaSuRCA v1.9.5 software.[42] NextGENe software (SoftGenetics) was used to align the error corrected reads to the CanFam3.1 reference genome assembly.[91] The NextGENe Viewer application was used to manually scan the alignment across all coding regions of the 13 candidate genes previously associated with human NCL (Table 1) and to identify sequence variants (differences between the affected Golden Retriever genome sequence and the CanFam3.1 reference sequence).

To determine if sequence variants in the coding regions of the 13 candidate genes also occurred in unaffected dogs, we queried an in-house PostgreSQL database which contained the variant calls for the whole genome sequences of an additional 127 canid samples. These control whole genome sequences were from normal canids or from dogs with other diseases. Fifty eight of the whole genome sequences were generated by our group and 70 were provided by others as listed in the Acknowledgements.

An apparent *CLN5*c.934_935delAG 2 bp deletion (numbered relative to GenBank accession no. NC_006604.3) was verified by the automated Sanger sequencing of amplicons produced with PCR primers *5'-TTTGCTTTGGTGTTCACATAGG-3'* and *5'-CCCAAGTAGGTAGGTTCTCCA-3'*. A TaqMan allelic discrimination assay[44] was used to genotype DNA samples from individual dogs for the *CLN5c.934_935delAG* sequence variant.

For this assay, the PCR primer sequences were *5'-TCTGCTCAGTATCTTGCAAATTTTTGATG-3'* and

5'-AATAAAAGGAAATTTCATAGGTAAAAACCAATATTCAAAA-3' and the sequences of the competing probes were *5'-VIC-TGATTATACACAGAGAGTTT-MGB-3'* (reference allele) and *5'-FAM-AGTGATTATACACAGAGTTT-MGB-3'* (mutant allele).

Results

Fluorescence microscopy

As with the other canine NCLs, the clinically affected Golden Retriever littermates exhibited substantial accumulations of autofluorescent storage material in the cerebellum, cerebral cortex and retina (Fig. 2). In the cerebellum, the storage material was present primarily in the perinuclear areas of the Purkinje cells (Fig. 2A). There were only very minor amounts of autofluorescent material in the granular and molecular layers. In the cerebral cortex the accumulation of autofluorescent material was relatively uniform among neurons throughout the tissue (Fig. 2B), but this material was not observed in white matter tracts. In the retina, the autofluorescent storage material was most prominent in the ganglion cells and along the outer limiting membrane (Fig. 2C), but small amounts of this material could also be seen along the inner edge of the inner nuclear layer. In the heart, large amounts of the storage material were present in muscle fibers occurring in elongated clumps with long axes in parallel to the long axes of the muscle fibers (Fig. 2D).

Electron microscopy

Electron microscopic examination of the storage material from the cerebellar Purkinje cells (Fig. 3A), cerebral cortical neurons (Fig. 3B), and retinal ganglion cells (Fig. 3C) indicated that the

material within the membrane-bound inclusion bodies consisted primarily of membrane-like structures in a variety of arrangements. In the cerebellar Purkinje cells the profiles of the membrane-like structures ranged from elongated linear profiles to small circular profiles (Fig. 3A). These two types of profiles were interspersed. In the cerebral cortical neurons and retinal ganglion cells, the storage body contents consisted almost exclusively of linear membrane-like profiles; almost none of the circular profiles were observed (Fig. 3B and 3C). The cortical neuron storage bodies appeared to be composed of clumps of the linear structures (Fig. 3B).

Immunohistochemistry

The cerebellum samples from the affected dogs contained abundant cells with astrocytic morphology and strong GFAP immunostaining throughout all layers of the tissue. These GFAP-positive cells were particularly abundant adjacent to the Purkinje cell layer (Fig. 4A). By contrast, no strongly GFAP-positive cells were observed in the cerebellum from the normal control dog (Fig. 4B).

Molecular genetics

DNA from the proband was used to generate a whole genome sequence that had 21-fold average coverage and was found to contain 6.1 million potential sequence variants. The reads were deposited in the Sequence Read Archives (Accession SRR1784082) and aligned to the canine genome reference assembly. The alignment was manually scanned across each of the 130 coding exons in the canine orthologs of the 13 genes associated with human NCL. The results are summarized in Table 2. No variants were found in the coding exons or adjacent splice site signals in *PPT1*, *CLN3*, *DNAJC*, *MFSD8*, *CTSD*, or *KCTD7*. In and around the coding exons of

the other 7 candidate genes, we found 9 sequence variants including 3 homozygous and 2 heterozygous synonymous mutations that are unlikely to cause disease. In addition, we found individual missense mutations in *ATP13A2* and *CTSF*, a complex deletion-insertion in *GRN*, and a two base pair deletion and frame shift in *CLN5*. The missense mutations and the complex deletion-insertion were all frequent in the 127 control whole genome sequences; whereas, the *CLN5* frame shift, *CLN5:c.934_935delAG*, was homozygous in the proband but absent from all 127 control whole genome sequences. The homozygous *CLN5c.934_935delAG* deletion (Fig 5A) was verified by automated Sanger sequencing (Fig. 5B).

We used a TaqMan allelic discrimination assay to genotype individual dogs for the *CLN5:c.934_935delAG* deletion. Seventeen of the genotyped dogs were closely related to the proband, as shown in the pedigree (Fig. 1). Of the genotyped family members, only the three affected dogs (the proband, her clinically affected littermate sibling and her clinically affected sibling in the second litter) were *CLN5:c.934_935delAG* homozygotes. The other family members with available DNA were either heterozygotes (n = 11) or homozygous for the wild type allele (n = 4) and none of these dogs exhibited any neurological abnormalities. An additional 1062 randomly selected DNA samples from Golden Retrievers were genotyped for the *c.934_935delAG* deletion. One of these samples was homozygous for the deletion allele, 8 were heterozygotes and the rest were homozygous for the wild type allele. When contacted, the owner of the deletion-allele homozygote indicated that her dog had been euthanized due a progressive neurodegenerative disease with clinical signs and an age at onset that were similar to those of the Golden Retriever siblings in our study. Nonetheless, a clinical diagnosis of NCL could not be confirmed because necropsy tissues were not collected. This homozygote and one of the heterozygotes were related to the proband through a common ancestor: the paternal grandfather

of the proband's paternal grandfather. We could not identify family relationships between any of the other heterozygotes and the proband's relatives. The earliest birthdate of the heterozygotes was 1994. Thus, although the *c.934_935delAG* allele is rare, it may be rare but found in unrelated lines of Golden Retrievers. The *CLN5:c.934_935delAG* allele was not detected in the DNA samples from any of the 143 purebred dogs from 99 breeds other than Golden Retriever.

Discussion

Disease phenotype

We investigated a 3-generation Golden Retriever family in which 4 siblings developed a progressive neurological disease with signs that included anxiety, constant circling, tremors, aggression, ataxia, localized and generalized seizures, and visual impairment. The proband and her affected female littermate exhibited abnormal accumulations of autofluorescent cytoplasmic material primarily in the Purkinje cells of the cerebellum, in neurons throughout the cerebral cortex and in the ganglion cells of the retina (Fig. 3). Examination by electron microscopy revealed that the cytoplasmic storage material consisted of membrane-bounded inclusions containing primarily rectilinear and circular profiles. The progressive neurodegeneration and the autofluorescence and ultrastructure of the cytoplasmic storage material in the brain and retina are distinguishing features of the NCLs.

The neurological abnormalities in the affected Golden Retrievers became apparent to their owners when the dogs reached 15 months of age or older. In other breeds, dogs with genetically defined NCL have developed many of the same neurological signs as these Golden Retrievers; however, the ages at onset of the disease signs have varied among the various canine NCL models. Compared to the affected Golden Retrievers, Dachshunds with mutations in *PPT1* or

TPP1, a Chinese Crested Dog with a mutation in *MFSD8*, and English Setters and a mixed breed dog with different *CLN8* mutations all had diseases with earlier onsets of neurodegenerative signs.[22,25,26,73,74] The onset of disease signs in Tibetan Terriers with NCL due to a mutation in *ATP13A2* has occurred at much later age (4 to 6 years) than that of the Golden Retrievers.[92] The ages at onset of clinical signs for Border Collies with a *CLN5* mutation, an Australian Shepherd with a *CLN6* mutation and American Bulldogs with a *CTSD* mutation were similar to those of the Golden Retrievers described here.[76,82,85]

Disease genotype

Alignment of the adapter-trimmed, error-corrected reads of the proband's whole genome sequence to canine reference genome sequence CanFam3.1, allowed us to scan for sequence variants in all coding regions in the canine orthologs of the 13 genes known to harbor NCL-causing mutations in human patients. This was the third whole genome sequence we generated with DNA from a dog with a confirmed case of NCL. Earlier, we used this approach to identify a *CLN8* nonsense mutation as the probable cause of NCL in a mixed-breed dog and to identify a single-base-pair deletion and frameshift in *MFSD8* as the probable cause of NCL in a Chinese Crested Dog.[25,26] In the proband's whole genome sequence, a scan of the coding regions of the 13 canine othologs of the genes known to harbor human NCL mutations identified only one plausible causal sequence variant: a 2 bp deletion and frame shift in *CLN5*. Although the *CLN5 c.934_935delAG* allele was rare among Golden Retrievers, homozygosity for this mutation was completely concordant with clinical signs of NCL in the genotyped members of the proband's family.

CLN5 was first identified as a gene that harbored distinct homozygous mutations in 3 human NCL patients in 1998.[93] By now at least 35 different *CLN5* mutations have been identified as the

39

likely causes of disease in human NCL patients with (http://www.ucl.ac.uk/ncl/CLN5mutationtable.htm). In addition, nullizygous transgenic knockout mice developed NCL[94] and naturally occurring *CLN5* mutations have been associated with NCL in Devon cattle, Borderdale sheep, and Border Collie dogs.[75,95,96] The Border Collie mutation was a *CLN5:c.619C>T* transition that predicted a CLN5:p.Q207X nonsense mutation. The whole genome sequence from the Golden Retriever proband indicated that the dog was homozygous for a novel *CLN5* mutation, *CLN5:c.934_935delAG*, which predicted a frameshift and premature termination codon, CLN5:p.E312Vfs*6.

Mutant CLN5 protein

CLN5 encodes a polypeptide, CLN5, with no recognized functional domains and no detected amino acid sequence homology to other proteins; however, human CLN5 contains 8 N-glycosylation consensus sequences (Asn-X-Thr/Ser),[93] all of which are conserved in canine CLN5. In addition, human CLN5 contains 2 clusters of hydrophobic amino acid residues, one near the N-terminal end and the other near the C-terminal end,[93] that are also conserved in canine CLN5. GenBank accession NP_006484, indicates that the full-length human CLN5 polypeptide consists of 407 amino acids. However, there are also 3 shorter isoforms which start with in-fame initiator methionines at codon positions 30, 50, and 62.[97,98] Signal peptides of various lengths are cleaved from newly synthesized CLN5 while still in the endoplasmic reticulum.[98] The online software program, SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/),[99] predicts that the cleavage site will be between p.S91 and p.R92 or between p.G95 and p.I96. The only initiator methionine available for translating canine CLN5 corresponds to the initiator at position 50. SignalP 4.1 predicts that the signal peptide will be removed from canine CLN5 by cleavage

between p.G38 and p.G39, which corresponds to the predicted cleavage site between codons 95 and 96 in the human *CLN5*. If these predicted cleavage sites are correct, the signal peptide from the human isoform starting at codon 50 would be 46 amino acids in length and would have 17% sequence homology with the 38 amino-acid-long canine signal peptide. Furthermore, both the canine and human propeptides would be 312 amino acids in length and the sequence homology between them would be 91% (Fig 6).

Mature normal CLN5 protein is primarily located in the lysosomes.[98,100,101] As the human CLN5 propeptide passes through the golgi apparatus on its way to the lysosomes, it acquires mannose-rich oligosaccharides at all 8 of its N-glycosylation consensus sequences.[102] Site-directed mutagenesis experiments with the human gene suggested that the glycosylation of at least 6 of these sites is required for normal protein function.[102] Glycosylation at the furthest C-terminal of these sites was required for normal trafficking to the lysosomes.[102] This site is lost in the truncated protein predicted from the mutant Golden Retriever *CLN5* (Fig 6). The C-terminal hydrophobic region of CLN5 was reported to form an amphipathic helix that tightly associates with biomembranes, but does not cross them.[103] This membrane anchoring may be required for normal CLN5 trafficking to the lysosomes.[103] The amphipathic helix is conserved in canine CLN5 where it occupies amino acids 296 through 335. This region is interrupted by the frameshift at p.E312 as predict by the *CLN5:c.934_935delAG* Golden Retriever mutation (Fig 6). Thus, the truncated CLN5 predicted by the Golden Retriever *CLN5* mutation lacks the 39 C-terminal amino acids, the last N-glycosylation site and part of the amphipathic helix. In contrast, the truncated CLN5 predicted by the Border Collie *CLN5* mutation lacks 143 C-terminal amino acid residues, four of the eight N-glycosylation sites and the entire amphipathic helix (Fig 6); yet, both canine CLN5 models have indistinguishable clinical presentations. This suggests that

little or no biological function is retained as a result of either mutation. That both mutations result in a complete loss of function is further supported by the fact that human patients that are homozygous for the major Finnish CLN5 variant, CLN5:p.Y392X, had classical late-infantile disease signs even though only 16 amino acid residues were missing from the truncated gene product (Fig.6).[93]

Animal models for CLN5

The progression of clinical signs we and others have reported for dogs with homozygous truncating mutations in *CLN5,*[75,82] also occurred in human CLN5 patients. As with the dogs, these patients exhibited an initial mild loss of coordination. This was followed by increasingly severe mental and motor deterioration, visual impairment, and seizures and culminated with premature death.[104–108] Atypical human CLN5 disease variants with delayed onset and a slower progression of clinical signs have been described.[105–108] Nonetheless, the onsets of clinical signs for patients with typical late-infantile CLN5 disease have occurred at 4 to 7 years of age.[104,105] Most of these patients were blind before their 11th birthday and they seldom survived past the age of 30.[104,105] In both the *CLN5:c.619T* homozygous Border Collies and the *CLN5:c.934_935delAG* homozygous Golden Retrievers described here, the clinical signs became apparent at about 13 months of age when dogs were young but sexually mature adults.[75,82] Compared to their human counterparts, it not unusual for canine neurodegenerative disease models to have disease onsets at earlier absolute ages but at later developmental stages. The dogs with truncating *CLN5* mutations were usually euthanized before their third birthday due to their deteriorating health and aggressive behavior. In these potential canine CLN5 models, the 1.5 to 2.0 year time span from the onset of clinical signs to euthanasia may prove to be ideal for evaluating and optimizing therapeutic interventions.

Studies with the *CLN5*-knockout mice have provided insights into the nature and spatiotemporal distribution of histopathologic and biochemical lesions during disease development.[94,109,110] Nonetheless, large animal models are often preferable to rodents for investigating therapeutic interventions for neurodegenerative diseases because their gyrencephalic brains are more similar to human brains in size and complexity and because the durations of clinical signs for the large animal model diseases are closer to those of the human diseases. Thus, a colony of *CLN5*-deficient sheep is maintained as a large animal model for CLN5 disease that is currently being utilized to evaluate the potential efficacy of *CLN5* gene therapy.[96,111] Compared to the sheep model, a *CLN5*-deficient dog model would present fewer problems for housing and management, particularly in urban research settings. In addition, the dogs have more complex cognitive behaviors than do sheep. This is important in a model for a disease in which one of the major symptoms is progressive decline in cognitive abilities. Border Collies, the breed with the previously reported *CLN5* mutation,[75] are noted for their high-energy, workaholic attitude. The easy going temperaments that typify Golden Retrievers would make them more adaptable to research colony life. We have obtained and stored frozen semen from the heterozygous father of the proband, thereby preserving an opportunity for us or others to establish a Golden Retriever based model for CLN5 disease.

Acknowledgements

Fig. 1: Pedigree of the Golden Retriever family showing relationships among clinically normal dogs (open figures) and affected dogs (solid figures). Males are represented by squares and females by circles. The arrow points to the figure representing the proband. The genotypes at *CLN5c.934_935delAG* are shown under the figures representing dogs with available DNA. Del = deletion allele; Wt = wild type allele.

Fig. 2. Fluorescence micrographs of cryostat sections of cerebellum (A), cerebral cortex (B), retina (C), and heart muscle (D) from a Golden Retriever dog that was euthanized after exhibiting clinical signs of NCL. In the cerebellum, the perinuclear regions of the Purkinje cells (pc) contained large amounts of autofluorescent material. In the cerebral cortex the majority of neurons contained this material (arrows in B) in varying amounts. In the retina the autofluorescent storage material occurred predominantly in the ganglion cells (gc) and along the outer limiting membrane (olm). Autofluorescent storage material was also present in the heart muscle fibers (arrows in D). Bar in (A) indicates the magnification for both (A) and (B).

46
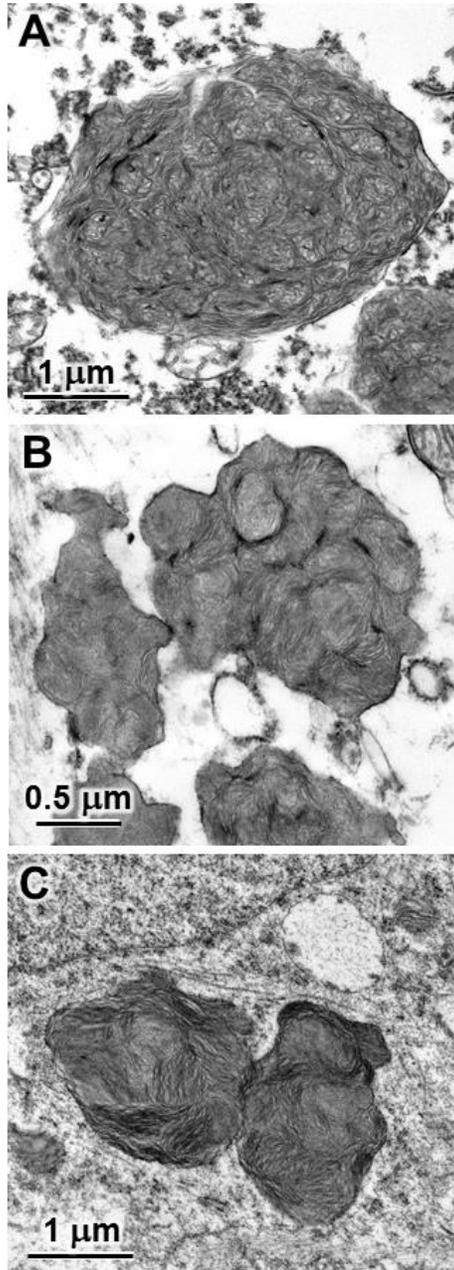
Fig. 3. Electron micrographs of an affected dog's storage bodies from a cerebellar Purkinje cell (A), cerebral cortical neuron (B) and retinal ganglion cell (C).
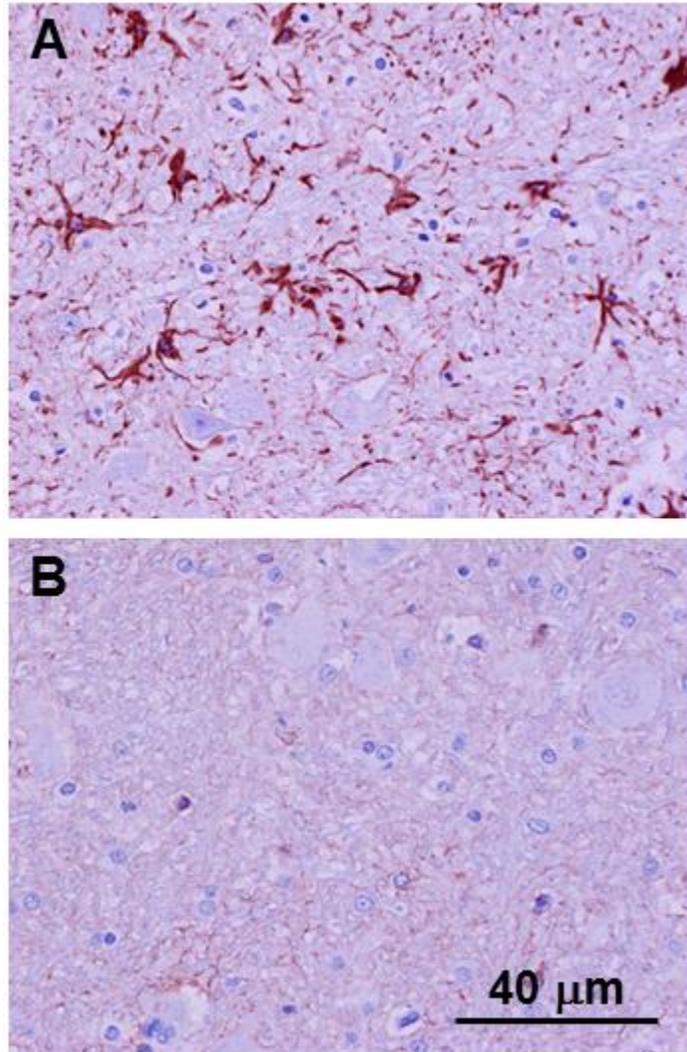
Fig. 4. GFAP-immunostained sections of the cerebellar medulla from an affected Golden Retriever (A) and from a normal dog that did not exhibit any neurological abnormalities at the time of euthanasia (B). The density of GFAP-positive cells and the amount of GFAP staining was much higher in the affected dog. Bar in (B) indicates magnification of both micrographs.
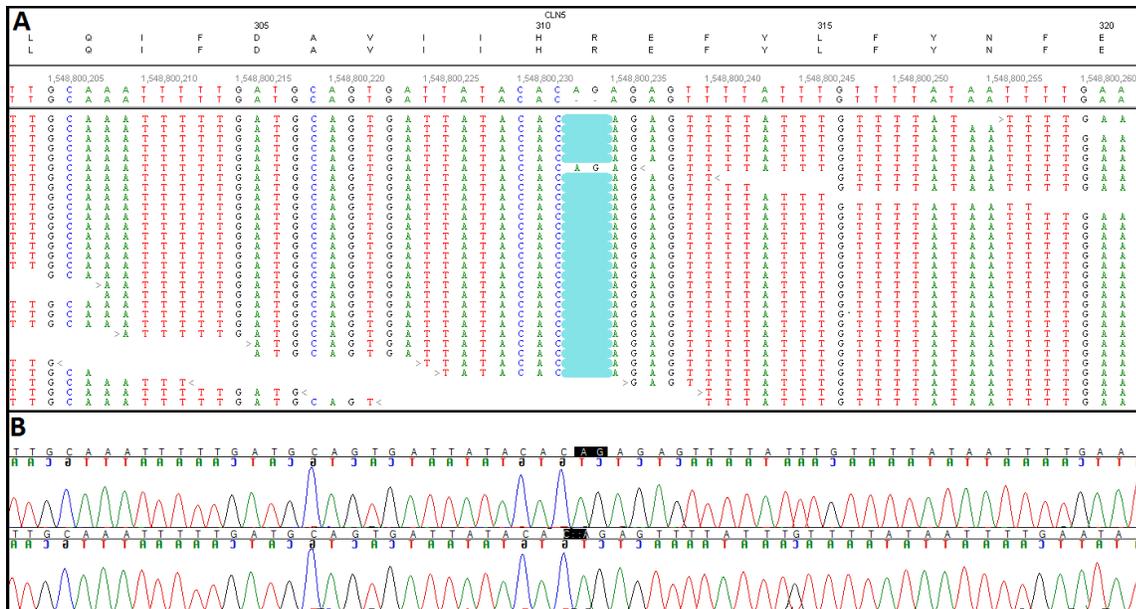
Fig. 5. Canine genomic DNA sequences around *CLN5:c.934_935delAG*. (A) Screen capture of the NextGENe Viewer displaying the regional alignment of whole genome sequence reads from an affected Golden Retriever showing the homozygous *CLN5:c.934_935AG* deletion. (B) Automated Sanger sequences from a normal control dog (top) and an affected Golden Retriever (bottom) verify the homozygous *CLN5:c.934_935AG* deletion.
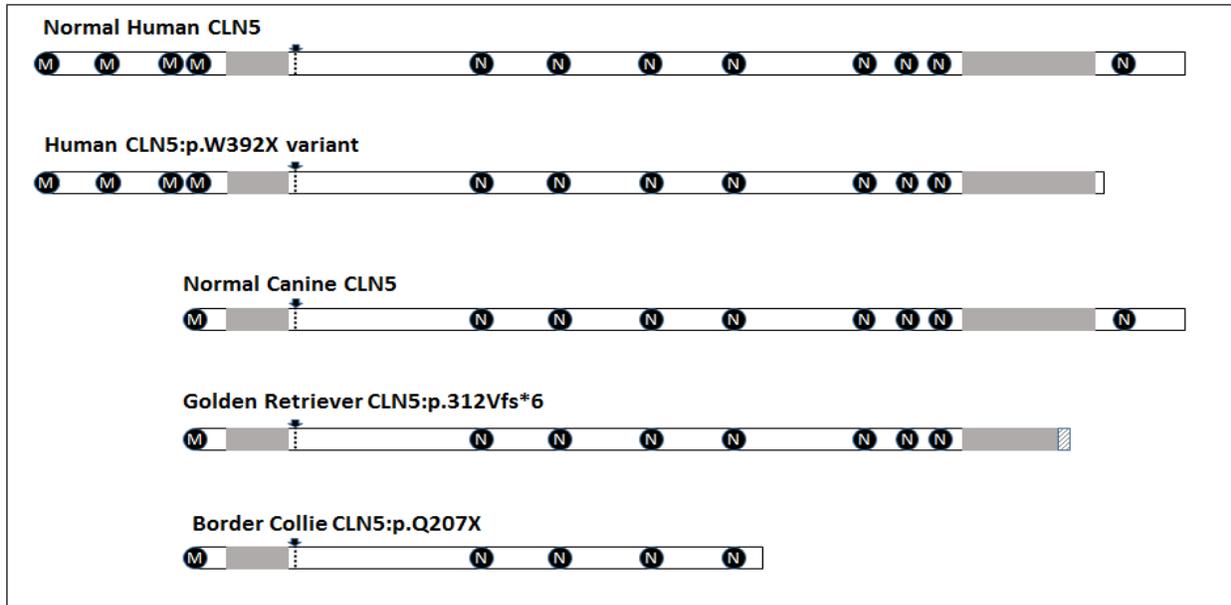
Fig. 6: Diagrams showing the relative positions of features in normal and mutant human and canine CLN5. Normal human CLN5 has 4 potential inframe initiator methionines (encircled Ms), 8 N-glycosylation consensus sequences (encircled Ns) and 2 hydrophobic regions (shaded regions). The arrow shows the predicted signal peptide cleavage site. Normal canine CLN5 has only 1 potential inframe initiator methionine; however the other features are similar to the human protein. The Golden Retriever CLN5:p.312fs*6 variant protein lacks part of the C-terminal hydrophobic region and one of the N-glycosylation consensus sequences and has acquired 6 variant amino acid residues between the frameshift and the first in-frame termination codon (diagonal shading). The Border Collie CLN5:p.Q207X variant protein lacks the entire C-terminal hydrophobic region and 4 of the N-glycosylation consensus sequences. The human CLN5:p.Y392X variant protein only lacks 1 of the N-glycosylation consensus sequences.

50

Table 1: Human NCLs with Canine Models

Human Disease[†]	Causal Human Gene	Observed Nucleotide Sequence Change in Canine Ortholog[‡]	Predicted Amino Acid Sequence Change in Canine Gene Product	Affected Dog Breed
CLN1	*PPT1*	*PPT1:c.736_737insC*	p.F246Lfs*29	Dachshund[73]
CLN2	*TPP1*	*TPP1:c.325delC*	p.A108Pfs*6	Dachshund[74]
CLN5	*CLN5*	*CLN5:c.619C>T*	p.Q207X	Border Collie[75]
CLN6	*CLN6*	*CLN6:c.829T>C*	p.W277R	Australian Shepherd[76]
CLN7	*MFSD8*	*MFSD8c.843delT*	p.F282Lfs*13	Chinese Crested Dog[26]
CLN8	*CLN8*	*CLN8:c.491T>C*	p.L164P	English Setter[22]
		CLN8:c.585G>A	p.W195X	Australian Shepherd Mix[25]
CLN10	*CTSD*	*CTSD:c.597G>A*	p.M199I	American Bulldog[77]
CLN12	*ATP13A2*	*ATP13A2:c.1623delG*	p.P541fs*56[♯]	Tibetan Terrier[23,78]

[†]There are as yet no known canine models for *CLN3*, *CLN4*, *CLN11*, *CLN13* or *CLN14*; *CLN9* is not in current use.

[‡]A canine lysosomal storage disease originally described as a NCL[79] is more appropriately classified as a mucopolysaccharidosis[80,81].

Table 2: Sequence variants from the proband's whole genome sequence that are within or near the coding regions of the 13 NCL candidate genes.

Gene	Exons with no variants	Variants				
		Exon	Zygosity	cDNA change	Amino Acid Change	Comment
PPT1	1-9					
TPP1	1-6, 8-13	7	Homozygous	*c.711A>G*	p.A237A	Silent mutation
		7	Homozygous	*c.885C>T*	p.P295P	Silent mutation
CLN3	1-15					
DNAJC	1-4					
CLN5	1-3	4	Homozygous	*c.934_935delAG*	p.E312Vfs*6	Unique mutation
CLN6	1-3, 5-7	4	Heterozygous	*c.327G>T*	p.L109L	Silent mutation
MFSD8	1-12					
CLN8	2	1	Heterozygous	*c.327A>G*	p.T109T	Silent mutation
CTSD	1-9					
GRN	2-9	1	Homozygous	*c.96G>A*	p.L32L	Silent mutation
		10	Homozygous	*c.1326_1344 delinsCTGC*	p.LPPAPTH 422_448FC	Common allele
ATP13A2	2-28	1	Homozygous	*c.208G>A*	p.A70T	Common allele
CTSF	1-5,7-13	6	Heterozygous	*c.G735T*	p.H245Q	Common allele
KCTD7	1-4					

CHAPTER IV

DILATED CARDIOMYOPATHY

An *RBM20* Frameshift Mutation in a Canine Model of Dilated Cardiomyopathy

Doug Gilliam, Stacey Leach, Barbara Gandolfi, Ana Kolicheski, Liz Hansen, Tendai Mhlanga-Mutangadura, Jeremy F. Taylor, Robert D. Schnabel, Gary S. Johnson

Abstract

Young-adult onset dilated cardiomyopathy (DCM) segregates in the Standard Schnauzer dog breed in a pattern consistent with an autosomal recessive mode of inheritance. To identify the molecular genetic cause of Standard Schnauzer DCM, DNA from an affected dog was used to generate a whole genome sequence with 31-fold average coverage. Among the sequence variants in this whole genome sequence was a 22-bp deletion and frameshift in *RBM20*, the canine ortholog of a gene previously associated with human DCM. The *RBM20* deletion allele was homozygous in the whole genome sequence of the affected Schnauzer, but absent from 101 whole genome sequences of normal canids or dogs with other diseases. An additional 753 Standard Schnauzers, including 21 with DCM, were genotyped for the *RBM20* deletion. In this cohort, all 20 of the deletion-allele homozygotes had DCM, only one of the 183 heterozygous dogs had DCM, and all of the reference-allele homozygotes were DCM free. *RBM20* deficiency is known to alter exon-splicing patterns and produced aberrant titin isoforms in a rat model and in human DCM patients. To determine if the Standard Schnauzers with DCM had similar exon-splicing abnormalities, RNA prepared from their left ventricular walls was compared by RT-PCR to similarly prepared RNA from normal adult dogs. Titin transcripts with extensive exon skipping were detected in the normal RNA but not in the RNA from the dogs with DCM. Conversely, titin transcripts with retained exons were detected in the RNA from dogs with DCM

but not in the normal-dog RNA. Thus, we have identified a canine model for human *RBM20*-associated DCM.

Introduction

Dilated cardiomyopathy (DCM), characterized by left ventricular enlargement and reduced systolic function, is an important source of human morbidity and mortality and the leading cause for heart transplantation.[112] An estimated 30-50% of DCM cases are familial with inheritance patterns most commonly consistent with an autosomal dominant mode of inheritance.[113] Mutations in >50 different genes have been reported in patients with dilated cardiomyopathy.[112,113]

DCM is also an important veterinary disease. This disease condition is a frequent cause of disability and premature death in dogs.[114] Breed-to-breed differences in ages at onset,[115,116] survival times,[117] inheritance patterns[116,118–120] and myocardial histopathology[115,121–123] are indicative of major genetic contributions.[114] Although many efforts to identify molecular genetic risks factors for DCM by genome mapping,[124–127] linkage with regional markers associated with human DCM genes,[128–130] or the resequencing of canine orthologs of these genes[121,131–135] have failed to find plausible candidate mutations, others appear to have been more successful. A 16 bp deletion at a donor splice site in *PKD4* is reported to increase the risk of DCM in one Doberman Pinscher population,[136] but not in another Doberman Pinscher population.[137] In addition, an 8 bp deletion in the 3'-untranslated region of *STRN*, which had previously been associated with arrhythmogenic right ventricular cardiomyopathy in Boxers,[138] is both common and likely causal in Boxers that develop DCM.[139] Finally, an X chromosome microdeletion that encompassed the entire dystrophin gene caused both skeletal myopathy and DCM in German Shorthair Pointers.[118]

Although DCM has probably occurred in most if not all dog breeds, several reports have focused on the DCM-associated epidemiology, clinical characteristics, laboratory findings and histopathology in certain breeds with a high prevalence of the disease including Boxer,[140] Doberman Pinscher,[119,141–150] Great Dane,[151,152] Irish Wolfhound,[120,153–155] Newfoundland,[156,157] and Portuguese Water Dog.[116,158,159] Harmon et al[160] have recently described DCM in 15 young-

adult Standard Schnauzers. Both genders were equally represented (7 males, 8 females). The ages at onset of clinical signs ranged from 0.9 to 2.9 years. The familial relationships between affected dogs was consistent with an autosomal recessive mode of inheritance. Here we report that a whole genome sequence from one of the affected Standard Schnauzers contained a rare an *RBM20* deletion and frameshift that is the apparent cause of DCM in the breed.

Materials and Methods

Acquisition and storage of biological samples:

All studies were conducted with the approval of the University of Missouri, Animal Care and Use Committee and with the informed consent of the dogs' owners. EDTA-anticoagulated blood and buccal swabs were collected from privately owned dogs. Blood samples were shipped over ice to the University of Missouri. The biological material from buccal swabs was stabilized on Whatman FTA-elute cards (Thermo Fisher Scientific, Waltham, MA) and shipped to the University of Missouri at room temperature. DNA was extracted from the blood and buccal swab samples by previously described procedures.[22,161]

Myocardial tissue samples from left ventricular walls were collected at necropsy, shortly after euthanasia, from two Standard Schnauzers with DCM and congestive heart failure and from two adult experimental dogs euthanized for reasons unrelated to heart disease. The sample were immediately stabilized in RNA*later* (Qiagen Inc. Valencia, CA).

Generation and analysis of a whole genome sequence: DNA from a Standard Schnauzer with DCM was submitted to the University of Missouri DNA Core Facility for the preparation of 2 PCR-free libraries (fragment sizes of approximately 350 bp and 550 bp) with the TruSeq DNA PCR-Free Sample Preparation Kit (Illumina, San Diego, CA) and for 2X100 paired end sequencing in 2 flow-cell lanes in an Illumina HiSeq 2000 Sequencer. The adaptors were trimmed with custom Perl scripts and the adaptor-trimmed reads were deposited in the NCBI Sequence Read Achieve (accession SRX863111). The resulting FASTQ files were processed as previously described.[26] After error correction with MaSuRCA v2.2.2 software[42] the reads were aligned to the CanFam3.1 reference genome assembly with NextGENe software (SoftGenetics, State College, PA), which was also used for the identification and initial categorization of the sequence variants. Likely false positive variant calls were identified and removed with custom

Perl scripts. The genome-wide average coverage was 31 fold. The 6.2 million potential sequence variants (differences from the reference sequence) were uploaded to a custom PostgreSQL database that also contained the variant calls from another 101 canid whole genome sequences which served as controls. Forty-three of these control WGSs were from our group and 48 were from others listed in the Acknowledgements.

Variant verification: PCR amplification and direct Sanger sequencing was used to determine the validity of a candidate sequence variant, *RBM20:c.2472_2493delGAAGGTCAAAATCTGCCCAGAA* (numbered as in GenBank accession NC_006610), identified from the whole genome sequence. The amplification was done with GoTaq Flexi DNA polymerase (Promega Corp., Fitchburg, WI) and flanking PCR primers 5′-GAGAGAGCTGGTATCCCAC-3′ and 5′-CATTGGTGAAATCCTTGGCTA-3′ for 40 thermocycles at an annealing temperature of 60°C. The size of the amplicons were evaluated by capillary electrophoresis with an eGene HAD-GT12 electrophoresis apparatus and a QIAxcel cartridge (Qiagen). Samples were purified with QIAquick PCR spin columns (Qiagen) and shipped to Sequetech Corp. (Mountain View, CA) for Sanger sequencing with BigDye-Terminator technology and an ABI 3730xl sequencer.

Genotyping: A PCR restriction fragment length polymorphism (PCR-RFLP) assay was used to genotype individual DNA samples for the 22-bp *RBM20* deletion by PCR amplification with primers

5'-CATTTTGGTTCAGGAGCAAGAGT-3' and *5'-GGCATTTCCACGTCCATGTCA-3'* as described above, except that the annealing temperature was 62°C. The resulting amplicons were digested for 1 hr with restriction enzyme Hpy 188I (New England Biolabs, Inc., Ipswich, MA) under conditions recommended by the supplier. Hpy 188I hydrolyzes the ancestral allele but not the variant allele. The sizes of the restriction fragments were evaluated by capillary electrophoresis.

Evaluation of titin transcripts: A PureLink RNA Mini Kit (Invitrogen, Carlsbad, CA) was used to extract RNA from the left ventricular wall tissue sections obtained from 4 dogs and stored in RNA*later* as described above. The 4 RNA preparations were compared to each other in reverse transcriptase-PCR (RT-PCR) experiments. In each experiment, SuperScript III reverse transcriptase (Invitrogen) was used for first strand synthesis. The same oligonucleotide serve as

primer the reverse transcriptase and as antisense primer in the PCR amplification. In the first experiment, the sense primer was *5'-CACCTCGCCAAATTCACCT-3'* and the antisense primer was *5'-ATTTCAGGCACTTTTACTGGTT-3'*. The primer pairs used for the other two RT-PCR experiments were *5'-AATACAAAATCAGTGTCACCG-3'* (sense) with *5'-AGCCACTATACATTCTAAGTCA-3'* (antisense) and *5'-AAATTCAAAAGTATATCGAGCCT-3'* (sense) with *5'-TTCAGGAAGTAATTTGCGAACT-3'* (antisense). For each experiment, the resulting amplicons were detected by capillary electrophoresis and their identity was established by automated Sanger sequencing as described above.

Results

Illumina massively parallel sequencing technology was used for whole genome sequencing of a male Standard Schnauzer with DCM. Alignment of the processed whole genome sequence reads to the reference canine genome sequence (CanFam 3.1) indicated 31-fold average coverage with 7.2 million potential sequence variants (differences from the reference canine genome CanFam 3.1). Because DCM was believed to segregate in Standard Schnauzers as an autosomal recessive trait,[160] we queried the custom PostgreSQL database that contained these sequence variants along with the variants from 101 other control canid whole genome sequences to identify only those variants with the following three criteria: homozygous in the genome sequence from the affected Standard Schnauzer; absent from the 101 control genome sequences; and, predicted to alter the amino acid sequence of the gene product. Fifty-nine sequence variants in 57 genes met these criteria Among them were 2 nonsense mutations, 3 deletions that caused a reading-frame shift, 5 in-frame deletions, 1 in-frame insertion and 48 missense mutations in 46 different genes (Table 1). Only one of these sequence variants, a 22 bp deletion and frameshift in *RBM20* (*RBM20:c.2472_2493delGAAGGTCAAAATCTGCCCAGAA*), was in a gene previously associated with human dilated cardiomyopathy. This variant was verified by the direct automated Sanger sequencing of an amplicon produced with PCR primers spanning the deletion.

A PCR-RFLP assay was used to genotype the DNA sample used whole genome sequencing and DNA samples from 753 other privately owned Standard Schnauzers. Twenty two of these dogs

had clinical histories of DCM. The results are summarized in Table 2. All but one of the Standard Schnauzers with DCM were homozygous for the 22 bp *RBM20* deletion. The only exception, a female Standard Schnauzer, was a heterozygous for the DCM deletion. This dog received an echocardiographic diagnosis of DCM when 2 and a half years old. Unlike the other Standard Schnauzers with DCM, the heterozygote had very frequent arrhythmias (ventricular premature complexes and runs of ventricular tachycardia) that has not responded to anti-arrhythmic therapy. None of the other 182 heterozygotes and none of the 550 Standard Schnauzers that were homozygous for the reference allele had been diagnosed with DCM. It is noteworthy that at least 7 of the heterozygotes were >11 years old without developing DCM and 3 of them were DCM-free past their 15th birthday.

The cardiac pathology in an *RBM20* deficient rat model and in human DCM patients has been attributed to aberrant exon-splicing patterns in transcripts of genes expressed in the heart such as *TTN*, the gene for titin (see Discussion). To determine if a similar pathogenesis contributed to the development of the DCM in Standard Schnauzers, exon splicing patterns of titin transcripts in RNA preparations from left ventricular wall myocardial tissue from two adult dogs with normal hearts were compared to titin transcript patterns in comparable RNA preparations from two DCM-affected Standard Schnauzers. These comparisons were made with three RT-PCR amplifications that targeted different classes of titin transcripts. One of the RT-PCR amplifications was done with primers designed to anneal to titin exon 48 and to adjacent parts of exons 187 and 188. These are constituently expressed exons that flank an extensive region of alternatively expressed exons (Fig. 1). RT-PCR amplification with these primers produced a prominent amplicon of approximately 330 bp with RNA from normal dogs. Direct automated Sanger sequencing confirmed that this amplicon resulted from the splicing of exon 48 with exon 187. On the other hand, no RT-PCR products were detected when these same primers were used amplify RNA preparations from the ventricles of DCM-affected Standard Schnauzers. This result is consistent with expectation that the titin transcripts in RNA from DCM-affected Schnauzers will have retained many or most of the 138 intervening exons thereby separating the amplification primers by as much as 23 kb. The other two RT-PCR amplifications were attempted with primers designed to anneal to exons 59 and 60 and to exons 182 and 185. These exons are normally removed by exon skipping from adult titin transcripts. As expected, when

the RNA preparations from the two normal dogs were used as template for RT-PCR amplifications with these primer pairs, no products were detected. On the other hand, with the RNA preparations from the two DCM-affected Standard Schnauzers, the primer pairs from exons 59 and 60 produced prominent amplicons of approximately 230 bp and the primer pairs from exons 182 and 185 produced prominent amplicons of approximately 300 bp. Direct Sanger sequencing of these amplicons confirmed that they were generated from titin transcripts with exon 59 spliced to exon 60 and from the consecutively splicing of exons 182 through 185.

Discussion

We identified a homozygous 22 bp deletion in *RBM20* from the whole-genome sequence of a single DCM-affected Standard Schnauzer and showed that homozygosity for this deletion was associated with DCM in this dog breed. *RBM20* encodes RNA-binding motif protein 20 (RBM20), one of the hundreds of RNA-binding proteins that occupy the ribonucleoprotein complexes which regulate a wide variety of cellular processes including the posttranscriptional modifications of messenger RNA.[162] The first report that mutations in human *RBM20* cause dilated cardiomyopathy appeared in 2009 by Brauch et al.[163] These investigators used linkage analysis to map an autosomal dominant form of dilated cardiomyopathy in two large families and identified loci on overlapping segments of chromosome 10. Affected members of each family had distinct heterozygous *RBM20* missense mutations. These two missense mutations were separated by only three codons in exon 9. Examination of coding exons from other dilated cardiomyopathy patients identified three additional heterozygous missense mutation within this same five-codon segment of exon 9.[163] More recently dilated cardiomyopathy patients have been found to harbor at least 10 additional heterozygous *RBM20* mutations both within and beyond the five codon exon-nine hotspot.[164–166]

Much of our current knowledge of RBM20 function comes from investigations of a colony of mutant rats that express an abnormal pattern of cardiac titin isoforms.[167,168] Titin, the largest known protein, is expressed in skeletal and cardiac muscle where it functions as a sarcomeric filament that spans a half-sarcomere length (>1 mµ) with the N-terminal end in the Z-disk and the C-terminal near the M-band (Fig. 2).[169] The segment of titin within the A band of the

sarcomere is fixed in length but the section in the L band has an elastic nature that can change with changes in sarcomere length.[169] Titin's flexability results from extensible elements that act as a series of molecular springs and provide passive tension as the sarcomere extends. Passive tension in the myocardium is important for efficient diastolic ventricular filling. This passive tension results from three biochemically distinct titin I band domains that respond sequentially to increasing tension.[170,171] The first extension under low tension occurs in a tandem series of immunoglobulin-like repeats, the poly-Ig domain. Next, increased tension reversibly extends the PEVK domain, so named because it is rich in proline (P), glutamic acid (E), valine, and lysine (K). Yet higher tension produces reversible extension of cardiac-specific N2-B domain.

Titin is encoded by a single gene, *TTN*. Bang et al[172] identified 363 exons in human *TTN*; however, the 4,200 kDa protein that would result from a transcript containing all 363 exons has not been detected. Instead, preparations from both skeletal and cardiac muscle contain highly heterozygous mixtures of titin isoforms that result from alternative exon splicing. Although imprecise, the exon splicing of titin transcripts is regulated such that titin isoform composition depends upon muscle type (skeletal or cardiac), host organism, developmental stage, location within organ, and physiologic and pathologic states. Classes of titin isoforms have traditionally been distinguished by the presence or absence of epitopes located near the N_2 line of the sarcomere.[173] The N2-B epitope is only expressed in the myocardium; whereas, the N2-A epitope is expressed in skeletal muscle isoforms and in some, but not all cardiac muscle isoforms. Thus, the three recognized classes of titin isoforms are N2-A (from skeletal muscle), N2-B (from cardiac muscle) and N2-BA (also from cardiac muscle).[173,174] With molecular weights of approximately 3,000 kDa, the N2-B titin isoforms are substantially smaller than the >3,300 kDa N2-BA isoforms The size differences are primarily due to greatly shortened poly-Ig and PEVK domains, which explains why myocardium rich in N2-B titin isoforms is relatively stiff, while myocardium rich in N2-BA isoforms are more compliant.[175]

The ventricles of human adults and adult cats, dogs, sheep and pigs contain approximately equal amounts of the N2-BA and N2-B titin isoforms.[176–179] Adult cattle ventricles have higher N2-BA to N2-B ratios and their atria express N2-BA almost exclusively.[177] In contrast, the N2-B class of titin isoforms is predominate in the ventricles of the adult rats, mice, hamsters and rabbits.[174,176,177,180] Their high heart rates and thin ventricle walls may require a stiffer

myocardium. Neonatal rat ventricles contain both N2-BA and N2-B classes of titin isoforms, although only traces of the N2-BA remain in the 2 week old rat.[168,180,181] N2-BA isoforms with unusually high molecular weight distributions (around 3,700 kDa) are the first forms of titin to appear in the developing rat fetus. The highly compliant myocardium produced by these extra-large isoforms facilitate ventricular filling at the low filling pressures of the developing heart.[182] The extra-large isoforms are gradually replaced by smaller N2-BA isoforms so that the N2-BA remnants in the week old rat have a mass of about 3,300 kDa.[168,181] As with rats, extra-large fetal type titin N2-BA isoforms have been detected in neonatal, but not adult mice, rabbits, and pigs.[176,181] Taken together these reports indicate that the distribution of cardiac titin isoforms and the stiffness of the myocardium they populate are under developmental control. The extra-large N2-BA isoforms that first appear in the fetus are gradually replaced by N2-BA isoforms with more extensive exon skipping to produce shorter poly-Ig and PEVK domains. In addition, the N2-B titin isoforms, with more extensive exon skipping, first appear in the neonatal period and during maturation partially or completely replace the N2-BA isoforms to provide a stiffer adult myocardium.

Mutant rats from the previously mentioned colony with abnormal distributions of titin isoforms do not undergo the normal switch from the highly extensible extra-large fetal titin isoforms to the smaller and stiffer adult isoforms.[167,168] These rats have a 95 kb deletion that eliminates 13 of the 14 *Rbm20* coding exons.[166] In heterozygous rats, the switch from fetal to adult titin isoforms is delayed and incomplete. Rats with the homozygous *Rbm20* deletion retain extra-large fetal titin isoforms throughout their lives.[166] Unlike the human patients with *RBM20* mutations, the mutant rats do not develop congestive heart failure. They do, however, develop left ventricular dilation and sub-endothelial fibrosis,[166] which may compensate for their overly compliant myocardium. Abnormal splice variants in transcripts from 30 genes in addition to *Ttn* were identified by the global sequencing of mRNA from the cardiac tissue of rats that carried the *Rbm20* mutation.[166] These results imply that *Rbm20* makes a required contribution to the normal developmental regulation of exon splicing in a subset of genes expressed in the heart. With respect to titin transcripts, exon-splicing patterns are stalled at the embryonic state in rats without functional RBM20 because they are homozygous for the *Rbm20* deletion. The delayed developmental changes in the heterozygous rats likely result from reduced RBM20 concentration due to haploinsufficiency.

The absence of N2-B titin isoforms and the presence of unusually large N2-BA titin isoforms in cardiac biopsy material from a human patient with DCM caused by a *RBM20* mutation is further confirmation that RBM20 plays an important role in the developmental regulation of exon splicing in the heart.[166] If, as expected, a breakdown in the developmental regulation of exon splicing is also responsible for Standard Schnauzer DCM, myocardial RNA from affected dogs should contain extra-long fetal-type titin transcripts due to the inclusion of exons that are skipped in normal adult dogs. Our RT-PCR results support this expectation. RT-PCR amplification with primers designed to anneal to constitutively expressed exons spanning a broad region of differentially expressed exons produced a prominent amplicon with extensive exon skipping from RNA samples from normal dogs but not from the RNA of affected Standard Schnauzers. Furthermore, RT-PCR primers designed to anneal to differentially expressed titin exons produced prominent amplicons with ventricular RNA from the affected Standard Schnauzers but not from the adult normal control dogs. These results imply that the same failure to switch from fetal to adult transcript patterns previously described in the *Rbm20* deficient rats and human DCM patients with *RBM20* mutations is also responsible for the DCM which develops in Schnauzers that are homozygous for the 22-bp *RBM20* deletion.

Similarities between *RBM20*-associated human and canine DCM suggest that the canine disease could become a valuable large animal model for investigating underlying disease mechanisms and/or for determining the efficacy of therapeutic interventions. Yet to be determined is the significance of an apparent species difference in the mode of inheritance. So far all human patients with *RBM20*-associated DCM have carried the mutant *RBM20* allele in the heterozygous state. [163,165,166,183]Fourteen of the 15 known human DCM-associated *RBM20* mutations were missense mutations. The exception was a nonsense mutation, RBM20:p.1031X, predicted to truncate the gene product by 16%. Eight of the 14 *RBM20* missense mutation*s* involved 5 consecutive codons (codon numbers 634 through 638) located within an SR-rich domain likely to participate in protein-protein interactions. The predominance of missense mutations and their clustering suggests the DCM may result from a toxic gain of function rather than haploinsufficiency. In the *Rbm20* rat model, haploinsufficiency is clearly responsible for the abnormalities found in the heterozygotes since their causal microdeletion eliminates nearly the entire *Rbm20* gene. Both homozygous and heterozygous rats develop left ventricular dilation and subendocardial fibrosis; however, neither homozygous nor heterozygous rats develop DCM

or congestive heart failure.[166] Based on pedigree analysis, DCM was reported to be a recessive trait.[160] Most of the genotype results reported here support this conclusion. Recessive Standard Schnauzer DCM appears to be highly penetrant as all 21 of the Standard Schnauzers that were homozygous for the deletion allele developed DCM. Also consistent with a recessive mode of inheritance, the canine *RBM20* deletion and frameshift reported here is predicted to delete the C-terminal third of the normal gene product and is likely to result in a mutant gene product with little or no residual activity. Nonetheless, one heterozygous Standard Schnauzer developed DCM at 2.5 years of age. This could be a phenocopy that developed DCM because of inherited or acquired factors unrelated to the dog's single *RBM20* deletion allele. Alternatively, the single deletion allele may have caused or contributed to the development of DCM. If heterozygous Standard Schnauzers are at higher than normal risk of developing DCM, that risk is much lower than carried by deletion-allele homozygotes since 182 of 183 heterozygotes have not developed DCM and several of them have remained DCM-free into old age.

A detailed clinical description of genetically defined Standard Schnauzer DCM and associated *post mortem* findings is in preparation and will be submitted for publication in the veterinary literature. In the future, it will be interesting to learn if DCM is more common in heterozygous Standard Schnauzers than in reference-allele homozygotes. When tissue samples from the hearts of heterozygous Standard Schnauzers become available at necropsy, histopathologic analysis and analysis of the RNA extracted from theses tissue samples should reveal if the heterozygous state induces subclinical pathology as it does in the rat model. Future RNA-seq experiments with the canine *RBM20* model should be particular interesting as they should reveal which genes in addition to *TTN* produce transcript with RBM20-controled exon splicing pattern.
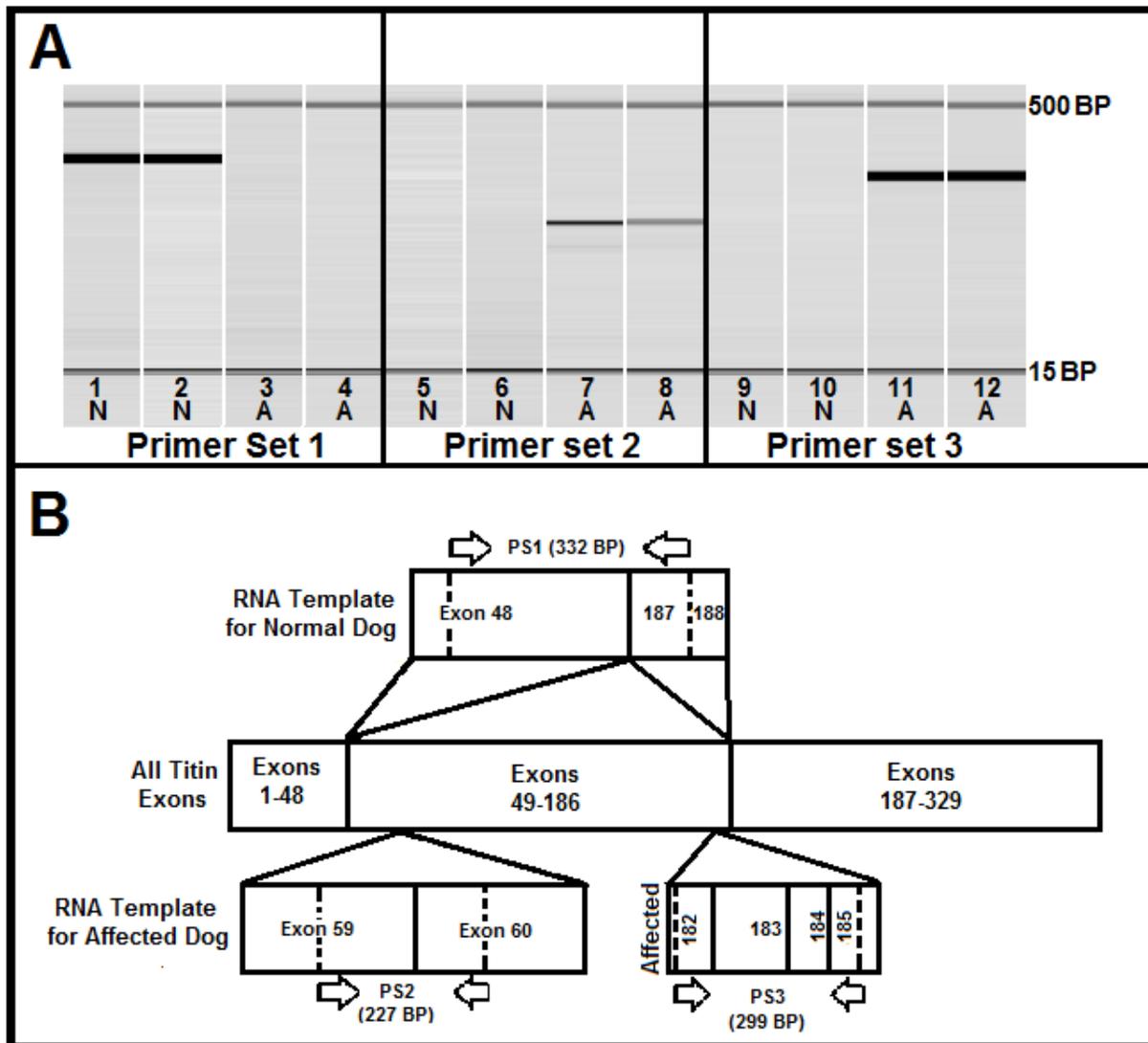
Figure 1. RT-PCR of Titin and in two DCM affected dogs and phenotypically normal, DCM unaffected dogs (A) Microcapillary electrophoretograms of Titin-specific amplicons from total RNA. RT-PCR was performed with cardiac tissue of two DCM-unaffected dogs (lanes 1 and 2, 5 and 6, 9 and 10) and cardiac tissue of two DCM-affected Standard Schnauzer dogs (lanes 3 and 4, 7 and 8, 11 and 12). RT-PCR was performed with three primer sets spanning exons 48-188, spanning exons 59-60) and spanning exons 182-185. (B) Diagram of the position of primers used for RT-PCR amplification relative to an artificial Titin transcript in which all exons are represented.
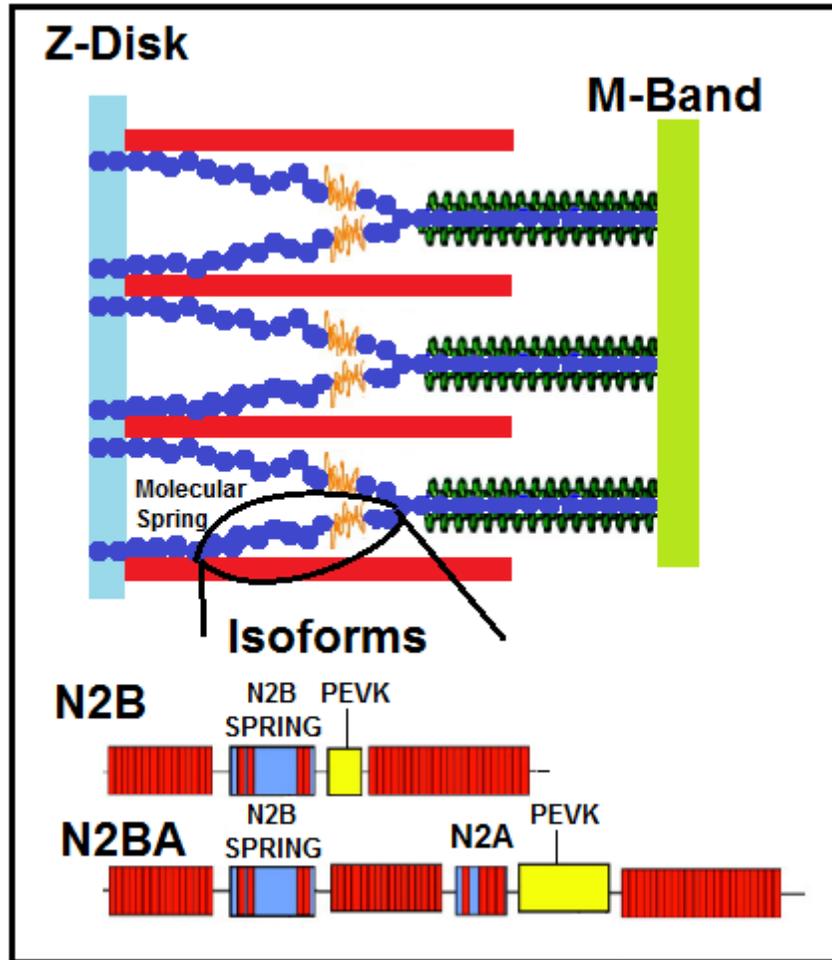
Figure 1. Carciac Sarcomere with relation to Titin. Titin extends from the Z-Disk of the sarcomere where it binds to the Myosin thick filament to the M-Band. Two important cardiac isoforms are shown with middle-IG and molecular spring sections diagrammed.

Table 1: Unique coding sequence variants

Gene	Chromo-some	Chromosome Position	Reference Base	Variant	A.A. Change
ABCC5	34	16860883	C	T	1393G>S
ADAMTSL3	3	55357925	T	G	360S>A
AKR1C3	2	30680225	C	T	92R>C
ANKRD31	3	31139161	G	A	1182D>N
ARHGEF10	16	54367853	G	A	1063A>T
ARHGEF40	15	18163988	C	G	533S>R
BCL2A1	3	57453987	G	T	105V>L
BPTF	9	12643091	C	T	1587V>M
C6H7orf26	6	11924238	G	A	418V>M
CEP72	34	11649338	C	T	729A>T
CHD1L	17	57803800	A	delAGC	In-Frame
CPEB2	3	65266319	G	delGCGGCCGGG	In-Frame
CPEB2	3	65266329	G	delGGG	In-Frame
CYP2U1	32	28425491	C	delCC	FS
DTNBP1	35	14862576	G	delGCGGGGG	FS
EVL	8	68445848	A	C	201T>P
FAM179B	8	22447703	G	A	881G>R
FOXE3	15	13293642	T	delTCGGGGCCGGGG	In-Frame
FOXL2	23	35146662	T	G	36S>R
GJD4	2	1361530	C	G	169G>R
HES5	5	57389229	A	G	19L>P
HHIPL1	8	68059568	C	T	543P>S
ICOSLG	31	38094895	G	C	40R>G

ISM2	8	50285808	T	G	21M>L
ISM2	8	50285811	T	G	20M>L
ITIH2	2	28368055	A	T	930H>Q
KIAA0284	8	72391999	G	C	1122Q>H
KLHL33	15	17764075	A	G	51S>P
LOC100682843	5	63555605	C	G	58A>P
LOC100684157	35	10330650	C	T	232R>W
LOC100686872	7	14556241	G	C	355A>P
LOC100687523	14	5968248	C	T	48E>K
LOC100688803	27	33836969	G	A	119P>S
LOC100688834	6	55275930	G	C	308A>P
LOC100856052	21	27080997	C	T	148V>M
LOC100856118	5	56482848	C	T	16S>F
LOC100856208	9	11524410	C	A	105D>E
LOC476669	20	44823213	C	delCGG	In-Frame
LOC483880	19	29773106	A	C	336L>R
LOC485836	24	22543389	C	T	278R>X
LOC485836	24	22543492	G	T	312C>F
LOC606782	34	228202	C	G	42C>S
LOC607629	6	5597554	T	C	285L>P
LOR	17	61994430	T	G	38C>G
MPHOSPH9	26	6326335	A	C	78Q>P
NOL7	35	13059130	C	G	65A>G
PGAP3	9	22805894	T	G	241S>A
PLEC	13	37458844	C	T	1934R>Q
RANBP3	20	54122045	G	C	36E>Q

RBM20	28	22146844	G	delGAAGGTCAAAATCTGCCCAGAA	FS
RORC	17	60827348	C	A	241A>S
SDSL	26	10759557	C	A	320Q>K
SEPHS2	6	17669238	G	A	444E>K
SH3TC1	3	60019372	C	T	169S>L
SIX3	10	47340777	G	insCGG	In-Frame
SLC6A15	15	26487978	C	T	699G>R
SLC6A16	1	107162770	C	T	17Q>X
VWA5B2	34	17124632	A	T	73I>F
VWA5B2	34	17124637	T	G	74H>Q

Table 2: Distributions of DCM-related phenotypes and RBM20-related genotypes in a Standard Schnauzer cohort.

| | Genotype | | | |
Phenotype	Reference/Reference	Reference/Deletion	Deletion/Deletion	Total
DCM-Affected	0	1	21	22
DCM-Free	550	182	0	732
Total	550	183	21	754

CHAPTER V

Conclusions and Future Studies

We successfully identified the likely causal mutations associated with three different inherited diseases in dogs, all of which are in genes associated with equivalent human pathologies. I will briefly discuss the consequences of the findings of these mutations and suggest future direction of studies for each in turn. Furthermore, I will suggest further paths that can be undertaken to find more causal mutations, which can be plausibly used as models for human disease.

The first finding was the discovery of the missense mutation in *KCNJ10* that causes Spinocerebellar Ataxia in the Jack-Russell Terrier, Parson-Russell Terrier, Russell Terrier as well as several other breeds with ancestry from the British hunting terriers.[30,184] KCNJ10 encodes for the inwardly rectifying potassium channel Kir4.1. Mutations in KCNJ10 in humans have been identified in patients with two autosomal recessive hereditary disease syndromes with overlapping signs: SeSAME (seizures, sensorineural deafness, ataxia, mental retardation, and electrolyte imbalances) and EAST (epilepsy, ataxia, sensorineural deafness, and tubulopathy).[51,52] While these are both rare diseases, further studies to elucidate how mutations in *KCNJ10* affect the function of Kir4.1 and the lack of potassium buffering could potentially shed light on excessive neuronal membrane excitability in other forms of ataxia, myokymia or epilepsy. Another study that could be undertaken would be to determine whether heterozygous carriers have any fitness advantage due to altered levels of *KCNJ10* and the effects that the myokymia has on muscle mass. Finally, while we've discovered the cause of SaMS in Russell-Group Terriers, further genetic studies could be undertaken on the study dog (or another RGT that is

phenotypically and genotypically affected) to determine the genetic etiology of myokymia, which was not present in all study dogs.

The second discovery was a 2bp deletion in canine *CLN5* that is responsible for NCL in Golden Retrievers, a *CLN5:c.934_935delAG* predicting a frameshift and premature stop codon, CLN5:p.E312Vfs*6.[24] *CLN5* was first identified as the causal gene for human NCL in 1998 and Border Collie dogs with NCL have a *CLN5:c.619C>T* transition predicting a CLN5:p.Q207X nonsense mutation .[75,93] The progression of clinical signs that both the Border Collies with the *CLN5:c.619C>T* and the Golden Retrievers with the aforementioned *CLN5:c.934_935delAG* also occurs in human CLN5 patients. One dissimilarity is age at onset: the age at onset for patients with the typical late-infantile CLN5 disease occur from ages 4-7,[104,105] whereas clinical signs in both the Border Collies and Golden Retrievers with NCL become apparent at 13 months of age, when dogs are young but sexually mature adults.[24,75,82]

One potential future direction for this mutation would be the establishment of a large-animal model for CLN5. Currently, a colony of *CLN5*-deficient sheep are maintained as a large-animal model to evaluate the potential efficacy of *CLN5* gene therapy[96,111], a canine *CLN5*-deficient dog model would be easier to house and manage in an urban setting. Another important aspect of a canine *CLN5* model is that dogs have a more complex set of cognitive behaviors than do sheep, which is incredibly important for a model of disease where one of the major symptoms is the progressive decline in cognitive ability. However, one does not establish a large animal model colony without taking into account temperament. Of the two potential canine *CLN5*-deficient dog models, Border Collies are noted for their work ethic and high energy. Golden Retrievers, on the other hand, are easy going and would likely be very adaptable to life in a research colony. Studies involving a colony of *CLN5*-deficient Golden Retrievers would allow for the study of

CLN5drug treatment efficacy as well as to allow furthering of understanding of the biology of CLN5 disease phenotype.

The third finding was a 22bp deletion and frameshift in *RBM20* that causes Dilated Cardiomyopathy (DCM) in Standard Schnauzers.[19] RBM20 encodes RNA binding motif protein 20, a spliceosomal protein highly expressed in the heart. Mutations in RBM20 are known to cause DCM in human patients, generally in a five amino acid hotspot in exon 6.[163,183,185,186,165] In humans, the age at onset can range from infancy through late adulthood with most patients diagnosed from 20-50 years old[187]. Standard Schnauzers with DCM have a median age at DCM diagnosis of 1.6 years of age.[160] The clinical manifestation of DCM in Standard Schnauzers is earlier than that seen within most human DCM cases caused by RBM20 mutations, which we propose makes them an ideal large-animal model for the study of possible clinical interventions. This early manifestation would allow astute researchers to identify homozygous carriers of the disease early on (as early as canine disease mutation screening would allow) and follow gene transcription and RNASeq of left ventricular cardiac tissue at serial timelines. At the time of the serial biopsies, echocardiograms could be taken to show heart function and possible disease progression. Studies such as these could yield biomarkers that could lead to human medical treatments for human DCM. The role that RBM20 mutations show via aberrant splicing could function as diagnostic tools for earlier treatment. It is plausible that many of the same aberrant splicing events that take place in human DCM patients take place within the *RBM20*-deficient Standard Schnauzers.

Our lab is currently using whole genome sequencing with DNA from dogs affected with a wide range of heritable canine diseases. DNA sequencing technology has improved rapidly over the last few years, allowing for increasingly inexpensive sequencing and the reduction in time

required to generate them. In 2014, the canine reference sequence was updated to include RNAseq data on many canine tissues, allowing for better annotation. Technology has improved across the board; from sequencers to annotations, to software to chemistries. It is clear to us that whole genome sequencing will soon be a key tool for the discovery of disease-causing variants in dogs as well as other species. However, it is not a complete tool as I'll illustrate below.

Of the 145 dogs of which our lab has sequences, we've only found causal mutations for roughly 30% of them. This is due to a wide variety of reasons, which are seen in the study of genetic diseases in many species. Many of the dogs in our study cohort show an autosomal recessive mode of inheritance. As a result of our study methodologies, we look within coding regions of genes. This is not the only location of plausible causal mutations, nor is it likely to yield the best results. Our lab and every other lab conducting similar studies simply picks from the low-hanging fruit of Mendelian genetics. Not all diseases will have causal mutations that reside within the coding regions. Some may be caused by structural variation. Some may be caused by copy-number variation. Some disease phenotype could be due to gene regulation from RNAi, or microRNA interference. Furthermore, it's possible that epigenetic changes currently misunderstood or unknown may hold the key to discovery.

Luckily for our lab, answers for some of these quandaries may be forthcoming or already here. Usher Syndrome, the leading cause of inherited deaf-blindness in humans, is divided into three subtypes. Usher type 2 is the most common and mutations in the USH2A gene account for up to 80% of all cases.[188] A significant portion of those who suffer from Usher type 2 have only one heterozygous mutation in USH2A; Steele-Stallard et al.[188] tested for intronic mutations in nine genes that, when aberrant, are shown to cause Usher Syndrome phenotype and identified a second mutation in 35% of cases, including a common intronic mutation. [188] It is possible that

not just complex-heterozygosity of a single gene could play a role in disease, but complex heterozygosity of a panel of possible genes may lead to disease phenotype within some of the dogs in our disease cohort. It should be possible to screen the WGSs of our dogs and filter for certain subsets of genes in order to determine whether mutations in separate genes could lead to disease phenotype.

Furthermore, the study of copy number variation could yield positive results in the search for causal mutations. According to Lobo[189], copy number variation is an often overlooked cause of disease. Recently, Wang et al.[190] showed that one can use WGS and statistical tools (as well as their free, proprietary protocol) in order to predict/detect copy number variation. Integrating this or a similar tool for detecting copy would allow our lab to possibly find the genetic cause of disease, possibly, within some of those who have been sequenced but no causal mutation determined. In another study, Chen et al[191] used WGS and a novel algorithm in order to detect copy number variants as well as analysis of coverage depth to detect structural variation. They applied it to cancer genetics and found novel CNVs along with complex rearrangements and CNVs that had been missed previously by other methods.[191]

Another contributing plausible cause for the inherited diseases in the subject dogs that we haven't found yet may lie in the field of RNAi (RNA interference), a form of gene regulation previously mostly a niche section of genetics until it was discovered to be the cause of Fragile-X syndrome in 2002.[192] As recently as post-author's dissertation defense, Penso-Dolfin et al.[193] published an analysis and release of an improved microRNA annotation for the canine genome. They identified novel miRNAs, which are known to play a critical role in regulating gene expression. Knowing the sequence of these miRNAs will allow us to elucidate how they function

to regulate gene expression as well as show how disease phenotype can arise when this function

is aberrant.

Table: Disease-causing mutations discovered as part of the thesis research.

Dog Breed(s)	Disease	Molecular Method	Mutant Gene	Corresponding Human Disease
Russell-Group Terriers	SAMS	Whole Genome Sequencing	KCNJ10	SeSAME/EAST
Golden Retrievers	Neuronal Ceroid Lipofuscinosis	Whole Genome Sequencing	CLN5	Neuronal Ceroid Lipofuscinosis
Standard Schnauzers	Dilated Cardiomyopathy	Whole Genome Sequencing	RBM20	Dilated Cardiomyopathy

Literature Cited

1. Vilà C, Savolainen P, Maldonado JE, et al. Multiple and ancient origins of the domestic dog. *Science*. 1997;276(5319):1687-1689. doi:10.1126/science.276.5319.1687.

2. Ostrander EA, Galibert F, Patterson DF. Canine genetics comes of age. *Trends Genet*. 2000;16(3):117-124. doi:10.1016/S0168-9525(99)01958-7.

3. Savolainen P, Zhang Y, Luo J, Lundeberg J, Leitner T. Genetic evidence for an East Asian origin of domestic dogs. *Science (80-)*. 2002;298(5598):1610-1613. doi:10.1126/science.1073906.

4. Wayne R, Geffen E, Girman DJ, Koepfli KP, Lau LM, Marshall CR. MOLECULAR SYSTEMATICS OF THE CANIDAE. *Syst Biol*. 1997;46(4):622-653. doi:doi:10.1093/sysbio/46.4.622.

5. Ostrander EA KL. Unleashing the Canine Genome. *Genome Res*. 2000;10(9):1271-1274.

6. Asher L, Diesel G, Summers JF, McGreevy PD CL. Inherited defects in pedigree dogs. Part 1: Disorders related to breed standards. *Vet J*. 2009;182(3):402-411.

7. Nowend KL, Starr-Moss AN et al. The function of dog models in developing gene therapy strategies for human health. *Mammal Genome*. 2011;22:476-485.

8. Lequarre AS, Andersson L, Andre C, et al. LUPA: A European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. *Vet J*. 2011;189(2):155-159. doi:10.1016/j.tvjl.2011.06.013.

9. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 2005;438(7069):803-819.

10. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008;299(11):1335-1344. doi:10.1001/jama.299.11.1335.

11. Ostrander E a, Beale HC. Leading the way: finding genes for neurologic disease in dogs using genome-wide mRNA sequencing. *BMC Genet*. 2012;13:56. doi:10.1186/1471-2156-13-56.

12. Forman OP, De Risio L, Stewart J, Mellersh CS, Beltran E. Genome-wide mRNA sequencing of a single canine cerebellar cortical degeneration case leads to the identification of a disease associated SPTBN2 mutation. *BMC Genet*. 2012;13(1):55. doi:10.1186/1471-2156-13-55.

13. Broeckx BJ, Coopman F, Verhoeven GE, Bavegems V, De Keulenaer S, De Meester E, Van Niewerburgh F DD. Development and performance of a targeted whole exome sequencing enrichment kit for the dog (Canis Familiaris Build 3.1). *Sci Rep*. 2014;(4):5597.

14. Bahassi EM, Stambrook PJ. Next-generation sequencing technologies: Breaking the sound

barrier of human genetics. *Mutagenesis*. 2014;29(5):303-310. doi:10.1093/mutage/geu031.

15. Priest JR, Ceresnak SR, Dewey FE, et al. Molecular diagnosis of long QT syndrome at 10 days of life by rapid whole genome sequencing. *Heart Rhythm*. 2014;11(10):1707-1713. doi:10.1016/j.hrthm.2014.06.030.

16. Ku CS, Polychronakos C, Tan EK, Naidoo N, Pawitan Y, Roukos DH, Mort M CD. A new paradigm emerges from the study of de novo mutations in the context of neurodevelopmental disease. *Mol Psychiatry*. 2013;18(2):141-153.

17. Metzker M. Sequencing technologies of the next generation. *Nat Rev Genet*. 21010;11(1):31-46.

18. Awano T, Johnson GS, Wade CM, et al. Genome-wide association analysis reveals a SOD1 mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A*. 2009;106:2794-2799. doi:10.1073/pnas.0812297106.

19. Gilliam D, Harmon MW, Gandolfi B, Johnson GS, Mhlanga-Mutangadura T, Hansen L, Taylor JF, Schnabel RD LS. Standard Schnauzer dogs with Dilated Cardiomyopathy have a 22 bp Deletion and Frameshift in RBM20. *Unkn J*. 2016.

20. Sanders DN, Zeng R. et al. GM2 ganaliosidosis associated with a HEXA missense mutation in japanese chin dogs: a potential model for Tay Sachs disease. *Mol Genet Metab*. 2013;108:70-75.

21. Farias FH, Zeng R, Johnson GS, Shelton GD, Paquette D, O'Brien DP. A L2HGDH initiator methionine codon mutation in a Yorkshire terrier with L-2-hydroxyglutaric aciduria. *BMC Vet Res*. 2012;8:124. doi:10.1186/1746-6148-8-124.

22. Katz ML, Khan S, Awano T, Shahid SA, Siakotos AN, Johnson GS. A mutation in the CLN8 gene in English Setter dogs with neuronal ceroid-lipofuscinosis. *Biochem Biophys Res Commun*. 2005;327(2):541-547. doi:10.1016/j.bbrc.2004.12.038.

23. Farias FHG, Zeng R, Johnson GS, et al. A truncating mutation in ATP13A2 is responsible for adult-onset neuronal ceroid lipofuscinosis in Tibetan terriers. *Neurobiol Dis*. 2011;42(3):468-474. doi:10.1016/j.nbd.2011.02.009.

24. Gilliam D, Kolicheski A, Johnson GS, et al. Golden Retriever dogs with neuronal ceroid lipofuscinosis have a two-base-pair deletion and frameshift in CLN5. *Mol Genet Metab*. 2015;115(2-3):101-109. doi:10.1016/j.ymgme.2015.04.001.

25. Guo J, Johnson GS, Brown HA, et al. A CLN8 nonsense mutation in the whole genome sequence of a mixed breed dog with neuronal ceroid lipofuscinosis and Australian shepherd ancestry. *Mol Genet Metab*. 2014;112(4):302-309. doi:10.1016/j.ymgme.2014.05.014.

26. Guo J, O'Brien DP, Mhlanga-Mutangadura T, et al. A rare homozygous MFSD8 single-base-pair deletion and frameshift in the whole genome sequence of a Chinese Crested dog with neuronal ceroid lipofuscinosis. *BMC Vet Res*. 2014;10:960. doi:10.1186/s12917-014-

0181-z.

27.   Chen X, Johnson GS, Schnabel RD, et al. A neonatal encephalopathy with seizures in standard poodle dogs with a missense mutation in the canine ortholog of ATF2. *Neurogenetics*. 2008;9(1):41-49. doi:10.1007/s10048-007-0112-2.

28.   Zeng R, Farias FHG, Johnson GS, et al. A Truncated Retrotransposon Disrupts the GRM1 Coding Sequence in Coton de Tulear Dogs with Bandera's Neonatal Ataxia. *J Vet Intern Med*. 2011;25(2):267-272. doi:10.1111/j.1939-1676.2010.0666.x.

29.   Farias FHG, Johnson GS, Taylor JF, et al. An ADAMTS17 splice donor site mutation in dogs with primary lens luxation. *Investig Ophthalmol Vis Sci*. 2010;51(9):4716-4721. doi:10.1167/iovs.09-5142.

30.   Rohdin C, Gilliam D, O'Leary CA, O'Brien DP, Coates JR, Johnson GS JK. A KCNJ10 mutation previously identified in the Russell group of terriers also occurs in Smooth-Haired Fox Terriers with hereditary ataxia and in related breeds. *Acta Vet Scand*. 2015;23(57):26.

31.   Hartley WJ PA. Ataxia in Jack Russell Terriers. *Acta Neuropathol*. 1973;26:71-74.

32.   Coates JR, Carmichael KP, Shelton GD, O'Brien DP and JG. Preliminary characterization of a cerebellar ataxia in Jack Russell Terriers. *J Vet Int Med*. 1996;10:176.

33.   Bhatti SF, Vanhaesebrouck AE, Van S, I et al. Myokymia and neuromyotonia in 37 Jack Russell terriers. *Vet J*. 2011;189:284-288.

34.   Simpson K, Eminaga S CG. Hereditary ataxia in Jack Russell terriers in the UK. *VetRec*. 2012;170:548.

35.   Vanhaesebrouck AE, van Soens I, Poncelet L, et al. Clinical and electrophysiological characterization of myokymia and neuromyotonia in Jack Russell Terriers. *J Vet Intern Med*. 2010;24(4):882-889. doi:10.1111/j.1939-1676.2010.0525.x.

36.   Wessmann A, Goedde T, Fischer A, et al. Hereditary ataxia in the Jack Russell Terrier--clinical and genetic investigations. *J Vet Intern Med*. 2004;18(4):515-521. doi:10.1111/j.1939-1676.2004.tb02577.x.

37.   Van Ham L, Bhatti S, Polis I, Fatzer R, Braund K, Thoonen H. "Continuous muscle fibre activity" in six dogs with episodic myokymia, stiffness and collapse. *Vet Rec*. 2004;155(24):769-774. doi:10.1136/vr.155.24.769.

38.   Browne DL, Gancher ST, Nutt JG et al. Episodic ataxia/myokymia syndrome is associated with point mutations in the human potassium channel gene, KCNA1. *Nature-Genetics*. 1994;8:136-140.

39.   Dedek K, Kunath B, Kananura C et al. Myokymia and neonatal epilepsy caused by a mutation in the voltage sensor of the KCNQ2 K+ channel. *Proc Natl Acad Sci U S A*. 2001;98:12272-12277.

40. Van Poucke M, Vanhaesebrouck AE, Peelman LJ, Van Ham L. Experimental validation of in silico predicted KCNA1, KCNA2, KCNA6 and KCNQ2 genes for association studies of peripheral nerve hyperexcitability syndrome in Jack Russell Terriers. *Neuromuscul Disord*. 2012;22(6):558-565. doi:10.1016/j.nmd.2012.01.008.

41. Forman OP, De Risio L, Mellersh CS. Missense Mutation in CAPN1 Is Associated with Spinocerebellar Ataxia in the Parson Russell Terrier Dog Breed. *PLoS One*. 2013;8(5):1-8. doi:10.1371/journal.pone.0064627.

42. Zimin A V., Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler. *Bioinformatics*. 2013;29(21):2669-2677. doi:10.1093/bioinformatics/btt476.

43. Adzhubei IA, Schmidt S, Peshkin L et al. A method and server for predicting damaging missense mutations. *Nature-Methods*. 2010;7:248-249.

44. Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. In: *Genetic Analysis - Biomolecular Engineering*. Vol 14. ; 1999:143-149. doi:10.1016/S1050-3862(98)00019-9.

45. Sims MH, Moore RE. Auditory-evoked response in the clinically normal dog: Early latency components. *Am J Vet Res*. 1984;45(10):2019-2027.

46. Walker TL, Redding RW, Braund KG. Motor nerve conduction velocity and latency in the dog. *Am J Vet Res*. 1979;40(10):1433-1439. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=525865.

47. Gutman G a, Chandy KG, Adelman JP, et al. International Union of Pharmacology. XLI. Compendium of voltage-gated ion channels: potassium channels. *Pharmacol Rev*. 2003;55(4):583-586. doi:10.1124/pr.55.4.9.

48. Djukic B, Casper KB, Philpot BD et al. Conditional knock-out of Kir4.1 leads to glial membrane depolarization, inhibition of potassium and glutamate uptake, and enhanced short-term synaptic potentiation. *J Neurosci*. 2007;27:11354-11365.

49. Hibino H, Inanobe A, Furutani K, Murakami S, Findlay I, Kurachi Y. Inwardly rectifying potassium channels: their structure, function, and physiological roles. *Physiol Rev*. 2010;90(1):291-366. doi:10.1152/physrev.00021.2009.

50. Neusch C, Rozengurt N, Jacobs RE, Lester H a, Kofuji P. Kir4.1 potassium channel subunit is crucial for oligodendrocyte development and in vivo myelination. *J Neurosci*. 2001;21(15):5429-5438. doi:21/15/5429 [pii].

51. Scholl UI, Choi M, Liu T, et al. Seizures, sensorineural deafness, ataxia, mental retardation, and electrolyte imbalance (SeSAME syndrome) caused by mutations in KCNJ10. *Proc Natl Acad Sci U S A*. 2009;106(14):5842-5847. doi:10.1073/pnas.0901749106.

52. Bockenhauer D, Feather S, Stanescu HC, et al. Epilepsy, ataxia, sensorineural deafness,

tubulopathy, and KCNJ10 mutations. *N Engl J Med*. 2009;360:1960-1970. doi:10.1016/S1090-3798(09)70148-5.

53. Vanhaesebrouck AE, Bhatti SFM, Franklin RJM et al. Myokymia and neuromyotonia in veterinary medicine: A comparison with peripheral nerve hyperexcitability syndrome in humans. *Vet J*. 2013;197:153-162.

54. Ando M, Takeuchi S. Immunological identification of an inward rectifier K+ channel (Kir4.1) in the intermediate cell (melanocyte) of the cochlear stria vascularis of gerbils and rats. *Cell Tissue Res*. 1999;298(1):179-183. doi:10.1007/s004419900066.

55. Rozengurt N, Lopez I, Chiu CS et al. Time course of inner ear degeneration and deafness in mice lacking the Kir4.1 potassium channel subunit. *HearRes*. 2003;177:71-80.

56. Kaji M SH. Myokymia. *Nihon Rinsho*. 1993;51:2866-2870.

57. Vit JP, Ohara PT, Bhargava A et al. Silencing the Kir4.1 potassium channel subunit in satellite glial cells of the rat trigeminal ganglion results in pain-like behavior in the absence of nerve injury. *J Neurosci*. 2008;28:4161-4171.

58. Bandulik S, Schmidt K, Bockenhauer D, et al. The salt-wasting phenotype of EAST syndrome, a disease with multifaceted symptoms linked to the KCNJ10 K+ channel. *Pflugers Arch Eur J Physiol*. 2011;461(4):423-435. doi:10.1007/s00424-010-0915-0.

59. Kofuji P, Ceelen P, Zahs KR, Surbeck LW, Lester HA, Newman EA. Genetic inactivation of an inwardly rectifying potassium channel (Kir4.1 subunit) in mice: phenotypic impact in retina. *J Neurosci*. 2000;20(15):5733-5740. doi:10.1037/a0013262.Open.

60. Thompson DA, Feather S, Stanescu HC, et al. Altered electroretinograms in patients with KCNJ10 mutations and EAST syndrome. *J Physiol*. 2011;589(Pt 7):1681-1689. doi:10.1113/jphysiol.2010.198531.

61. Cross JH, Arora R, Heckemann RA, et al. Neurological features of epilepsy, ataxia, sensorineural deafness, tubulopathy syndrome. *Dev Med Child Neurol*. 2013;55(9):846-856. doi:10.1111/dmcn.12171.

62. Dr??gem??ller M, Jagannathan V, Howard J, et al. A frameshift mutation in the cubilin gene (CUBN) in Beagles with Imerslund-Gr??sbeck syndrome (selective cobalamin malabsorption). *Anim Genet*. 2014;45(1):148-150. doi:10.1111/age.12094.

63. Owczarek-Lipska M, Jagannathan V, Drogemuller C et al. A frameshift mutation in the cubilin gene (CUBN) in Border Collies with Imerslund-Grasbeck syndrome (selective cobalamin malabsorption). *PLoS One*. 2008;8:e61144.

64. Haltia M, Goebel HH. The neuronal ceroid-lipofuscinoses: A historical introduction. *Biochim Biophys Acta - Mol Basis Dis*. 2013;1832(11):1795-1800. doi:10.1016/j.bbadis.2012.08.012.

65. Anderson GW, Goebel HH, Simonati A. Human pathology in NCL. *Biochim Biophys Acta - Mol Basis Dis*. 2013;1832(11):1807-1826. doi:10.1016/j.bbadis.2012.11.014.

66. Mink JW, Augustine EF, Adams HR, Marshall FJ KJ. Genetic basis and phenotypic correlations of the neuronal ceroid lipofusinoses. *J Child Neurol*. 2013;28(9):1101-1105.

67. Vesa J, Hellsten E, Verkruyse L a, et al. Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. *Nature*. 1995;376:584-587. doi:10.1038/376584a0.

68. Doggett NA, Meincke LJ, Liu Z, et al. Isolation of a novel gene underlying Batten disease, CLN3. The International Batten Disease Consortium. *Cell*. 1995;82(6):949-957. doi:0092-8674(95)90274-0 [pii].

69. Cotman SL, Karaa A, Staropoli JF SK. Neuronal ceroid lipofuscinosis: impact of recent genetic advances and expansion of the clinicopathologic spectrum. *Curr Neurol Neurosci Rep*. 2013;13(8):366.

70. Berkovic SF, Carpenter S, Andermann F, Andermann E, Wolfe LS. Kufs' disease: a critical reappraisal. *Brain*. 1988;111 (Pt 1:27-62. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3284607.

71. Williams R. *Appendix 1: NCL Incidence and Prevalence Data. In:The Neuronal Ceroid Lipofuscinoses (Batten Disease) 2nd Edition*.; 2011.

72. Bond M, Holthaus SMK, Tammen I, Tear G, Russell C. Use of model organisms for the study of neuronal ceroid lipofuscinosis. *Biochim Biophys Acta - Mol Basis Dis*. 2013;1832(11):1842-1865. doi:http://dx.doi.org/10.1016/j.bbadis.2013.01.009.

73. Sanders DN, Farias FH, Johnson GS, et al. A mutation in canine PPT1 causes early onset neuronal ceroid lipofuscinosis in a Dachshund. *Mol Genet Metab*. 2010;100(4):349-356. doi:10.1016/j.ymgme.2010.04.009.

74. Awano T, Katz ML, O'Brien DP, et al. A frame shift mutation in canine TPP1 (the ortholog of human CLN2) in a juvenile Dachshund with neuronal ceroid lipofuscinosis. *Mol Genet Metab*. 2006;89(3):254-260. doi:10.1016/j.ymgme.2006.02.016.

75. Melville SA, Wilson CL, Chiang CS, Studdert VP, Lingaas F, Wilton AN. A mutation in canine CLN5 causes neuronal ceroid lipofuscinosis in Border collie dogs. *Genomics*. 2005;86(3):287-294. doi:10.1016/j.ygeno.2005.06.005.

76. Katz ML, Farias FH, Sanders DN, et al. A missense mutation in canine CLN6 in an Australian Shepherd with neuronal ceroid lipofuscinosis. *J Biomed Biotechnol*. 2011;2011. doi:10.1155/2011/198042.

77. Awano T, Katz ML, O'Brien DP, et al. A mutation in the cathepsin D gene (CTSD) in American Bulldogs with neuronal ceroid lipofuscinosis. *Mol Genet Metab*. 2006;87(4):341-348. doi:10.1016/j.ymgme.2005.11.005.

78. Wöhlke A, Philipp U, Bock P, et al. A one base pair deletion in the canine ATP13A2 gene causes exon skipping and late-onset neuronal ceroid lipofuscinosis in the Tibetan terrier. *PLoS Genet*. 2011;7(10). doi:10.1371/journal.pgen.1002304.

79. Abitbol M, Thibaud JL, Olby NJ, et al. A canine Arylsulfatase G (ARSG) mutation leading to a sulfatase deficiency is associated with neuronal ceroid lipofuscinosis. *Proc Natl Acad Sci U S A*. 2010;107(33):14775-14780. doi:10.1073/pnas.0914206107.

80. Kowalewski B, Lamanna WC, Lawrence R, et al. Arylsulfatase G inactivation causes loss of heparan sulfate 3-O-sulfatase activity and mucopolysaccharidosis in mice. *Proc Natl Acad Sci*. 2012;109(26):10310-10315. doi:10.1073/pnas.1202071109.

81. Kowalewski B, L??bke T, Kollmann K, et al. Molecular characterization of arylsulfatase G expression, processing, glycosylation, transport and activity. *J Biol Chem*. 2014;289(40):27992-28005. doi:10.1074/jbc.M114.584144.

82. Studdert VP, Mitten RW. Clinical features of ceroid lipofuscinosis in border collie dogs. *Aust Vet J*. 1991;68(4):137-140. http://www.ncbi.nlm.nih.gov/pubmed/2069541.

83. Mizukami K, Kawamichi T, Koie H, et al. Neuronal Ceroid Lipofuscinosis in Border Collie Dogs in Japan: Clinical and Molecular Epidemiological Study (2000–2011). *Sci World J*. 2012;2012:1-7. doi:10.1100/2012/383174.

84. Kluth S, Eckardt J, Distl O. Selection response to DNA testing for canine ceroid lipofuscinosis in Tibetan terriers. *Vet J*. 2014;201(3):433-434. doi:10.1016/j.tvjl.2014.05.029.

85. Evans J, Katz ML, Levesque D, Shelton GD, de Lahunta A, O'Brien D. A variant form of neuronal ceroid lipofuscinosis in American bulldogs. *J Vet Intern Med*. 2005;19(1):44-51. doi:15715047.

86. Vuillemenot BR, Kennedy D, Cooper JD, et al. Nonclinical evaluation of CNS-administered TPP1 enzyme replacement in canine CLN2 neuronal ceroid lipofuscinosis. *Mol Genet Metab*. 2015;114(2):281-293. doi:10.1016/j.ymgme.2014.09.004.

87. Katz ML, Coates JR, Sibigtroth CM, et al. Enzyme replacement therapy attenuates disease progression in a canine model of late-infantile neuronal ceroid lipofuscinosis (CLN2 disease). *J Neurosci Res*. 2014;92(11):1591-1598. doi:10.1002/jnr.23423.

88. Whiting REH, Narfstr??m K, Yao G, et al. Enzyme replacement therapy delays pupillary light reflex deficits in a canine model of late infantile neuronal ceroid lipofuscinosis. *Exp Eye Res*. 2014;125:164-172. doi:10.1016/j.exer.2014.06.008.

89. Sofroniew M V., Vinters H V. Astrocytes: Biology and pathology. *Acta Neuropathol*. 2010;119(1):7-35. doi:10.1007/s00401-009-0619-8.

90. Wininger FA, Zeng R, Johnson GS, Katz ML, Johnson GC, Bush WW, Jarboe JM CJ. Degenerative myelopathy in a bernese mountain dog with a novel SOD1 missense mutation. *J Vet Intern Med*. 2011;25:1166-1170.

91. Hoeppner MP, Lundquist A, Pirun M, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*. 2014;9(3). doi:10.1371/journal.pone.0091172.

92.   Katz ML, Sanders DA, Sanders DN, Hansen EA, Johnson GS. Assessment of plasma carnitine concentrations in relation to ceroid lipofuscinosis in Tibetan Terriers. *Am J Vet Res*. 2002;63(6):890-895.

93.   Savukoski M, Klockars T, Holmberg V, Santavuori P, Lander ES, Peltonen L. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. *Nat Genet*. 1998;19(3):286-288. doi:10.1038/975.

94.   Kopra O, Vesa J, von Schantz C, et al. A mouse model for Finnish variant late infantile neuronal ceroid lipofuscinosis, CLN5, reveals neuropathology associated with early aging. *Hum Mol Genet*. 2004;13(23):2893-2906. doi:10.1093/hmg/ddh312.

95.   Houweling PJ, Cavanagh JAL, Palmer DN, et al. Neuronal ceroid lipofuscinosis in Devon cattle is caused by a single base duplication (c.662dupG) in the bovine CLN5 gene. *Biochim Biophys Acta - Mol Basis Dis*. 2006;1762(10):890-897. doi:10.1016/j.bbadis.2006.07.008.

96.   Frugier T, Mitchell NL, Tammen I, Houweling PJ, Arthur DG, Kay GW, van Diggelen OP, Jolly RD PD. A new large animal model of CLN5 neuronal ceroid lipofuscinosis in Borderdale sheep is caused by a nucleotide substitution at a consensus splice site (c.571+1G>A) leading to excision of exon 3. *Neurobiol Dis*. 2008;29(2):306-315.

97.   Vesa J, Chin MH, Oelgeschläger K, et al. Neuronal ceroid lipofuscinoses are connected at molecular level: interaction of CLN5 protein with CLN2 and CLN3. *Mol Biol Cell*. 2002;13(7):2410-2420. doi:10.1091/mbc.E02-01-0031.

98.   Schmiedt ML, Bessa C, Heine C, Ribeiro MG, Jalanko A, Kytt??l?? A. The neuronal ceroid lipofuscinosis protein CLN5: New insights into cellular maturation, transport, and consequences of mutations. *Hum Mutat*. 2010;31(3):356-365. doi:10.1002/humu.21195.

99.   Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785-786. doi:10.1038/nmeth.1701.

100.   Isosomppi J, Vesa J, Jalanko A, Peltonen L. Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. *Hum Mol Genet*. 2002;11(8):885-891. doi:11971870.

101.   Holmberg V, Jalanko A, Isosomppi J, Fabritius AL, Peltonen L, Kopra O. The mouse ortholog of the neuronal ceroid lipofuscinosis CLN5 gene encodes a soluble lysosomal glycoprotein expressed in the developing brain. *Neurobiol Dis*. 2004;16(1):29-40. doi:10.1016/j.nbd.2003.12.019.

102.   Moharir A, Peck SH, Budden T, Lee SY. The Role of N-Glycosylation in Folding, Trafficking, and Functionality of Lysosomal Protein CLN5. *PLoS One*. 2013;8(9). doi:10.1371/journal.pone.0074299.

103.   Larkin H, Ribeiro MG, Lavoie C. Topology and Membrane Anchoring of the Lysosomal Storage Disease-Related Protein CLN5. *Hum Mutat*. 2013;34(12):1688-1697. doi:10.1002/humu.22443.

104. Santavuori P, Rapola J, Nuutila A, et al. The spectrum of Jansky-Bielschowsky disease. *Neuropediatrics*. 1991;22(2):92-96. doi:10.1055/s-2008-1071423.

105. Kousi M, Lehesjoki AE, Mole SE. Update of the mutation spectrum and clinical correlations of over 360 mutations in eight genes that underlie the neuronal ceroid lipofuscinoses. *Hum Mutat*. 2012;33(1):42-63. doi:10.1002/humu.21624.

106. Cannelli N, Nardocci N, Cassandrini D, et al. Revelation of a novel CLN5 mutation in early juvenile neuronal ceroid lipofuscinosis. *Neuropediatrics*. 2007;38(1):46-49. doi:10.1055/s-2007-981449.

107. Xin W, Mullen TE, Kiely R, et al. CLN5 mutations are frequent in juvenile and late-onset non-finnish patients with NCL. *Neurology*. 2010;74(7):565-571. doi:10.1212/WNL.0b013e3181cff70d.

108. Mancini C, Nassani S, Guo Y, et al. Adult-onset autosomal recessive ataxia associated with neuronal ceroid lipofuscinosis type 5 gene (CLN5) mutations. *J Neurol*. 2014;262(1):173-178. doi:10.1007/s00415-014-7553-y.

109. von Schantz C, Kielar C, Hansen SN, et al. Progressive thalamocortical neuron loss in Cln5 deficient mice: Distinct effects in Finnish variant late infantile NCL. *Neurobiol Dis*. 2009;34(2):308-319. doi:10.1016/j.nbd.2009.02.001.

110. Schmiedt ML, Blom T, Blom T, et al. Cln5-deficiency in mice leads to microglial activation, defective myelination and changes in lipid metabolism. *Neurobiol Dis*. 2012;46(1):19-29. doi:10.1016/j.nbd.2011.12.009.

111. Hughes SM, Hope KM, Xu JB, Mitchell NL, Palmer DN. Inhibition of storage pathology in prenatal CLN5-deficient sheep neural cultures by lentiviral gene therapy. *Neurobiol Dis*. 2014;62:543-550. doi:10.1016/j.nbd.2013.11.011.

112. Cuenca S, Ruiz-Cano MJ, Gimeno-Blanes JR, Jurado A, Salas C, Gomez-Diaz I, Padron-Barthe L, Grillo JJ, Vilches C, Segovia J, Pascual-Figal D, Lara-Pezzi E, Monserrat L, Alonso-Pulpon L G-PP. Genetic basis of familial dilated cardiomyopathy patients undergoing heart transplantation. *J Hear Lung Transpl*. 2016:DOI: http://dx.doi.org/10.1016/j.healun.2015.12.01.

113. Towbin JA. Inherited cardiomyopathies. *Circ J*. 2014;78(10):2347-2356. doi:DN/JST.JSTAGE/circj/CJ-14-0893 [pii].

114. Simpson S, Edwards J, Ferguson-Mignan TFN, Cobb M, Mongan NP, Rutland CS. Genetics of human and canine dilated cardiomyopathy. *Int J Genomics*. 2015;2015. doi:10.1155/2015/204823.

115. Legge CH, López a, Hanna P, Côté E, Hare E, Martinson S a. Histological characterization of dilated cardiomyopathy in the juvenile toy Manchester terrier. *Vet Pathol*. 2013;50(6):1043-1052. doi:10.1177/0300985813480509.

116. Alroy J, Rush JE, Freeman L, et al. Inherited infantile dilated cardiomyopathy in dogs: Genetic, clinical, biochemical, and morphologic findings. *Am J Med Genet*.

2000;95(1):57-66. doi:10.1002/1096-8628(20001106)95:1<57::AID-AJMG12>3.0.CO;2-O.

117. Martin MWS, Stafford Johnson MJ, Strehlau G, King JN. Canine dilated cardiomyopathy: A retrospective study of prognostic findings in 367 clinical cases. *J Small Anim Pract*. 2010;51(8):428-436. doi:10.1111/j.1748-5827.2010.00966.x.

118. Schatzberg SJ, Olby NJ, Breen M, et al. Molecular analysis of a spontaneous dystrophin "knockout" dog. *Neuromuscul Disord*. 1999;9(5):289-295. doi:10.1016/S0960-8966(99)00011-5.

119. Meurs KM, Fox PR, Norgard M, et al. A prospective genetic evaluation of familial dilated cardiomyopathy in the Doberman pinscher. *J Vet Intern Med*. 2007;21(5):1016-1020. doi:10.1892/0891-6640(2007)21[1016:APGEOF]2.0.CO;2.

120. Distl O, Vollmar a C, Broschk C, Hamann H, Fox PR. Complex segregation analysis of dilated cardiomyopathy (DCM) in Irish wolfhounds. *Heredity (Edinb)*. 2007;99(4):460-465. doi:10.1038/sj.hdy.6801024.

121. Meurs KM, Hendrix KP, Norgard MM. Molecular evaluation of five cardiac genes in Doberman Pinschers with dilated cardiomyopathy. *Am J Vet Res*. 2008;69(8):1050-1053. doi:10.2460/ajvr.69.8.1050.

122. Tidholm A, Jönsson L, Jonsson L. Histologic Characterization of Canine Dilated Cardiomyopathy. *Vet Pathol*. 2005;42(1):1-8. doi:10.1354/vp.42-1-1.

123. Lobo L, Carvalheira J, Canada N, Bussadori C, Gomes JL, Faustino AMR. Histologic Characterization of Dilated Cardiomyopathy in Estrela Mountain Dogs. *Vet Pathol*. 2010;47(4):637-642. doi:10.1177/0300985810364511.

124. Philipp U VA and DO. Evaluation of six candidate genes for dilated cardiomyopathy in Irish wolfhounds. *Anim Genet*. 2008;39:88-89.

125. Werner P, Raducha MG, Prociuk U, Sleeper MM, Van Winkle TJ, Henthorn PS. A novel locus for dilated cardiomyopathy maps to canine chromosome 8. *Genomics*. 2008;91(6):517-521. doi:10.1016/j.ygeno.2008.03.007.

126. Mausberg TB, Wess G, Simak J, Keller L, Drogemuller M, Drogemuller C, Webster MT, Stephenson H D-MJ and LT. A locus on chromosome 5 is associated with dilated cardiomyopathy in Doberman Pinschers. *PLoS One*. 2011;(6):e20042.

127. Philipp U, Vollmar A, Häggström J, Thomas A, Distl O. Multiple loci are associated with dilated cardiomyopathy in irish wolfhounds. *PLoS One*. 2012;7(6). doi:10.1371/journal.pone.0036691.

128. Wiersma AC, Stabej P, Leegwater PA, Van Oost BA OW and D-MJ. Evaluation of 15 candidate genes for dilated cardiomyopathy in the Newfoundland dog. *J Hered*. 2008;99:73-80.

129. Stabej P, Leegwater PA, Imholz S, Versteeg SA, Zijlstra C, Stokhof AA D-PA and van

OB. The canine sarcoglycan delta gene: BAC clone contig assembly, chromosome assignment and interrogation as a candidate gene for dilated cardiomyopathy in Dobermann dogs. *Cytogenet Genome Res*. 2005;111:140-146.

130. Stabej P, Imholz S, Versteeg SA, Zijlstra C, Stokhof AA, Domanjko-Petric A LP and van OB. Characterization of the canine desmin (DES) gene and evaluation as a candidate gene for dilated cardiomyopathy in the Dobermann. *Gene*. 2004;340:241-249.

131. Philipp U VA and DO. Evaluation of the titin-cap gene (TCAP) as candidate for dilated cardiomyopathy in Irish wolfhounds. *Anim Biotechnol*. 2008;19:231-236.

132. Lynne O'Sullivan M, O'Grady MR, Glen Pyle W, Dawson JF. Evaluation of 10 genes encoding cardiac proteins in Doberman Pinschers with dilated cardiomyopathy. *Am J Vet Res*. 2011;72(7):932-938. doi:10.2460/ajvr.72.7.932.

133. Philipp U, Broschk C, Vollmar A, Distl O. Evaluation of tafazzin as candidate for dilated cardiomyopathy in Irish wolfhounds. In: *Journal of Heredity*. Vol 98. ; 2007:506-509. doi:10.1093/jhered/esm045.

134. Stabej P, Leegwater PA, Stokhof AA, Domanjko-Petrič A, van Oost BA. Evaluation of the phospholamban gene in purebred large-breed dogs with dilated cardiomyopathy. *Am J Vet Res*. 2005;66(3):432-436. doi:10.2460/ajvr.2005.66.432.

135. Meurs KM, Magnon AL, Spier AW, Miller MW, Lehmkuhl LB, Towbin JA. Evaluation of the cardiac actin gene in Doberman Pinschers with dilated cardiomyopathy. *Am J Vet Res*. 2001;62(1):33-36.

136. Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA ME and L-TK. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. *Hum Genet*. 2012;131:1319-1325.

137. Owczarek-Lipska M, Mausberg TB, Stephenson H, Dukes-McEwan J WG and LT. A 16-bp deletion in the canine PDK4 gene is not associated with dilated cardiomyopathy in a European cohort of Doberman Pinschers. *Anim Genet*. 2013;44:239.

138. Meurs KM, Mauceli E, Lahmers S, Acland GM, White SN, Lindblad-Toh K. Genome-wide association identifies a deletion in the 3 untranslated region of Striatin in a canine model of arrhythmogenic right ventricular cardiomyopathy. *Hum Genet*. 2010;128(3):315-324. doi:10.1007/s00439-010-0855-y.

139. Meurs KM, Stern JA, Sisson DD, et al. Association of dilated cardiomyopathy with the striatin mutation genotype in boxer dogs. *J Vet Intern Med*. 2013;27(6):1437-1440. doi:10.1111/jvim.12163.

140. Palermo V, Stafford Johnson MJ, Sala E, Brambilla PG, Martin MWS. Cardiomyopathy in Boxer dogs: A retrospective study of the clinical presentation, diagnostic findings and survival. *J Vet Cardiol*. 2011;13(1):45-55. doi:10.1016/j.jvc.2010.06.005.

141. Petric AD SP and ZA. Dilated cardiomyopathy in Doberman Pinschers: Survival, Causes

of Death and a Pedigree Review in a Related Line. *J Vet Cardiol*. 2002;4:17-24.

142. Wess G, Schulze A, Butz V, et al. Prevalence of dilated cardiomyopathy in Doberman Pinschers in various age groups. *J Vet Intern Med*. 2010;24(3):533-538. doi:10.1111/j.1939-1676.2010.0479.x.

143. Calvert CA, Jacobs GJ. Heart rate variability in Doberman Pinschers with and without echocardiographic evidence of dilated cardiomyopathy. *Am J Vet Res*. 2000;61(5):506-511. doi:10.2460/javma.2001.219.1084.

144. Everett RM, McGann J, Wimberly HC, Althoff J. Dilated cardiomyopathy of Doberman pinschers: retrospective histomorphologic evaluation of heart from 32 cases. *Vet Pathol*. 1999;36(3):221-227. doi:10.1354/vp.36-3-221.

145. Beier P, Reese S, Holler PJ, Simak J, Tater G, Wess G. The role of hypothyroidism in the etiology and progression of dilated cardiomyopathy in doberman pinschers. *J Vet Intern Med*. 2015;29(1):141-149. doi:10.1111/jvim.12476.

146. Vollmar A, Fox PR, Meurs KM, Liu SK. Dilated Cardiomyopathy in Juvenile Doberman Pinschers. *J Vet Cardiol*. 2003;5(1):23-27. doi:10.1016/S1760-2734(06)70041-8.

147. Calvert C a, Pickus CW, Jacobs GJ, Brown J. Signalment, survival, and prognostic factors in Doberman pinschers with end-stage cardiomyopathy. *J Vet Intern Med*. 1997;11(6):323-326. doi:10.1111/j.1939-1676.1997.tb00474.x.

148. Calvert CA, Hall G, Jacobs G, Pickus C. Clinical and pathologic findings in Doberman pinschers with occult cardiomyopathy that died suddenly or developed congestive heart failure: 54 cases (1984-1991). *J Am Vet Med Assoc*. 1997;210(4):505-511.

149. Gilbert SJ, Wotton PR, Bailey AJ, Sims TJ, Duance VC. Alterations in the organisation, ultrastructure and biochemistry of the myocardial collagen matrix in Doberman pinschers with dilated cardiomyopathy. *Res Vet Sci*. 2000;69(3):267-274. doi:10.1053/rvsc.2000.0423.

150. Minors SL, O'Grady MR. Resting and dobutamine stress echocardiographic factors associated with the development of occult dilated cardiomyopathy in healthy Doberman pinscher dogs. *J Vet Intern Med*. 1998;12(5):369-380.

151. Phillips DE, Harkin KR. Hypothyroidism and myocardial failure in two Great Danes. *J Am Anim Hosp Assoc*. 2003;39(2):133-137.

152. Meurs KM, Miller MW, Wright N a. Clinical features of dilated cardiomyopathy in Great Danes and results of a pedigree analysis: 17 cases (1990-2000). *J Am Vet Med Assoc*. 2001;218(5):729-732. doi:10.2460/javma.2001.218.729.

153. Vollmar AC, Aupperle H. Cardiac pathology in Irish wolfhounds with heart disease. *Journal of Veterinary Cardiology*. 2016:18:57-70.

154. Vollmar AC. The prevalence of cardiomyopathy in the Irish wolfhound: a clinical study of 500 dogs. *J Am Anim Hosp Assoc*. 2000;36(2):125-132. doi:10.5326/15473317-36-2-125.

155.   Brownlie SE, Cobb MA. Observations on the development of congestive heart failure in Irish wolfhounds with dilated cardiomyopathy. *J Small Anim Pract*. 1999;40(8):371-377. doi:10.1111/j.1748-5827.1999.tb03102.x.

156.   Tidholm A, Jonsson L. Dilated cardiomyopathy in the Newfoundland: a study of 37 cases (1983-1994). *J Am Anim Hosp Assoc*. 1996;32(6):465-470. doi:10.5326/15473317-32-6-465.

157.   Tidholm A, Häggström J, Jönsson L. Detection of attenuated wavy fibers in the myocardium of Newfoundlands without clinical or echocardiographic evidence of heart disease. *Am J Vet Res*. 2000;61(3):238-241.

158.   Dambach DM, Lannon a, Sleeper MM, Buchanan J. Familial dilated cardiomyopathy of young Portuguese water dogs. *J Vet Intern Med*. 1999;13(1):65-71.

159.   Sleeper MM, Henthorn PS, Vijayasarathy C, et al. Dilated cardiomyopathy in juvenile Portuguese Water Dogs. *J Vet Intern Med*. 2002;16(1):52-62. doi:10.1111/j.1939-1676.2002.tb01606.x.

160.   Harmon MW LS. DCM Phenotype Paper.

161.   Zeng R, Coates JR, Johnson GC, et al. Breed distribution of SOD1 alleles previously associated with canine degenerative myelopathy. *J Vet Intern Med*. 2014;28(2):515-521. doi:10.1111/jvim.12317.

162.   Lukong KE, Chang K, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. *Trends Genet*. 2008;24(8):416-425. doi:10.1016/j.tig.2008.05.004.

163.   Brauch KM, Karst ML, Herron KJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol*. 2009;54(10):930-941. doi:10.1016/j.jacc.2009.05.038.

164.   Li D, Morales A, Gonzalez-Quintana J, et al. Identification of novel mutations in RBM20 in patients with dilated cardiomyopathy. *Clin Transl Sci*. 2010;3(3):90-97. doi:10.1111/j.1752-8062.2010.00198.x.

165.   Refaat MM, Lubitz SA, Makino S, et al. Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. *Hear Rhythm*. 2012;9(3):390-396. doi:10.1016/j.hrthm.2011.10.016.

166.   Guo W, Schafer S, Greaser ML, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med*. 2012;18(5):766-773. doi:10.1038/nm.2693.

167.   Greaser ML, Warren CM, Esbona K, et al. Mutation that dramatically alters rat titin isoform expression and cardiomyocyte passive tension. *J Mol Cell Cardiol*. 2008;44(6):983-991. doi:10.1016/j.yjmcc.2008.02.272.

168.   Greaser ML, Krzesinski PR, Warren CM, Kirkpatrick B CK and MR. Developmental changes in rat cardiac titin/connectin: transitions in normal animals and in mutants with a delayed pattern of isoform transition. *J Muscle Res Cell Motil*. 2005;26:325-332.

169. Furst DO, Osborn M NR and WK. The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line. *J Cell Biol*. 1988;106:1563-1572.

170. Linke WA, Rudy DE, Centner T, et al. I-band titin in cardiac muscle is a three-element molecular spring and is critical for maintaining thin filament structure. *J Cell Biol*. 1999;146(3):631-644. doi:10.1083/jcb.146.3.631.

171. Anderson BR, Granzier HL. Titin-based tension in the cardiac sarcomere: Molecular origin and physiological adaptations. *Prog Biophys Mol Biol*. 2012;110(2-3):204-217. doi:10.1016/j.pbiomolbio.2012.08.003.

172. Bang ML, Centner T, Fornoff F, et al. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ Res*. 2001;89(11):1065-1072. doi:10.1161/hh2301.100981.

173. Gautel M LE and PF. Assembly of the cardiac I-band region of titin/connectin: expression of the cardiac-specific regions and their structural relation to the elastic segments. *J Muscle Res Cell Motil*. 1996;17:449-461.

174. Neagoe C, Opitz CA, Makarenko I, Linke WA. Gigantic variety: Expression patterns of titin isoforms in striated muscles and consequences for myofibrillar passive stiffness. In: *Journal of Muscle Research and Cell Motility*. Vol 24. ; 2003:175-189. doi:10.1023/A:1026053530766.

175. Freiburg a, Trombitas K, Hell W, et al. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circ Res*. 2000;86(11):1114-1121. doi:10.1161/01.RES.86.11.1114.

176. Lahmers S, Wu Y, Call DR, Labeit S, Granzier H. Developmental Control of Titin Isoform Expression and Passive Stiffness in Fetal and Neonatal Myocardium. *Circ Res*. 2004;94(4):505-513. doi:10.1161/01.RES.0000115522.52554.86.

177. Cazorla O, Freiburg A, Helmes M, et al. Differential expression of cardiac titin isoforms and modulation of cellular stiffness. *Circ Res*. 2000;86(1):59-67. doi:10.1161/01.RES.86.1.59.

178. Wu Y, Bell SP, Trombitas K, et al. Changes in titin isoform expression in pacing-induced cardiac failure give rise to increased passive muscle stiffness. *Circulation*. 2002;106(11):1384-1389. doi:10.1161/01.CIR.0000029804.61510.02.

179. Roberts AM, Ware JS, Herman DS, Schafer S, Baksi J, Bick AG, Buchan RJ, Walsh R, John S, Wilkinson S, Mazzarotto F, Felkin LE, Gong S, MacArthur JA, Cunningham F, Flannick J, Gabriel SB, Altshuler DM, Macdonald PS, Heinig M, Keogh AM, Hayward CS, Banner N SC and CS. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci Transl Med*. 2015;7:270ra6.

180.	Warren CM, Krzesinski PR, Campbell KS, Moss RL, Greaser ML. Titin isoform changes in rat myocardium during development. *Mech Dev*. 2004;121(11):1301-1312. doi:10.1016/j.mod.2004.07.003.

181.	Opitz CA, Leake MC, Makarenko I, Benes V, Linke WA. Developmentally Regulated Switching of Titin Size Alters Myofibrillar Stiffness in the Perinatal Heart. *Circ Res*. 2004;94(7):967-975. doi:10.1161/01.RES.0000124301.48193.E1.

182.	Walker JS, De Tombe PP. Titin and the Developing Heart. *Circ Res*. 2004;94(7):860-862. doi:10.1161/01.RES.0000126698.37440.B0.

183.	Li D, Morales A, Gonzalez-Quintana J, et al. Identification of novel mutations in RBM20 in patients with dilated cardiomyopathy. *Clin Transl Sci*. 2010;3(3):90-97. doi:10.1111/j.1752-8062.2010.00198.x.

184.	Gilliam D, O'Brien DP, Coates JR, et al. A homozygous KCNJ10 mutation in jack russell terriers and related breeds with spinocerebellar ataxia with myokymia, seizures, or both. *J Vet Intern Med*. 2014;28(3):871-877. doi:10.1111/jvim.12355.

185.	Rampersaud E, Siegfried JD, Norton N, Li D, Martin E, Hershberger RE. Rare variant mutations identified in pediatric patients with dilated cardiomyopathy. *Prog Pediatr Cardiol*. 2011;31(1):39-47. doi:10.1016/j.ppedcard.2010.11.008.

186.	Millat G, Bouvagnet P, Chevalier P, et al. Clinical and mutational spectrum in a cohort of 105 unrelated patients with dilated cardiomyopathy. *Eur J Med Genet*. 2011;54(6). doi:10.1016/j.ejmg.2011.07.005.

187.	Dec GW, Fuster V. Idiopathic dilated cardiomyopathy. *N Engl J Med*. 1994;331:1564-1575.

188.	Steele-Stallard HB, Le Quesne Stabej P, Lenassi E, Luxon LM, Claustres M, Roux AF, Webster AR B-GM. Screening for duplications, deletions and a common intronic mutation detects 35% of second mutations in patients with USH2A monoallelic mutations on Sanger sequencing. *Orphanet J Rare Dis*. 2013;8:122.

189.	Lobo I. Copy Number Variation and Genetic Disease. *Nat Educ*. 2008;1(1):65.

190.	Wang W, Wang W, Sun W, Crowley JJ, Szatkiewicz JP. Allele-specific copy-number discovery from whole-genome and whole-exome sequencing. *Nucleic Acids Res*. 2015;43(14):e90. doi:10.1093/nar/gkv319.

191.	Chen X, Gupta P, Wang J, et al. CONSERTING: integrating copy-number analysis with structural-variation detection. *Nat Methods*. 2015;12(6):527-530. doi:10.1038/nmeth.3394.

192.	Caudy AA, Myers M, Hannon GJ HS. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev*. 2002;16(19):2491-2496.

193. Penso-Dolfin L, Swofford R, Johnson J, Alföldi J, Lindblad-Toh K, Swarbreck D, Moxon S DPF. An Improved microRNA Annotation of the Canine Genome. *PLoS One*. 2016;11(4):e0153453.

VITA

Douglas H Gilliam, Jr. was born in 1983 in Houston, Texas. He obtained his Bachelor's degree in Biology in May of 2011 at Missouri State University in Springfield, Missouri. In 2011, he started his PhD studies at the University of Missouri-Columbia. He completed his PhD in Genetics in May of 2016