

8-2017

Identification of Genetic Variants Underlying Simple and Complex Canine Myopathies

Jacquelyn Marie Evans

Clemson University, jmevans1230@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

Evans, Jacquelyn Marie, "Identification of Genetic Variants Underlying Simple and Complex Canine Myopathies" (2017). *All Dissertations*. 2007.

https://tigerprints.clemson.edu/all_dissertations/2007

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

IDENTIFICATION OF GENETIC VARIANTS UNDERLYING SIMPLE AND
COMPLEX CANINE MYOPATHIES

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Genetics

by
Jacquelyn Marie Evans
August 2017

Accepted by:
Dr. Leigh Anne Clark, Committee Chair
Dr. Miriam Konkel
Dr. Stephen Kresovich
Dr. Margaret Ptacek
Dr. Rajandeep Sekhon

ABSTRACT

The unique population structure of dog breeds, arising from their domestication from gray wolves and subsequent breed formation, and the similarity of their inherited diseases to human disorders make them an ideal comparative genetics model. Herein the genetic basis of three canine muscle diseases, each a model for human disease, is investigated using genome-wide approaches. Nemaline rod myopathy (NM) is one of the most common congenital myopathies in people and is characterized by rod bodies in the muscle fibers, muscle weakness, and reduced muscle tone. We characterized the first large animal model of autosomal recessive NM in a family of American bulldogs and, through a combination of genome-wide SNP profiling and whole exome sequencing, identified a nonsense mutation in *NEB*.

Limb girdle muscular dystrophy (LGMD) affects the hip and shoulder muscles and may cause respiratory and cardiac muscle degeneration. We determined that a muscular dystrophy in a family of Boston terriers is a sarcoglycanopathy, a type of LGMD caused by mutations in one of six sarcoglycan genes, and identified a 2 bp deletion in *SGCD* through direct whole exome sequencing. Further, an unrelated Boston terrier having LGMD harbors a 19.4 kb deletion, omitting exons 7 and 8 of *SGCD*. The *NEB* and *SGCD* mutations are present only in the affected families.

Juvenile dermatomyositis (JDM) is an autoimmune disease with a complex mode of inheritance and an environmental trigger, affecting children ages 2 to 17. In dogs, dermatomyositis (DMS) affects collies and Shetland sheepdogs and causes skin lesions and muscle weakness. We conducted genome-wide association studies in both breeds and

identified linkage disequilibrium with SNPs on chromosomes 10 and 31. Through whole genome and transcriptome sequencing, we found mutations in two novel genes, *PAN2* (chr10) and *MAP3K7CL* (chr31). Additionally, we identified an association with a haplotype of the major histocompatibility complex class II genes on chromosome 12, alleles of which are also associated with JDM. When variants at all three loci are considered together, genotypes confer low, moderate, or high risk for DMS, with moderate- and high-risk genotypes explaining 93% of cases.

DEDICATION

For Mom, Rachel, and Adam.

ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Leigh Anne Clark, for bringing me into the lab as an undergraduate student, giving me an independent project, and encouraging me to pursue a PhD. I was very fortunate to join the lab as a graduate student when she procured NIH funding for the genetic study of dermatomyositis and even more so to travel with her to present that work in the U.S. and internationally. I am grateful for her mentorship and friendship and all of her help in achieving my goals. I thank my fellow Canine Genetics Laboratory members Drs. Kate Tsai and Alison Starr-Moss for their advice, encouragement, and friendship. Thank you to my committee members, Drs. Miriam Konkol, Stephen Kresovich, Margaret Ptacek, and Rajandeep Sekhon, for their guidance, support, and enthusiasm for my work.

I have several collaborators to thank. The American bulldog and Boston terrier projects would not have been possible without the contributions and support of Drs. Diane Shelton and Melissa Cox. I thank Dr. Tom Famula for being the Clemson group's favorite statistician; his contribution to the dermatomyositis research is invaluable. I am grateful to Dr. Rooksana Noorai for her work on dermatomyositis in collies; I was very lucky to be able to carry on that project. I thank the undergraduate and graduate students who contributed to this work, especially Kendall Anderson and Cody Hill. I also thank the breeders, owners, and veterinarians who provided samples, with special thanks to Dr. Nancy Kelso and the Collie Health Foundation, as well as Dr. Mary Mahaffey and the American Shetland Sheepdog Association.

Finally, thank you to my family for their lifelong support and their encouragement

throughout my academic career, and to Stephen for making life better every day.

TABLE OF CONTENTS

	Page
TITLE PAGE	i
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	
I. INTRODUCTION	1
Dog domestication	1
Dogs as a model	2
Identification of disease variants in dogs	4
Canine myopathies	8
References	12
II. EXOME SEQUENCING REVEALS A NEBULIN NONSENSE MUTATION IN A DOG MODEL OF NEMALINE ROD MYOPATHY	28
Abstract	29
Introduction	30
Materials and methods	31
Results	36
Discussion	42
Acknowledgments	44
References	45
III. EXOME SEQUENCING REVEALS INDEPENDENT SGCD DELETIONS CAUSING LIMB GIRDLE MUSCULAR DYSTROPHY IN BOSTON TERRIERS	51
Abstract	52

Table of Contents (Continued)

	Page
Background.....	53
Materials and methods	55
Results.....	60
Discussion	67
Conclusion	69
Abbreviations.....	69
Ethics approval.....	70
Availability of data	70
Competing interests	70
Funding	70
Author contributions	71
Acknowledgments.....	71
References.....	71
IV. BEYOND THE MHC: A CANINE MODEL OF DERMATOMYOSITIS SHOWS A COMPLEX PATTERN OF GENETIC RISK INVOLVING NOVEL LOCI	78
Abstract.....	79
Author summary	80
Introduction.....	81
Results and discussion	83
Methods.....	99
Accession numbers	107
Acknowledgments.....	107
Author contributions	107
References.....	108
V. CONCLUSION.....	119
Nemaline rod myopathy in American bulldogs	119
Limb girdle muscular dystrophy in Boston terriers	120
Dermatomyositis in collies and Shetland sheepdogs	120
APPENDICES	123
A: Association of DLA-DQB1 alleles with exocrine pancreatic insufficiency in Pembroke Welsh corgis	124
B: Supplemental information for Chapter 3	134
C: Supplemental information for Chapter 4	135

Table of Contents (Continued)

	Page
D: Permission to reprint published work	162

LIST OF TABLES

Table		Page
2.1	Genotypes for the nuclear family are reported for polymorphic SNPs representing each of the 10 candidate genes	39
4.1	Frequency of DLA-DRB1/-DQA1/-DQB1 haplotypes in collies and Shetland sheepdogs.....	84
4.2	Expressed exonic changes on chromosomes 10 and 31.....	89
4.3	Distribution of three-locus genotypes in 132 cases and 390 controls with penetrance, 95% confidence intervals, and <i>P</i> -values	94

LIST OF FIGURES

Figure		Page
2.1	Histopathology establishes NM	38
2.2	Filtering of NEB variants reveals S8042X	41
2.3	Nebulin is drastically reduced in skeletal muscles with S8042X	41
3.1	Histopathology of muscle biopsies from a female Boston terrier affected with sarcoglycanopathy (Case 3)	61
3.2	Loss of SGC staining in Cases 3 and 4	62
3.3	Representative immunofluorescence of cryosections from muscle of Cases 3 and 4, and control dog muscle	63
3.4	Electropherogram showing the 2 bp <i>SGCD</i> deletion in Case 3.....	65
3.5	Schematic and sequence showing the breakpoints of the 19,403 bp <i>SGCD</i> deletion in Case 1	66
4.1	Manhattan plots of DMS GWASs in two breeds.....	86
4.2	Regional plots depicting LD at two associated loci.....	87
4.3	Three-locus genotype disease probabilities and frequencies (132 cases, 390 controls)	93
4.4	Frequencies of DMS-associated polymorphisms across breeds	96

CHAPTER ONE

INTRODUCTION

Dog domestication

The earliest domesticated species, dogs (*Canis lupus familiaris*) descended from one or more populations of gray wolves 11,000 to 35,000 years ago alongside hunter-gatherer societies [1]. Putative domestication origins include Europe, East Asia, and the High Arctic. Changes in brain function, behavior, pigmentation, and diet and metabolism have been implicated in early dog domestication [2-6].

While previous studies have compared modern dogs to gray wolves to identify domestication genes, a recent study including village dogs, gray wolves, and three ancient dogs identified selective sweeps containing genes with roles in neural crest development and the minor spliceosome [7]. These genes may explain traits associated with the “domestication syndrome” that is characteristic of mammalian domesticates and includes floppy ears; depigmentation; and decreased body, skull, and jaw size [7,8]. Methylation patterns also appear to have been under selection during early domestication. Janowitz Koch et al. [9] identified hypermethylation of transposons in dogs compared to wolves, and suggest that transposons underwent expansion in dogs followed by epigenetic silencing. Furthermore, Banlaki et al. [10] found evidence for differential methylation between dogs and wolves specifically in the promoter regions of genes related to behavior and cognitive function. Taken together, these studies suggest that domestication phenotypes may be attributed to far-reaching, regulatory mechanisms rather than wide-spread coding variants.

After the first stage of domestication from wolves, with the emergence of agriculture and a more sedentary human lifestyle, dogs began to appear in a range of body sizes and became morphologically distinct from wolves [1,11]. Archaeological evidence suggests that dogs were bred to pull sleds around 9000 ya in Siberia, possibly as early as 15,000 ya; this is the earliest evidence of selective breeding for a job [12]. More recently, dogs were bred for retrieving, hunting, defense, speed, and companionship; modern breed formation during the Victorian era included artificial selection for novelty or “fancy” [13,14]. This final stage of domestication yielded the majority of dog diversity [1,15], and today, there is more phenotype variability among dogs than in the entire carnivore order [1].

Dogs as a model

There are over 400 breeds of dog [16], and they are afflicted with approximately 360 simple and complex naturally-occurring diseases analogous to human disease [17]. Dogs are superior disease models to rodents in a number of ways. As pets, dogs receive excellent medical care and typically have extensive health records and even family histories. Biological samples can be obtained from the pet population for identifying disease alleles, eliminating expenses associated with maintaining laboratory animals. As companion animals, dogs share their owners’ environment and thus are more likely than rodent models to be exposed to toxins and other environmental agents playing a role in complex disease [18].

Dogs have organ sizes similar to humans and may even be more similar genetically [17], which makes them particularly advantageous for gene therapy and drug trials. While rodent models are useful for developing gene therapies, scaling up the injections to human treatment poses a problem because host body weight, metabolic rate, and response are all factors that may not translate from a rodent model to humans [19]. Further, negligible contamination levels at a small-scale could be amplified and cause an immune response when the scale is increased for human trials [20]. Dog models are an ideal intermediate system to address these issues because of their more comparable size to humans [19], and dog models for hemophilia and inherited bleeding disorders have been successfully used to determine safe doses and prevent immune responses in trials that have translated to human treatment [21].

Dogs underwent two population bottlenecks: domestication from wolves and breed formation. The latter resulted in founder effects where a small number of individuals formed the basis for a breed, popular sire effects where studs who excel in competitions are selected to sire a disproportionately high number of puppies, and within breed population bottlenecks due to changes in breed standards or events like the World Wars, which led to near extinction of some breeds. These factors caused high genetic homogeneity and long stretches of linkage disequilibrium (LD) within breeds, on average extending 1 Mb, 100-fold greater than LD in humans [22-24]. Thus, fewer markers and fewer individuals are necessary to detect disease loci in dogs, tens of thousands of markers compared to millions needed in human studies [22,23]. However, the long LD can result in broad associated candidate intervals containing many genes for investigating

mutations. Because across-breed LD is comparable to what is observed in humans (~10 kb) [22,23], fine-mapping studies may be successful using multiple breeds with the same phenotype to narrow these critical intervals.

A consequence of selective breeding practices is that deleterious alleles become overrepresented in certain breeds so that about half of heritable diseases are breed-specific, making breeds similar to geographically isolated human populations [25]. These alleles are passed on identical by descent within breeds and, thus, approximately two-thirds of hereditary diseases in dogs are autosomal recessive [26]. For complex diseases in dogs, reduced locus heterogeneity within breeds is expected compared to human populations. It is likely that in dogs, a small number of genes with large effect drive complex diseases that in humans are governed by many loci of small effect [16]. This facilitates identification of associated genes; dogs may even have a recessive form of a human complex disease. Safra et al. [27] identified a novel gene in Weimereiners with an autosomal recessive neural tube defects, a disorder with complex inheritance patterns in humans. The gene was then found to be involved in human cases [27]. Complex disorders may also be studied across affected breeds to find multiple genes involved. For example, epilepsy manifests in different ways and certain breeds are affected with different types [16].

Identification of disease variants in dogs

The canine genome comprises 2.8 billion base pairs packaged into 38 acrocentric autosomes. In 2004, the dog reference genome was sequenced to 7.8x coverage in a

female Boxer, selected because she had low heterozygosity and provided full coverage of the X chromosome. A SNP map based on sequences from nine additional breeds and the 1.5x genome of a standard poodle was also made publicly available [23,28]. Since the release of the reference, single nucleotide polymorphism (SNP) arrays have been developed for genome-wide association studies (GWAS) [29]. GWAS generate genotypes for SNPs distributed throughout the genome and compare allele frequencies between unrelated affected and healthy individuals. SNP markers proximal to the causative mutation will be in LD, yielding an associated region. Development of the Affymetrix 50K platinum panel and Illumina 22K SNP arrays enabled successful GWAS [30-34], including identification of a novel ichthyosis gene in Golden retrievers, which led investigators to mutations in the same gene in human cases [35].

Today an Illumina CanineHD Bead Chip is available containing over 173,000 SNPs, which within breeds yields between 80,000 and 120,000 informative markers, sufficient coverage to identify loci associated with simple Mendelian disorders and some complex diseases [29]. With this array, simple disorders can be mapped with very small numbers of cases. For example, the Safra et al. [27] study of an autosomal recessive neural tube defect in Weimereiners used only 4 affected dogs in the GWAS.

After identifying an associated region through GWAS, the best candidate gene(s) within the critical interval is typically selected for Sanger sequencing to identify the causative mutation [36-42]. This approach can be costly and time intensive, particularly if the candidate genes are too numerous or too large. Recently, next generation sequencing (NGS), or high through-put sequencing, technologies have been developed that perform

parallel sequencing of millions of DNA fragments; the sequences are then aligned to a reference genome. NGS may be used to sequence the entire (whole genome resequencing – WGS) or a portion of the genome (targeted-resequencing), only the exons and sometimes introns and regulatory regions (whole exome sequencing – WES), or the transcriptome (RNAsequencing) of multiple individuals concurrently. Targeted and WGS technologies were first used in dog studies in 2011-2012, followed by the less-expensive WES in 2013, and the cost of NGS continues to decline [29]. NGS technologies now allow for rapid variant detection and screening following GWAS [43-55].

Recently, WGS has been utilized to identify mutations without prior mapping. When there are several known candidate genes for a recessive disease, WGS of a single case has been effective for simultaneously investigating all candidates and identifying a deleterious variant [56-58]. WGS has also been employed in family trio studies to filter variants for recessive mutations in the affected individual and obligate carrier parents without positional or functional candidate genes [59]. Other studies have identified homozygous mutations underlying recessive disease by scanning the entire genome of a single case and ruling out all variants present in the genomes of dogs of other breeds until a unique, deleterious variant is identified [60-61]. Bauer et al. [62] even identified a dominant *de novo* mutation this way for a German shepherd with ichthyosis.

Whole exome sequencing (WES) kits provide a more cost-effective alternative to WGS in these simple Mendelian disorders that are likely to be caused by coding mutations. A handful of studies have already used this technique alone or with mapping to identify mutations in progressive retinal atrophy [63], glaucoma [64], and neuroaxonal

dystrophy [65]. Genotyping by WES was also used to conduct a GWAS for glaucoma in Petit Basset Griffon Vendéens [66]. Variant calling and filtering for minor allele frequency less than 5% in 12 cases and 12 controls yielded 61,977 SNPs, and allele frequency analysis revealed an association on chromosome 3. Subsequent WGS of a single case detected an inversion in *ADAMTS17*.

In contrast to simple recessive or dominant diseases, complex polygenic traits in dogs may require denser SNP arrays than the Illumina 170k and more individuals to detect associations in GWAS. Hayward et al. [67] conducted GWAS with 4,200 dogs and 180,000 SNPs to identify across-breed associations with multiple complex diseases, including hip and elbow dysplasia and morphological traits like fur length and shedding. The authors suggest that increased SNP density will still be necessary for complex diseases and were unable to detect significant associations for several other disorders they investigated like lymphoma, mast cell tumor, and mitral valve degeneration.

Associations with multigenic disease have been identified using fewer than the 100 cases and 100 controls initially proposed by Lindblad-Toh et al. [23]. Karyadi et al. [68] mapped squamous cell carcinoma of the digit (SCCD) using 31 case and 34 control standard poodles to a region on chromosome 15. They narrowed the critical interval, where all affected dogs possessed at least one copy of the risk haplotype, from 813 kb to 145 kb by investigating these SNPs in cases from two other breeds also affected with SCCD, Briards and Giant Schnauzers, and ultimately found a strongly associated copy number variant in *KITLG* [68]. Across-breed fine mapping approaches can be highly effective in the case of complex disease because the LD is much narrower compared to

within breeds. SCCD disproportionately affects dogs with dark coats, so after the initial GWAS with only dark-coated standard poodles, the authors performed a second GWAS comparing 24 control dark-coated to 24 light-coated control poodles. Because no significant associations outside of *MCRI*, which controls coat color and is a known skin cancer susceptibility locus, were detected, the authors propose that a compensatory mutation in this gene in light-coated dogs likely confers protection from SCCD.

Multibreed GWAS, potentially with denser SNP arrays, followed by whole genome and/or transcriptome sequencing will likely be necessary to detect variants underlying complex disease in dogs, especially since many will involve regulatory mutations rather than coding changes.

Canine myopathies

Myopathy refers to any disease affecting the skeletal muscle and causing muscle weakness, without disorders of innervation or the neuromuscular junction. Two examples of canine myopathies and their utility as genetic models are Duchenne muscular dystrophy (DMD) and X-linked myotubular myopathy (XLMTM). DMD is characterized by the absence of dystrophin, a protein essential for muscle cell stability. Mouse and canine DMD (cDMD) models have been available for study for the last 30 years, but cDMD is more clinically and histologically similar to DMD and can replicate human immune response to gene therapy [19]. XLMTM was identified in a family of Labrador retrievers [69] found to harbor a mutation in *MTM1* [70], the same gene that causes XLMTM in humans [71]. Affected dogs closely recapitulate the human symptoms [70],

and gene therapy first developed in a mouse model was successfully refined in dogs to correct muscle weakness [72]. In addition to muscular dystrophies and XLMTM, cases of nemaline rod myopathy, central core-like myopathy, myofibrillar myopathy with desmin storage, and centronuclear myopathy have been described in dogs [73]. Herein the genetic basis of three canine myopathies is described.

Nemaline rod myopathy

Nemaline rod myopathy (NM) accounts for 17% of congenital myopathies in humans and causes reduced muscle strength and tone, involving the neck, face, and respiratory muscles [74]. NM is characterized by rod bodies in the skeletal muscle fiber and is most often caused by mutations in 10 genes encoding proteins of the sarcomere thin filament, which is essential for muscle contraction. The majority of cases are attributed to dominant mutations in *ACTA1* or recessive mutations in *NEB* [74]. *ACTA1* encodes the primary component of the thin filament, actin, and *NEB* produces the giant protein nebulin, which stabilizes actin and allows the thin filament to reach the appropriate length [75]. The number of rod bodies in the muscle fiber is not correlated with disease severity. Of the 10 genes known to cause NM [76], mouse and/or zebrafish models exist for 5, including *ACTA1* and *NEB* [77-88]; however, no large animal models have been described. NM has been reported in Border collie and a Schipperke, but the genetic basis is unknown [73]. Herein, the first molecularly characterized NM in dogs is described.

Limb girdle muscular dystrophy

Limb girdle muscular dystrophies (LGMD) are characterized by progressive deterioration of the shoulder and hip muscles and may involve the heart and respiratory muscles. Age of onset and severity are variable, and 31 genes have been implicated [89]. Autosomal recessive forms are most common with an incidence of 1 in 15,000 births [90]. Four of the 23 recessive subtypes of LGMD are caused by mutations in any of the four genes encoding subunits of the sarcoglycan complex and are termed “sarcoglycanopathies” [89]. Sarcoglycans function in stabilizing the cell membrane of muscle fibers. Mouse models exist for each form [91], and a spontaneous hamster model is available for δ -sarcoglycanopathy [92,93]. Gene therapy trials are already underway in humans for α -sarcoglycanopathy [94,95].

Sarcoglycanopathy has been reported in a Boston terrier, Cocker Spaniel, and Chihuahua [73], but the cause is unknown in each case. Herein, the genetic basis of a sarcoglycanopathy in two Boston terriers is described; this is the first molecularly characterized large animal model of this disease.

Dermatomyositis

Juvenile dermatomyositis (JDM) is the most common inherited childhood inflammatory myopathy [96]. JDM causes a rash on the bony prominences (e.g. across the cheekbones, elbows, knees, back, and chest) [96]. Muscle weakness is common, and affected children describe their muscles as achy and tender and may have difficulty rising from a seated position or climbing stairs [97]. Like most autoimmune diseases in humans, females are more often affected than males [98]. JDM is strongly associated with a

haplotype of the human leukocyte antigen genes of the major histocompatibility complex; however, other genetic factors of major effect are believed to be involved [99,100].

Canine dermatomyositis (DMS), a model for JDM, is an autoimmune vasculopathy that causes alopecia and skin lesions on the bony prominences (e.g. the face, ears, tail tip, and across the tops of the feet) [101-103]. Severe cases also involve the muscles, which may result in an abnormal high-stepping gate, megaesophagus, and/or aspiration pneumonia [101,103]. Decreased muscle tone may be evident in the head musculature [101]. Clinical signs typically appear subsequent to an environmental trigger such as viral infection or vaccination [104]. Lesions may wax and wane throughout a dog's lifetime and can be exacerbated by estrus or exposure to UV light [101,103]. Age of onset is variable with most dogs developing clinical signs between seven weeks and six months of age, but others developing signs in late adulthood [103,105]. This makes the disease very difficult to eliminate from the two breeds almost exclusively diagnosed with DMS: the collie and Shetland sheepdog.

The goal of this work is to identify mutations causing two simple recessive diseases: NM in American bulldogs and LGMD in Boston terriers, and one complex disease, DMS in collies and Shetland sheepdogs. In the case of NM, SNP arrays were used to identify regions of the genome inherited in an autosomal recessive manner in a nuclear family. Cross-referencing with known candidate genes resulted in two genes of interest, and subsequent WES identified a nonsense mutation in *NEB*. Two mutations underlying LGMD in independent Boston terrier families were identified in a known

candidate gene via WES of the affected dogs and relatives. Finally, multi-breed GWAS, WGS, and transcriptome sequencing were all employed to identify three polymorphisms that in combination confer low, moderate, or high risk for DMS.

References

1. Freedman AH, Wayne RK. Deciphering the origin of dogs: from fossils to genomes. *Annu Rev Anim Biosci.* 2017;5:281-307.
2. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, et al. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature.* 2013;495:360-364.
3. Wang GD, Zhai W, Yang HC, Fan RX, Cao X, Zhong L, et al. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat Commun.* 2013;4:1860.
4. Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature.* 2010;464:898-902.
5. Trut L, Oskina I, Kharlamova A. Animal evolution during domestication: the domesticated fox as a model. *Bioessays.* 2009;31:349-60.
6. Freedman AH, Schweizer RM, Ortega-Del Vecchyo D, Han E, Davis BW, Gronau I, et al. Demographically-Based Evaluation of Genomic Regions under Selection in Domestic Dogs. *PLoS Genet.* 2016;12:e1005851.

7. Pendleton AL, Shen F, Taravella AM, Emery S, Veeramah KR, Boyko AR, et al. Selective sweep analysis using village dogs highlights the pivotal role of the neural crest in dog domestication. bioRxiv. 2017;118794;doi:<https://doi.org/10.1101/118794>.
8. Wilkins AS, Wrangham RW, Fitch WT. The "domestication syndrome" in mammals: a unified explanation based on neural crest cell behavior and genetics. *Genetics*. 2014;197:795-808.
9. Janowitz Koch I, Clark MM, Thompson MJ, Deere-Machemer KA, Wang J, Duarte L, et al. The concerted impact of domestication and transposon insertions on methylation patterns between dogs and grey wolves. *Mol Ecol*. 2016;25:1838–1855.
10. Banlaki Z, Cimorelli G, Viranyi Z, Kubinyi E, Sasvari-Szekely M, Ronai Z. DNA methylation patterns of behavior-related gene promoter regions dissect the gray wolf from domestic dog breeds. *Mol Genet Genomics*. 2017;292:685-697.
11. Larson G, Karlsson EK, Perri A, Webster MT, Ho SY, Peters J, et al. Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA*. 2012;109:8878-8883.
12. Pitulko VV, Kasparov AK. Archaeological dogs from the early Holocene Zhokhov site in the eastern Siberian arctic. *J Archaeol Sci*. 2017;13:491-515.
13. Wilcox B, Walkowicz C. *The Atlas of Dog Breeds of the World*. 5th ed. TFH Publications; 1995.

14. American Kennel Club. *The Complete Dog Book*. 20th ed. Ballantine Books; 2006.
15. Shearin AL, Ostrander EA. Canine morphology: hunting for genes and tracking mutations. *PLoS Biol*. 2010;8:e1000310.
16. Ostrander EA, Franklin H. Epstein Lecture. Both ends of the leash--the human links to good dogs with bad genes. *N Engl J Med*. 2012;367:636-646.
17. Shearin AL, Ostrander EA. Leading the way: canine models of genomics and disease. *Dis Model Mech*. 2010;3:27-34.
18. Tsai KL, Clark LA, Murphy KE. Understanding hereditary diseases using the dog and human as companion model systems. *Mamm Genome*. 2007;18:444-451.
19. Duan D. Duchenne muscular dystrophy gene therapy in the canine model. *Hum Gene Ther Clin Dev*. 2015;26:57-69.
20. Wright JF. Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Ther*. 2008;15:840-848.
21. Nichols TC, Hough C, Agersø H, Ezban M, Lillicrap D. Canine models of inherited bleeding disorders in the development of coagulation assays, novel protein replacement and gene therapies. *J Thromb Haemost*. 2016;14:894-905.
22. Sutter NB, Eberle MA, Parker HG, Pullar BJ, Kirkness EF, Kruglyak L, et al. Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res*. 2004;14:2388-2396.

23. 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;438:803-819.
24. Parker HG, Shearin AL, Ostrander EA. Man's best friend becomes biology's best in show: genome analyses in the domestic dog. *Annu Rev Genet*. 2010;44:309-336.
25. Sutter NB, Ostrander EA. Dog star rising: the canine genetic system. *Nat Rev Genet*. 2004;5:900-910.
26. Clark LA, Starr-Moss AN. Genetics and genomics of the domestic dog. In: Khatib H, editor. *Molecular and Quantitative Animal Genetics*. Hoboken, New Jersey. Wiley Blackwell; 2015. pp. 121-130.
27. Safra N, Bassuk AG, Ferguson PJ, Aguilar M, Coulson RL, Thomas N, et al. Genome-wide association mapping in dogs enables identification of the homeobox gene, NKX2-8, as a genetic component of neural tube defects in humans. *PLoS Genet*. 2013;9:e1003646.
28. Kirkness EF, Bafna V, Halpern AL, Levy S, Remington K, Rusch DB, et al. The dog genome: survey sequencing and comparative analysis. *Science*. 2003;301:1898-1903.
29. van Steenbeek FG, Hytönen MK, Leegwater PA, Lohi H. The canine era: the rise of a biomedical model. *Anim Genet*. 2016;47:519-527.

30. Awano T, Johnson GS, Wade CM, Katz ML, Johnson GC, Taylor JF, et al.
Genome-wide association analysis reveals a SOD1 mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. *Proc Natl Acad Sci USA*. 2009;106:2794-2799.
31. 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;128:315-324.
32. Seppälä EH, Koskinen LL, Gulløv CH, Jokinen P, Karlskov-Mortensen P, Bergamasco L, et al. Identification of a novel idiopathic epilepsy locus in Belgian Shepherd dogs. *PLoS One*. 2012;7:e33549.
33. Hytönen MK, Arumilli M, Lappalainen AK, Kallio H, Snellman M, Sainio K, et al. A novel GUSB mutation in Brazilian terriers with severe skeletal abnormalities defines the disease as mucopolysaccharidosis VII. *PLoS One*. 2012;7:e40281.
34. Ahonen SJ, Arumilli M, Seppälä E, Hakosalo O, Kaukonen MK, Komáromy AM, et al. Increased expression of MERTK is associated with a unique form of canine retinopathy. *PLoS One*. 2014;9:e114552.
35. Grall A, Guaguère E, Planchais S, Grond S, Bourrat E, Hausser I, et al. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. *Nat Genet*. 2012;44:140-147.

36. Gill JL, Tsai KL, Krey C, Noorai RE, Vanbellinghen JF, Garosi LS, et al. A canine BCAN microdeletion associated with episodic falling syndrome. *Neurobiol Dis.* 2012;45:130-136.
37. Kyöstilä K, Cizinauskas S, Seppälä EH, Suhonen E, Jeserevics J, Sukura A, et al. A SEL1L mutation links a canine progressive early-onset cerebellar ataxia to the endoplasmic reticulum-associated protein degradation (ERAD) machinery. *PLoS Genet.* 2012;8:e1002759.
38. Ahonen SJ, Kaukonen M, Nussdorfer FD, Harman CD, Komáromy AM, Lohi H. A novel missense mutation in ADAMTS10 in Norwegian Elkhound primary glaucoma. *PLoS One.* 2014;9:e111941.
39. Ekenstedt KJ, Becker D, Minor KM, Shelton GD, Patterson EE, Bley T, et al. An ARHGEF10 deletion is highly associated with a juvenile-onset inherited polyneuropathy in Leonberger and Saint Bernard dogs. *PLoS Genet.* 2014;10:e1004635.
40. Pemberton TJ, Choi S, Mayer JA, Li FY, Gokey N, Svaren J, et al. A mutation in the canine gene encoding folliculin-interacting protein 2 (FNIP2) associated with a unique disruption in spinal cord myelination. *Glia.* 2014;62:39-51.
41. Wolf ZT, Leslie EJ, Arzi B, Jayashankar K, Karmi N, Jia Z, et al. A LINE-1 insertion in DLX6 is responsible for cleft palate and mandibular abnormalities in a canine model of Pierre Robin sequence. *PLoS Genet.* 2014;10:e1004257.

42. Wiik AC, Ropstad EO, Ekesten B, Karlstam L, Wade CM, Lingaas F. Progressive retinal atrophy in Shetland sheepdog is associated with a mutation in the CNGA1 gene. *Anim Genet.* 2015;46:515-521.
43. Kuchtey J, Olson LM, Rinkoski T, Mackay EO, Iverson TM, Gelatt KN, et al. Mapping of the disease locus and identification of ADAMTS10 as a candidate gene in a canine model of primary open angle glaucoma. *PLoS Genet.* 2011;7:e1001306.
44. Olsson M, Meadows JR, Truvé K, Rosengren Pielberg G, Puppo F, Mauceli E, et al. A novel unstable duplication upstream of HAS2 predisposes to a breed-defining skin phenotype and a periodic fever syndrome in Chinese Shar-Pei dogs. *PLoS Genet.* 2011;7:e1001332.
45. Yokoyama JS, Lam ET, Ruhe AL, Erdman CA, Robertson KR, Webb AA, et al. Variation in genes related to cochlear biology is strongly associated with adult-onset deafness in border collies. *PLoS Genet.* 2012;8:e1002898.
46. Agler C, Nielsen DM, Urkasemsin G, Singleton A, Tonomura N, Sigurdsson S, et al. Canine hereditary ataxia in old english sheepdogs and gordon setters is associated with a defect in the autophagy gene encoding RAB24. *PLoS Genet.* 2014;10:e1003991.
47. Downs LM, Mellersh CS. An Intronic SINE insertion in FAM161A that causes exon-skipping is associated with progressive retinal atrophy in Tibetan Spaniels and Tibetan Terriers. *PLoS One.* 2014;9:e93990.

48. Downs LM, Wallin-Håkansson B, Bergström T, Mellersh CS. A novel mutation in TTC8 is associated with progressive retinal atrophy in the golden retriever. *Canine Genet Epidemiol.* 2014;1:4.
49. Drögemüller M, Jagannathan V, Becker D, Drögemüller C, Schelling C, Plassais J, et al. A mutation in the FAM83G gene in dogs with hereditary footpad hyperkeratosis (HFH). *PLoS Genet.* 2014;10:e1004370.
50. Arendt ML, Melin M, Tonomura N, Koltookian M, Courtay-Cahen C, Flindall N, et al. Genome-Wide Association Study of Golden Retrievers Identifies Germ-Line Risk Factors Predisposing to Mast Cell Tumours. *PLoS Genet.* 2015;11:e1005647.
51. Wolf ZT, Brand HA, Shaffer JR, Leslie EJ, Arzi B, Willet CE, et al. Genome-wide association studies in dogs and humans identify ADAMTS20 as a risk variant for cleft lip and palate. *PLoS Genet.* 2015;11:e1005059.
52. Forman OP, Hitti RJ, Pettitt L, Jenkins CA, O'Brien DP, Shelton GD, et al. An Inversion Disrupting FAM134B Is Associated with Sensory Neuropathy in the Border Collie Dog Breed. *G3 (Bethesda).* 2016;6:2687-2692.
53. Hytönen MK, Arumilli M, Lappalainen AK, Owczarek-Lipska M, Jagannathan V, Hundi S, et al. Molecular Characterization of Three Canine Models of Human Rare Bone Diseases: Caffey, van den Ende-Gupta, and Raine Syndromes. *PLoS Genet.* 2016;12:e1006037.

54. Plassais J, Rimbault M, Williams FJ, Davis BW, Schoenebeck JJ, Ostrander EA. Analysis of large versus small dogs reveals three genes on the canine X chromosome associated with body weight, muscling and back fat thickness. *PLoS Genet.* 2017;13:e1006661.
55. Wielaender F, Sarviaho R, James F, Hytönen MK, Cortez MA, Kluger G, et al. Generalized myoclonic epilepsy with photosensitivity in juvenile dogs caused by a defective DIRAS family GTPase 1. *Proc Natl Acad Sci USA.* 2017;114:2669-2674.
56. Drögemüller M, Jagannathan V, Howard J, Bruggmann R, Drögemüller C, Ruetten M, et al. A frameshift mutation in the cubilin gene (CUBN) in Beagles with Imerslund-Gräsbeck syndrome (selective cobalamin malabsorption). *Anim Genet.* 2014;45:148-50.
57. Gilliam D, Kolicheski A, Johnson GS, Mhlanga-Mutangadura T, Taylor JF, Schnabel RD, et al. Golden Retriever dogs with neuronal ceroid lipofuscinosis have a two-base-pair deletion and frameshift in CLN5. *Mol Genet Metab.* 2015;115:101-9.
58. Kolicheski A, Barnes Heller HL, Arnold S, Schnabel RD, Taylor JF, Knox CA, et al. Homozygous PPT1 Splice Donor Mutation in a Cane Corso Dog With Neuronal Ceroid Lipofuscinosis. *J Vet Intern Med.* 2017;31:149-157.

59. Sayyab S, Viluma A, Bergvall K, Brunberg E, Jagannathan V, Leeb T, et al. Whole-Genome Sequencing of a Canine Family Trio Reveals a FAM83G Variant Associated with Hereditary Footpad Hyperkeratosis. *G3 (Bethesda)*. 2016;6:521-527.
60. Gilliam D, O'Brien DP, Coates JR, Johnson GS, Johnson GC, Mhlanga-Mutangadura T, 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;28:871-877.
61. Guo J, Johnson GS, Brown HA, Provencher ML, da Costa RC, Mhlanga-Mutangadura T, 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;112:302-309.
62. Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, et al. A de novo variant in the ASPRV1 gene in a dog with ichthyosis. *PLoS Genet*. 2017;13:e1006651.
63. Ahonen SJ, Arumilli M, Lohi H. A CNGB1 frameshift mutation in Papillon and Phalène dogs with progressive retinal atrophy. *PLoS One*. 2013;8:e72122.
64. Ahram DF, Grozdanic SD, Kecova H, Henkes A, Collin RW, Kuehn MH. Variants in Nebulin (NEB) Are Linked to the Development of Familial Primary Angle Closure Glaucoma in Basset Hounds. *PLoS One*. 2015;10:e0126660.

65. Tsuboi M, Watanabe M, Nibe K, Yoshimi N, Kato A, Sakaguchi M, et al. Identification of the PLA2G6 c.1579G>A Missense Mutation in Papillon Dog Neuroaxonal Dystrophy Using Whole Exome Sequencing Analysis. PLoS One. 2017;12:e0169002.
66. Forman OP, Pettitt L, Komáromy AM, Bedford P, Mellersh C. A Novel Genome-Wide Association Study Approach Using Genotyping by Exome Sequencing Leads to the Identification of a Primary Open Angle Glaucoma Associated Inversion Disrupting ADAMTS17. PLoS One. 2015;10:e0143546.
67. Hayward JJ, Castelhana MG, Oliveira KC, Corey E, Balkman C, Baxter TL, et al. Complex disease and phenotype mapping in the domestic dog. Nat Commun. 2016;7:10460.
68. Karyadi DM, Karlins E, Decker B, vonHoldt BM, Carpintero-Ramirez G, Parker HG, et al. A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit. PLoS Genet. 2013;9:e1003409.
69. Cosford KL, Taylor SM, Thompson L, Shelton GD. A possible new inherited myopathy in a young Labrador retriever. Can Vet J. 2008;49:393-397.
70. Beggs AH, Böhm J, Snead E, Kozlowski M, Maurer M, Minor K, et al. MTM1 mutation associated with X-linked myotubular myopathy in Labrador Retrievers. Proc Natl Acad Sci USA. 2010;107:14697-14702.
71. Laporte J, Hu LJ, Kretz C, Mandel JL, Kioschis P, Coy JF, et al. A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast. Nat Genet. 1996;13:175-182.

72. Childers MK, Joubert R, Poulard K, Moal C, Grange RW, Doering JA, et al. Gene therapy prolongs survival and restores function in murine and canine models of myotubular myopathy. *Sci Transl Med.* 2014;6:220ra10.
73. Shelton GD, Engvall E. Canine and feline models of human inherited muscle diseases. *Neuromuscul Disord.* 2005;15:127-138.
74. Sewry CA, Wallgren-Pettersson C. Myopathology in congenital myopathies. *Neuropathol Appl Neurobiol.* 2017;43:5-23.
75. Pappas CT, Krieg PA, Gregorio CC. Nebulin regulates actin filament lengths by a stabilization mechanism. *J Cell Biol.* 2010;189:859-870.
76. Nowak KJ, Davis MR, Wallgren-Pettersson C, Lamont PJ, Laing NG. Clinical utility gene card for: Nemaline myopathy - update 2015. *Eur J Hum Genet.* 2015;23 doi: 10.1038/ejhg.2015.
77. Corbett MA, Robinson CS, Dungleison GF, Yang N, Joya JE, Stewart AW, et al. A mutation in alpha-tropomyosin (slow) affects muscle strength, maturation and hypertrophy in a mouse model for nemaline myopathy. *Hum Mol Genet.* 2001;10:317-328.
78. de Haan A, van der Vliet MR, Gommans IM, Hardeman EC, van Engelen BG. Skeletal muscle of mice with a mutation in slow alpha-tropomyosin is weaker at lower lengths. *Neuromuscul Disord.* 2002;12:952-957.

79. Joya JE, Kee AJ, Nair-Shalliker V, Ghoddusi M, Nguyen MA, Luther P, et al. Muscle weakness in a mouse model of nemaline myopathy can be reversed with exercise and reveals a novel myofiber repair mechanism. *Hum Mol Genet.* 2004;13:2633-2645.
80. Bang ML, Li X, Littlefield R, Bremner S, Thor A, Knowlton KU, et al. Nebulin-deficient mice exhibit shorter thin filament lengths and reduced contractile function in skeletal muscle. *J Cell Biol.* 2006;173:905–916.
81. Witt CC, Burkart C, Labeit D, McNabb M, Wu Y, Granzier H, et al. Nebulin regulates thin filament length, contractility, and Z-disk structure in vivo. *EMBO J.* 2006;25:3843–3855.
82. Nowak KJ, Ravenscroft G, Jackaman C, Filipovska A, Davies SM, Lim EM, et al. Rescue of skeletal muscle alpha-actin-null mice by cardiac (fetal) alpha-actin. *J Cell Biol.* 2009;185:903–915.
83. Nguyen MA, Joya JE, Kee AJ, Domazetovska A, Yang N, Hook JW, et al. Hypertrophy and dietary tyrosine ameliorate the phenotypes of a mouse model of severe nemaline myopathy. *Brain.* 2011;134:3516–3529.
84. Ravenscroft G, Jackaman C, Bringans S, Papadimitriou JM, Griffiths LM, McNamara E, et al. Mouse models of dominant ACTA1 disease recapitulate human disease and provide insight into therapies. *Brain.* 2011;134:1101–1115.
85. Garg A, O'Rourke J, Long C, Doering J, Ravenscroft G, Bezprozvannaya S, et al. KLHL40 deficiency destabilizes thin filament proteins and promotes nemaline myopathy. *J Clin Invest.* 2014;124:3529–3539.

86. Li F, Buck D, De Winter J, Kolb J, Meng H, Birch C, et al. Nebulin deficiency in adult muscle causes sarcomere defects and muscle-type-dependent changes in trophicity: novel insights in nemaline myopathy. *Hum Mol Genet.* 2015;24:5219–5233.
87. Tian L, Ding S, You Y, Li TR, Liu Y, Wu X, et al. Leiomodin-3-deficient mice display nemaline myopathy with fast-myofiber atrophy. *Dis Model Mech.* 2015;8:635–641.
88. Sztal TE, Zhao M, Williams C, Oorschot V, Parslow AC, Giousoh A, et al. Zebrafish models for nemaline myopathy reveal a spectrum of nemaline bodies contributing to reduced muscle function. *Acta Neuropathol.* 2015;130:389-406.
89. Nigro V, Savarese M. Genetic basis of limb-girdle muscular dystrophies: the 2014 update. *Acta Myol.* 2014;33:1-12.
90. Nigro V. Molecular bases of autosomal recessive limb-girdle muscular dystrophies. *Acta Myol.* 2003;22:35-42.
91. Ng R, Banks GB, Hall JK, Muir LA, Ramos JN, Wicki J, et al. Animal models of muscular dystrophy. *Prog Mol Biol Transl.* 2012;105:83.
92. Nigro V, Okazaki Y, Belsito A, Piluso G, Matsuda Y, Politano L, et al. Identification of the Syrian hamster cardiomyopathy gene. *Hum Mol Genet.* 1997;6:601-607.
93. Straub V, Duclos F, Venzke DP, Lee JC, Cutshall S, Leveille CJ, et al. Molecular Pathogenesis of Muscle Degeneration in the δ -Sarcoglycan-Deficient Hamster. *Am J Pathol.* 1998;153:1623-1630.

94. Mendell JR, Rodino-Klapac LR, Rosales-Quintero X, Kota J, Coley BD, Galloway G, et al. Limb-girdle muscular dystrophy type 2D gene therapy restores alpha-sarcoglycan and associated proteins. *Ann Neurol.* 2009;66:290-297.
95. Mendell JR, Rodino-Klapac LR, Rosales XQ, Coley BD, Galloway G, Lewis S, et al. Sustained alpha-sarcoglycan gene expression after gene transfer in limb-girdle muscular dystrophy, type 2D. *Ann Neurol.* 2010;68:629-638.
96. Rider LG, Nistala K. The juvenile idiopathic inflammatory myopathies: pathogenesis, clinical and autoantibody phenotypes, and outcomes. *J Intern Med.* 2016;280:24–38.
97. Papadopoulou C, Wedderburn LR. Treatment of Juvenile Dermatomyositis: An Update. *Paediatr Drugs.* 2017 doi: 10.1007/s40272-017-0240-6.
98. Stübgen JP. Juvenile dermatomyositis/polymyositis and lymphoma. *J Neurol Sci.* 2017;377:19-24.
99. Miller FW, Cooper RG, Vencovský J, Rider LG, Danko K, Wedderburn LR, et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis Rheum.* 2013;65:3239–3247.
100. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes Immun.* 2015;16:470–480.
101. Kunkle GA, Chrisman CL, Gross TL, Fadok V, Werner LL. Dermatomyositis in collie dogs. *Comp Contin Educ Pract Vet.* 1985;7:185–192.

102. Hargis A, Mundell A. Familial canine dermatomyositis. *Compend Contin Educ Vet.* 1992;4:855–864.
103. Scott DW, Miller WH, Griffin CE. Congenital and hereditary defects. In: Muller and Kirk's *Small Animal Dermatology*, 6th ed. Philadelphia: Saunders, 2000 pp. 940–946.
104. Hargis AM, Prieur DJ, Haupt KH, McDonald TL, Moore MP. Prospective study of familial canine dermatomyositis: correlation of the severity of dermatomyositis and circulating immune complex levels. *Am J Pathol.* 1986;123:465–479.
105. Dourmishev LA, Dourmishev AL. *Dermatomyositis: Advances in Recognition, Understanding and Management.* New York: Springer; 2009.

CHAPTER TWO

EXOME SEQUENCING REVEALS A NEBULIN NONSENSE MUTATION IN A DOG MODEL OF NEMALINE ROD MYOPATHY

Jacquelyn M. Evans¹, Melissa L. Cox², Jonathan Huska³, Frank Li⁴, Luis Gaitero³, Ling
T. Guo⁵ Margaret L. Casal⁶, Henk L. Granzier⁴, G. Diane Shelton^{5*}, Leigh Anne Clark^{1*}

¹ Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634,
USA

² CAG GmbH - Center for Animal Genetics, Tuebingen, Germany

³ Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada

⁴ Department of Cellular and Molecular Medicine, University of Arizona, Tucson, AZ
85724, USA

⁵ Department of Pathology, University of California San Diego, La Jolla, CA 92093,
USA

⁶ School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104,
USA

*Co-corresponding authors:

Leigh Anne Clark

e-mail: lclark4@clemson.edu

Telephone number: 864-656-4696

Fax number: 864-656-6879

G. Diane Shelton

e-mail: gshelton@ucsd.edu

Telephone number: 858-534-1537

Fax number: 858-534-0391

Published – Mammalian Genome

Reprinted with permission from Springer, see Appendix D

Abstract

Nemaline myopathy (NM) is a congenital muscle disorder associated with muscle weakness, hypotonia, and rod bodies in the skeletal muscle fibers. Mutations in 10 genes have been implicated in human NM, but spontaneous cases in dogs have not been genetically characterized. We identified a novel recessive myopathy in a family of line-bred American bulldogs (ABDs); rod bodies in muscle biopsies established this as NM. Using SNP profiles from the nuclear family, we evaluated inheritance patterns at candidate loci and prioritized *TNNT1* and *NEB* for further investigation. Whole exome sequencing of the dam, two affected littermates, and an unaffected littermate revealed a nonsense mutation in *NEB* (g.52734272 C>A, S8042X). Whole tissue gel electrophoresis and western blots confirmed a lack of full-length NEB in affected tissues, suggesting nonsense-mediated decay. The pathogenic variant was absent from 120 dogs of 24 other breeds and 100 unrelated ABDs, suggesting that it occurred recently and may be private to the family. This study presents the first molecularly characterized large animal model of NM, which could provide new opportunities for therapeutic approaches.

Introduction

One of the most common human congenital myopathies, accounting for 17% of cases, nemaline myopathy (NM) is characterized by the presence of rod bodies in the skeletal muscle fibers, muscle weakness, and hypotonia [1]. Proximal limb muscles, facial muscles, and neck flexors are most frequently affected by muscle weakness, while respiratory insufficiency is the primary cause of death [1]. NM is clinically heterogeneous and is classified into six subtypes based on severity, pattern of muscle weakness, and age of onset [2]. Ten genes have been implicated in NM: *ACTA1*, *CFL2*, *KBTBD13*, *KLHL40*, *KHL41*, *LMOD3*, *NEB*, *TNNT1*, *TPM2*, and *TPM3* [3]. These genes encode proteins associated with the skeletal muscle sarcomere thin filament or Kelch domain proteins [4].

While founder mutations have been reported in Ashkenazi Jewish (*NEB*) [5], Old Order Amish (*TNNT1*) [6], and Turkish (*TPM3*) [7] populations, most NM cases involve de novo mutations. Among cases explained at the genetic level, most are attributed to recessive mutations in *NEB* (usually in compound heterozygosity) and dominant mutations in *ACTA1* [1]. It is often infeasible to identify de novo mutations through Sanger sequencing due to numerous candidate genes and the prohibitive size of the *NEB* transcript (~26 kb, 183 exons). In recent years, next generation sequencing technologies that allow for simultaneous capture of candidate genes have expedited the genetic characterization of individual NM cases [8-11].

Mouse models harboring mutations in *ACTA1* [12-14], *LMOD3* [15], *KLHL40* [16], *NEB* [17-19], and *TPM3* [20,21] are available for study of NM subtypes, but no

large animal models have been described. Dog models have been instrumental in advancing therapeutic strategies for hereditary muscle disorders such as Duchenne muscular dystrophy and myotubular myopathy [22,23]. While spontaneously occurring forms of NM have been reported in a Border collie and a Schipperke, they were not characterized at the molecular level [24]. Herein, we describe a novel NM in a family of American bulldogs (ABDs) and determine the pathogenic variant through genome-wide SNP profiling and whole exome sequencing (WES).

Materials & Methods

Animals

A five month old male ABD was evaluated at the University of Guelph Veterinary Teaching Hospital for non-progressive generalized muscle weakness, exercise intolerance, and tremors beginning at approximately two months of age. A female littermate had similar clinical signs, while another female littermate was reported to be clinically unaffected. Two other littermates could not be located. The sire and dam of the litter were cousins, and there was no prior history of muscle disease in the family. Samples from the family were collected for diagnostic procedures and submitted with owner consent.

Histopathology and immunohistochemistry

Biopsies from the triceps, biceps femoris, and suprascapular muscles were collected from both affected dogs under general inhalational anesthesia following electrophysiological examination. The biopsy specimens were either snap frozen in

isopentane (pre-cooled in liquid nitrogen) or immersion fixed in 2.5% glutaraldehyde for electron microscopy. Sections (8 μm) were further processed using a standard panel of histological and histochemical stains and reactions [25]. Similar staining and reactions were performed on age-matched control muscles from the tissue archives of the Comparative Neuromuscular Laboratory.

Electron Microscopy

Glutaraldehyde-fixed muscle specimens were post-fixed in 1% aqueous osmium tetroxide prior to dehydration and embedding in araldite resin. Thick sections (1 μm) were stained with toluidine blue-basic fuchsin prior to light microscopic examinations, while thin sections (60-90 nm) were stained with uranyl acetate and lead citrate prior to examination in a Zeiss 10 electron microscope.

DNA preparation

DNA was extracted from muscles of both affected dogs and blood from the sire, dam, and unaffected littermate using the DNeasy extraction kit (Qiagen, Hilden, Germany). Whole blood samples from ABDs recruited for an unrelated study were obtained as controls [26]. DNA was isolated following the Genra PureGene protocol (Qiagen). Genomic DNAs from purebred dogs of other breeds were available from our private DNA bank at Clemson University.

Illumina SNP arrays

To date, 10 genes have been identified in human cases of NM; their canine counterparts lie on 10 different chromosomes. To investigate inheritance patterns at each gene, SNP profiles were generated for the nuclear family using the Illumina CanineHD BeadChip, containing 173,662 SNPs (Illumina, San Diego, CA, USA). Genotypes at polymorphic SNPs encompassing the candidate genes (flanking the genes within 1 Mb) were examined for consistency with an autosomal recessive mode of inheritance.

Whole exome sequencing

Whole exome sequencing (WES) data were generated for the dam, two affected siblings, and an unaffected sibling. Genomic DNA (1 µg) from each sample was mechanically sheared to fragments of approximately 180-250 bp (Covaris LE220, Woburn, MA, USA). Fragment sizes were verified for quality control (Fragment Analyzer, Advanced Analytical, Ankeny, IA, USA). The fragment library was hybridized with 120-mer biotinylated RNA baits from the SureSelect XT Canine All Exon kit (Agilent, Santa Clara, CA, USA), which was designed based on the UCSC CanFam2 Ensembl and Refseq tracks as well as human protein alignments. Magnetic streptavidin beads were used for purification according to the manufacturer's instructions. Library DNA was amplified, sequencing barcodes and adapters were added (Illumina), and equimolar amounts of each sample were pooled. The pool was sequenced on both lanes of a Rapid Flowcell on a HiSeq2500 instrument (Illumina), generating paired-end 2x100 bp sequences, comprising on average 6 GB per sample. Sequencing data were

demultiplexed (Illumina bcl2fastq 1.8.2), sequencing adapters were trimmed (skewer 0.1.116), and the resulting sequence mapped to the canine genome (CanFam3.1) using the Burrows-Wheeler Aligner (bwa 0.7.2-r351). PCR duplicates and low-quality alignments were removed (samtools 0.1.18 and internal software). Variant calling was performed using a dual pipeline of bcftools (0.1.17) and varscan (2.3.5), and internal software was used to combine these files into a single variant call file (VCF) per sample.

Reads were visualized using Integrated Genomics Viewer (IGV) [27]. Known variants from Ensembl dbSNP (Can Fam3.1) and from private whole genome sequence databases were excluded. *NEB* was manually screened for variants predicted to have an effect on the protein.

Genotyping of g.52734272

Primers for PCR were designed to capture a 489 bp region encompassing g.52734272 (Forward 5'-AAGTCCCAGCAGCAACATAA-3', Reverse 5'-GTCCAAAGTGGTCGGTCCT-3'). Products were amplified using ReddyMix master mix (Thermo Scientific, Waltham, MA, USA) with 0.4 μ M primers, 50 ng DNA, and water for a total volume of 25 μ L. Thermal cycling conditions were as follows: 95°C for 5 min; 5 touchdown cycles of 95°C for 30 s, 58°C for 30 s reducing 1°C per cycle, and 72°C for 1 min; 31 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 1 min; and a 10 min final elongation at 72°C. Direct sequencing was carried out with BigDye Terminator using an ABI 3730xl Genetic Analyzer to validate g.52734272.

Using the same primers and thermal cycling conditions described above, genotyping of g.52734272 in unrelated ABDs and other breeds was carried out through a restriction enzyme digest with 5 μ L PCR product and either BfaI CutSmart (New England Biolabs, Ipswich, MA, USA) or FspBI FastDigest (Thermo Scientific) for total reaction volumes of 25 μ L and 15 μ L, respectively. BfaI (FspBI) recognizes and cleaves (^) the following sequence: 5'- C^TAG -3'. Digests were visualized on a 1.2% agarose gel, where wild type alleles are uncut (489 bp) and mutant alleles are cut once (248 bp and 241 bp).

Gel electrophoresis and western blotting

Frozen tissue samples were ground to a fine powder using a glass mortar and pestle chilled in liquid nitrogen. After 20 min of priming at -20°C, tissue was resuspended in a 1:1 mixture of an 8 M urea buffer (8 M urea, 2 M thiourea, 0.05 M Tris-HCl, 0.075 M dithiothreitol, as well as 3% SDS and 0.03% bromophenol blue, pH 6.8) and 50% glycerol with protease inhibitors (0.04 mM E-64, 0.16 mM leupeptin, and 0.2 mM PMSF). The solutions were mixed in a 60°C water bath for 4 min, followed by a 10 min incubation at the same temperature. Residual debris was removed via centrifugation at 13000 rpm, and the supernatant was flash frozen for storage at -80°C. Initial protein analysis was done via a 2-7% gradient acrylamide gel. Western blot was performed on the samples using 0.8% agarose gels run at 15 mA per gel for 2 hours and 50 minutes. Following this, they were transferred onto Immobilon-P PVDF membranes (Millipore, Billerica, MA, USA) using a semi-dry transfer unit (Bio-Rad, Hercules, CA, USA) for

2.5 h at 164 mA. Membranes were briefly stained with Ponceau S to check for transferred proteins. Following removal of the stain, membranes were incubated with primary antibody at 4°C overnight. Nebulin N-terminal and C-terminal expression was quantified using primary antibodies specific to those regions (Myomedix #6969 and #6964 respectively, <http://www.myomedix.com>). Expression was normalized to the integrated optical density of myoglobin heavy chain (MHC) obtained from the Ponceau S staining. The fluorescence of the western blots was analyzed using Odyssey Infrared Imaging System. Ponceau S images were analyzed with One-D scan EX (Scanalytics Inc., Rockville, MD, USA).

Results

Affected dogs could independently ambulate, had generalized atrophy, and the myopathy was relatively non-progressive (Supplemental video). Atrophy of the cervical and dorsal thoracic limb muscles was noted with bilateral hypertrophy of the triceps muscles. Serum creatine kinase (CK) activities were mildly elevated. Electromyography (EMG) revealed spontaneous electrical activity, consisting mainly of fibrillation potentials, within the proximal appendicular muscles of the thoracic limbs and the cervical paraspinal musculature. Motor nerve conduction velocity (MNCV) testing showed a mild decrease in the latency of the tibial and ulnar nerves. Respiratory difficulties were not present.

A marked variability in myofiber size and generalized atrophy was present in muscles from the affected ABDs (Fig. 1a) compared to control muscle (Fig. 1d). The

predominant abnormality found in >50% of the muscle fibers from all affected muscles examined was rod-like inclusion bodies highlighted with the modified Gomori trichrome stain (Fig. 1b); these structures were not observed in control muscles (Fig. 1e) or with H&E staining (Fig. 1a,d). Rod-like structures tended to be centralized or peripherally distributed in the muscle fibers and were present in both slow twitch type 1 and fast twitch type 2 muscle fibers. Atrophic fibers were also of both fiber types (Fig. 1c,f). Numerous rods were apparent along the long axis parallel to that of the muscle filaments (Fig. 1g). The rods were in structural continuity with Z disks (Fig. 1h), had the same electron density as the Z lines of adjacent sarcomeres, and had a similar lattice pattern of periodic cross-striations.

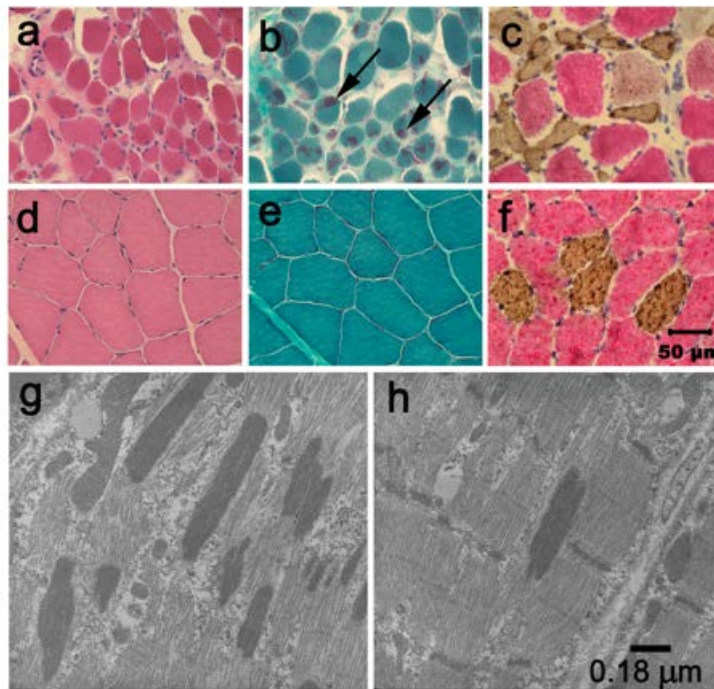
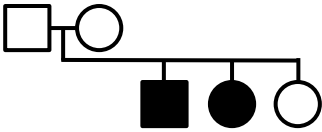


Figure 1. Histopathology establishes NM. Cryosections from the triceps muscle of an affected pup (a–c) and archived control triceps muscle (d–f) were stained with H&E (a, d), modified Gomori trichrome (b, e) and following incubation with monoclonal antibodies against type 1 and type 2 myosin heavy chains (c, f). Excessive variability in myofiber size and atrophy were observed in the affected muscle with the H&E stain (a) compared to control muscle (d). Numerous myofibers in the affected muscle contained rod bodies (b, *arrows*) not evident in control muscle (e). Both type 1 and type 2 fibers were atrophic (c) with fibers of both fiber types similar in size in control muscle (f). *Bar* 50 μm for images a–f. By electron microscopy, numerous electron dense rods were apparent along the long axis parallel to that of the muscle fiber (g). The rods were in structural continuity with Z disks (h), had the same electron density as the Z lines of adjacent sarcomeres, and had a similar lattice pattern of periodic cross striations. *Bar* 0.18 μm for images g and h

The absence of disease in the sire and dam indicates that transmission is likely autosomal recessive. We hypothesized that the parents were heterozygous for a deleterious allele inherited identical-by-descent through a common grandparent and that their affected progeny were homozygous. Allelic inheritance was manually evaluated for 28 polymorphic SNPs within or flanking the 10 candidate genes; only markers representing *TNNT1* (CFA1) and *NEB* (CFA19) were consistent with a simple autosomal recessive pattern (Table 1).



Gene (Chr)	SNP position					
<i>TNNT1</i> (1)	105264488	AG	AG	GG	GG	AG
	105432142	GG	AG	AG	AG	GG
	105598975	GG	CG	CG	CG	GG
	106497529	AG	AG	AA	AA	AG
<i>ACTA1</i> (4)	12713391	AG	AA	AG	AG	AA
	12875911	AG	AA	AG	AG	AA
<i>TPM3</i> (7)	45997866	CC	AC	CC	CC	CC
	46022207*	AT	AA	AA	AT	AT
	46034262*	AC	AA	AA	AC	AC
	46049585*	AG	AG	AA	AG	AG
	46303991	CC	AC	CC	CC	CC
<i>CFL2</i> (8)	16507437	CC	AC	CC	AC	CC
	16587360	GG	AG	GG	AG	GG
<i>TPM2</i> (11)	55241525*	AC	CC	CC	AC	AC
	55248320*	AA	AG	AA	AG	AA
<i>NEB</i> (19)	55761948*	GG	AG	GG	GG	GG
	55777049*	GG	AG	GG	GG	GG
	55824838*	AG	AG	AA	AA	AG
<i>LMOD3</i> (20)	25536336	AG	AA	AA	AA	AG
	25605665	AA	GG	AG	AG	AG
	25625642	AG	AA	AA	AA	AG
<i>KLHL40</i> (23)	14709678	AC	AA	AC	AC	AC
	14867876	AG	GG	AG	AG	AG
<i>KBTBD13</i> (30)	32370798	AG	GG	GG	GG	AG
	32443189	AG	GG	GG	GG	AG
	32458624	AC	AA	AA	AA	AC
<i>KLHL41</i> (36)	17244342	AG	AA	AA	AA	AG
	17280474	AG	GG	AG	AG	GG

Table 1. Genotypes for the nuclear family are reported for polymorphic SNPs (CanFam2) representing each of the 10 candidate genes. When possible, SNPs lying within the gene are used (asterisks); otherwise, the most proximal flanking SNPs are reported. When present, flanking SNPs (within 1 Mb) for which the affected dogs are homozygous are also shown. SNPs fitting a recessive pattern of inheritance are highlighted in bold

We conducted WES of four family members because of the large size of *NEB*. Across the samples, exome coverage, mappable reads, yields, and mean quality scores averaged 30X, 59 million, 6 Mb, and 37.1, respectively. VCFs were used to filter variants within the two candidate genes for those fitting recessive transmission (homozygous alternate in both affected dogs, heterozygous in the dam, and heterozygous or homozygous reference in the unaffected sibling). This step eliminated all 11 variants identified within *TNNT1* and 343 variants in *NEB*. Thirty-six remaining variants in *NEB* were filtered to remove common polymorphisms, non-coding variants, and non-damaging variants (Fig. 2a). Only one overtly deleterious variant was identified in *NEB*, a nonsense mutation g.52734272 C>A, corresponding to human exon 169 (S8042X; NP_001258137.1). In addition, we confirmed that no variants in the other eight candidate genes segregated with the phenotype. The *NEB* nonsense mutation was verified by Sanger sequencing (Fig. 2b) and was absent in 100 unrelated ABDs and five dogs each from 24 other breeds, determined by restriction digest genotyping.

Protein analysis of the triceps biopsies from the affected dogs suggested a loss of NEB, which was then confirmed via western blotting (Fig. 3a,b). These values were quantified, revealing NEB expression at 0.3% and 16.2% of their wild-type counterparts for the N-terminus and C-terminus, respectively (Fig. 3c,d).

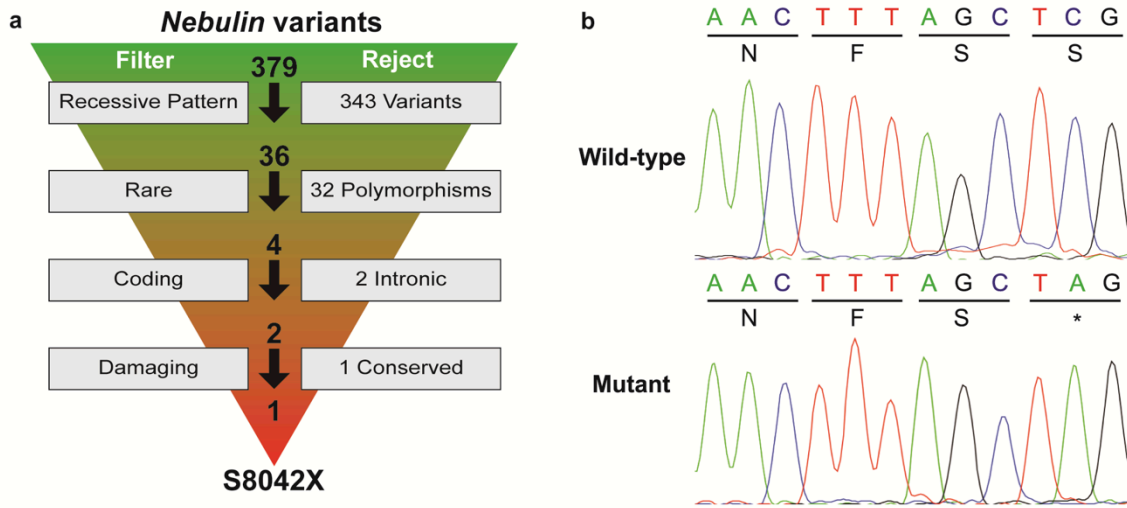


Figure 2. Filtering of *NEB* variants reveals S8042X. (a) Filtering parameters are shown to the left with the number of rejected variants to the right. Total variants prior to each filtering parameter are shown above the arrows, beginning with the total number of variants. (b) Chromatograms from Sanger sequencing show the wild-type and mutant alleles.

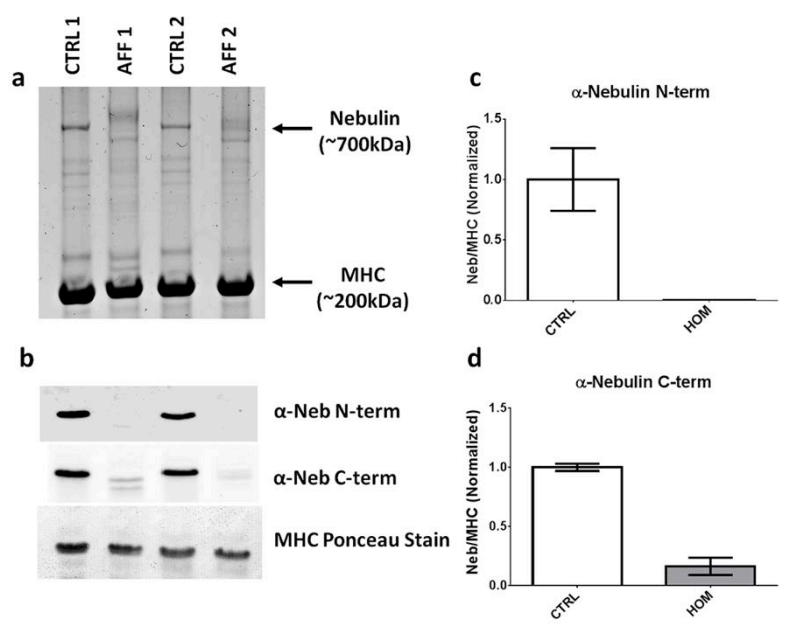


Figure 3. Nebulin is drastically reduced in skeletal muscles with S8042X. (a) Protein analysis on a 2-7% gradient acrylamide gel reveals a clear loss of nebulin in muscles from affected dogs compared to healthy controls. (b) Western blots for the same samples using antibodies to nebulin's N- and C-terminus. (c) and (d) Average quantification (CTRL 1 and CTRL 2; AFF 1 and AFF 2) of western blots, indicating a severe loss of nebulin protein in affected dogs.

Discussion

Histological and clinical findings from the affected ABDs are consistent with primary NM. While the nemaline rod bodies are hallmarks of the disease, these morphological features provide little indication of subtype or genetic etiology. The most common of the six subtypes is typical congenital NM, which is most often attributed to mutations in *NEB* [1,2]. The clinical presentation of independent ambulation and relatively non-progressive muscle weakness at two months of age in the ABDs is consistent with typical congenital NM [28].

We used a bipartite approach to uncover a nonsense mutation in *NEB* causing ABD NM. To refine the list of 10 known candidate genes, we generated genome-wide SNP profiles. Affected dogs inherited identical SNP haplotypes spanning *TNNT1* and *NEB*; therefore, these genes were prioritized for further study. Together, they comprise nearly 200 exons, rendering Sanger sequencing impractical. Instead, we utilized WES, an approach that provides high exon read coverage and the data to parse the rest of the exome if no mutations are found in candidate genes.

NEB, or nebulin, is a giant protein (600-900 kDa) that stabilizes the actin thin filament of skeletal muscle sarcomeres and is critical for proper thin filament formation and muscle contraction [29]. *NEB* mutations account for up to 50% of molecularly characterized NM cases [30]; most are truncating [31]. *NEB* has a multitude of isoforms achieved primarily through alternative splicing of three sets of exons (63-66, 143-144, and 167-177); truncating *NEB* mutations are frequently found in the latter set. In general, homozygous truncating mutations may be less tolerated in ubiquitously expressed exons

[31]. Consistent with these trends observed in human myopathies, S8042X lies in the alternatively spliced exons 167-177.

We elected to investigate NEB protein levels, rather than transcript presence, in cases and controls due to limited tissue quantities. Protein gels show little banding in the NEB region, a finding consistent with the absence of protein detected with the N-terminus anti-nebulin antibody. The C-terminus antibody showed a low level of protein (~16% of the levels in the control tissues) but with a mobility slightly larger than that of NEB in the control samples. It is possible that this is a unique splice isoform of nebulin that has low abundance and excludes the N-terminus. Alternatively, the C-terminus antibody may be cross-reacting with another protein similar in size or a degradation product of a larger protein. We consider the cross-reactivity hypothesis more likely because we have detected titin degradation products of a size similar to NEB that cross-react with NEB C-terminus antibodies in previous work. Regardless, our data suggest that S8042X results in very low levels of NEB, probably due to nonsense-mediated decay. The inability to successfully express normal levels of full-length nebulin likely underlies the pathology of the affected ABDs.

While we report the first mutation causing NM in dogs, this is not the first pathogenic variant described in canine *NEB*. Ahram et al. [32] identified *NEB* missense mutations associated with primary angle-closure glaucoma (PACG) in Basset hounds. Interestingly, affected Basset hounds display no clinical signs of NM [32]. Ocular abnormalities were not identified in either of the ABDs with NM, nor are they associated with human NM attributed to *NEB*.

The recent development of WES enrichment kits for canines [33,34] will facilitate faster discovery of pathogenic variants, particularly when candidate genes are sizable or numerous. Massive parallel sequencing of several exome libraries in a single hi-seq lane is an economical approach for sequencing of multiple individuals, relative to whole genome resequencing, and permits simultaneous identification and filtering of variants. To date, this report is one of few to use WES in a nuclear family to identify a pathogenic variant in dogs [32,35]. Rather than requiring the widespread distribution of the ABD mutation throughout the population in order to gather sufficient cases for analysis with traditional methods, WES allowed us to detect the mutation directly in the two probands.

Although no carriers were found in a survey of unrelated ABDs (extended family members were unavailable for study), a genetic test now exists for S8042X should NM arise again in the breed. The identification of carriers holds potential for the development of a colony for testing of therapeutic approaches for humans.

Acknowledgments

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Numbers R15AR062868 and R01AR053897. The authors wish to thank the dog owners for contributing samples to this study, and James R. Clark for helping prepare figure 2.

References

1. Romero NB, Clarke NF (2013) Congenital myopathies. *Handb Clin Neurol* 113, 1321-1336
2. Wallgren-Pettersson C, Sewry CA, Nowak KJ, Laing NG. Nemaline myopathies. *Semin Pediatr Neurol*. 2011;18:230-238.
3. Nowak KJ, Davis MR, Wallgren-Pettersson C, Lamont PJ, Laing NG. Clinical utility gene card for: Nemaline myopathy. *Eur J Hum Genet*. 2015:23
4. Malfatti E, Lehtokari VL, Böhm J, De Winter JM, Schäffer U, Estournet B, et al. Muscle histopathology in nebulin-related nemaline myopathy: ultrastructural findings correlated to disease severity and genotype. *Acta Neuropathol Commun*. 2014;2:44.
5. Anderson SL, Ekstein J, Donnelly MC, Keefe EM, Toto NR, LeVoci LA, et al. Nemaline myopathy in the Ashkenazi Jewish population is caused by a deletion in the nebulin gene. *Hum Genet*. 2004;115:185-190.
6. Johnston JJ, Kelley RI, Crawford TO, Morton DH, Agarwala R, Koch T, et al. A novel nemaline myopathy in the Amish caused by a mutation in troponin T1. *Am J Hum Genet*. 2000;67:814-821.
7. Lehtokari VL, Pelin K, Donner K, Voit T, Rudnik-Schöneborn S, Stoetter M, et al. Identification of a founder mutation in TPM3 in nemaline myopathy patients of Turkish origin. *Eur J Hum Genet*. 2008;16:1055-1061.

8. Chen Z, Wang JL, Tang BS, Sun ZF, Shi YT, Shen L, et al. Using next-generation sequencing as a genetic diagnostic tool in rare autosomal recessive neurologic Mendelian disorders. *Neurobiol Aging*. 2013;34:2442.e11-7.
9. Scoto M, Cullup T, Cirak S, Yau S, Manzur AY, Feng L, et al. Nebulin (NEB) mutations in a childhood onset distal myopathy with rods and cores uncovered by next generation sequencing. *Eur J Hum Genet*. 2013;21:1249-1252.
10. Güttsches AK, Dekomien G, Claeys KG, von der Hagen M, Huebner A, Kley RA, et al. Two novel nebulin variants in an adult patient with congenital nemaline myopathy. *Neuromuscul Disord* 2015;25:392-396.
11. Marra JD, Engelstad KE, Ankala A, Tanji K, Dastgir J, De Vivo DC, et al. Identification of a novel nemaline myopathy-causing mutation in the troponin T1 (TNNT1) gene: a case outside of the old order Amish. *Muscle Nerve*. 2015;51:767-772.
12. Nowak KJ, Ravenscroft G, Jackaman C, Filipovska A, Davies SM, Lim EM, et al. Rescue of skeletal muscle alpha-actin-null mice by cardiac (fetal) alpha-actin. *J Cell Biol*. 2009;185:903-915.
13. Nguyen MA, Joya JE, Kee AJ, Domazetovska A, Yang N, Hook JW, et al. Hypertrophy and dietary tyrosine ameliorate the phenotypes of a mouse model of severe nemaline myopathy. *Brain* 2011;134:3516-3529.
14. Ravenscroft G, Jackaman C, Bringans S, Papadimitriou JM, Griffiths LM, McNamara E, et al. Mouse models of dominant ACTA1 disease recapitulate human disease and provide insight into therapies. *Brain* 2011;134:1101-1115.

15. Tian L, Ding S, You Y, Li TR, Liu Y, Wu X, et al. Leiomodin-3-deficient mice display nemaline myopathy with fast-myofiber atrophy. *Dis Model Mech*. 2015;8:635-641.
16. Garg A, O'Rourke J, Long C, Doering J, Ravenscroft G, Bezprozvannaya S, et al. KLHL40 deficiency destabilizes thin filament proteins and promotes nemaline myopathy. *J Clin Invest*. 2014;124:3529-3539.
17. Bang ML, Li X, Littlefield R, Bremner S, Thor A, Knowlton KU, et al. Nebulin-deficient mice exhibit shorter thin filament lengths and reduced contractile function in skeletal muscle. *J Cell Biol*. 2006;173:905-916.
18. Witt CC, Burkart C, Labeit D, McNabb M, Wu Y, Granzier H, et al. Nebulin regulates thin filament length, contractility, and Z-disk structure in vivo. *EMBO J*. 2006;25:3843-3855.
19. Li F, Buck D, De Winter J, Kolb J, Meng H, Birch C, et al. Nebulin deficiency in adult muscle causes sarcomere defects and muscle-type-dependent changes in trophicity: novel insights in nemaline myopathy. *Hum Mol Genet*. 2015;24:5219-5233.
20. Corbett MA, Robinson CS, Duglison GF, Yang N, Joya JE, Stewart AW, et al. A mutation in alpha-tropomyosin (slow) affects muscle strength, maturation and hypertrophy in a mouse model for nemaline myopathy. *Hum Mol Genet*. 2001;10:317-328.

21. de Haan A, van der Vliet MR, Gommans IM, Hardeman EC, van Engelen BG
Skeletal muscle of mice with a mutation in slow alpha-tropomyosin is weaker at lower lengths. *Neuromuscul Disord.* 2002;12:952-957.
22. Kornegay JN, Bogan JR, Bogan DJ, Childers MK, Li J, Nghiem P, et al. Canine models of Duchenne muscular dystrophy and their use in therapeutic strategies. *Mamm Genome* 2012;2:85-108.
23. Childers MK, Joubert R, Poulard K, Moal C, Grange RW, Doering JA, et al. Gene therapy prolongs survival and restores function in murine and canine models of myotubular myopathy. *Sci Transl Med.* 2014;6:220ra10.
24. Shelton GD, Engvall E. Canine and feline models of human inherited muscle diseases. *Neuromuscul Disord.* 2005;15:127-138.
25. Dubowitz V, Sewry CA, Oldfors A. *Muscle Biopsy: A practical Approach.* 4th ed. Oxford, England: Saunders Elsevier; 2013.
26. Mauldin EA, Wang P, Evans E, Cantner CA, Ferracone JD, Credille KM, et al. Autosomal recessive congenital ichthyosis in American bulldogs is associated with NIPAL4 (ICHTHYIN) deficiency. *Vet Pathol.* 2014;52:654-662.
27. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013;14:178-192.

28. North KN and Ryan MM (2002) Nemaline Myopathy. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Fong CT, Mefford HC, Smith RJH, Stephens K editors. GeneReviews, University of Washington, Seattle; 2002.
29. Labeit S, Ottenheijm CA, Granzier H. Nebulin, a major player in muscle health and disease. FASEB J. 2011;25:822-829.
30. North KN. Clinical approach to the diagnosis of congenital myopathies. Semin Pediatr Neurol. 2011;18:216-220.
31. Lehtokari VL, Kiiski K, Sandaradura SA, Laporte J, Repo P, Frey JA, et al. Mutation update: the spectra of nebulin variants and associated myopathies. Hum Mutat. 2014;35:1418-1426.
32. Ahram DF, Grozdanic SD, Kecova H, Henkes A, Collin RW, Kuehn MH. Variants in *nebulin* (*NEB*) are linked to the development of familial primary angle closure glaucoma in Basset hounds. PLoS One 2015;10:e0126660.
33. Broeckx BJG, Coopman F, Verhoeven GEC, Bavegems V, De Keulenaer S, De Meester E, et al. Development and performance of a targeted whole exome sequencing enrichment kit for the dog (*Canis Familiaris* Build 3.1). Sci Rep. 2014;4:5597.
34. Broeckx BJG, Hitte C, Coopman F, Verhoeven GEC, De Keulenaer S, De Meester E, et al. Improved canine exome designs, featuring ncRNAs and increased coverage of protein coding genes. Sci Rep. 2015;5:12810.

35. Ahonen SJ, Arumilli M, Lohi H. A CNGB1 frameshift mutation in Papillon and Phalene dogs with progressive retinal atrophy. PLoS One, 2013;e72122.

CHAPTER THREE

EXOME SEQUENCING REVEALS INDEPENDENT SGCD DELETIONS CAUSING LIMB GIRDLE MUSCULAR DYSTROPHY IN BOSTON TERRIERS

Cox ML^{1†}, Evans JM^{2†}, Davis AG², Guo LT³, Levy JR⁴, Starr-Moss AN², Salmela E^{5,6},
Hytönen MK^{5,6}, Lohi H^{5,6}, Campbell KP⁴, Clark LA^{2*}, Shelton GD^{3*}

† These authors contributed equally to this work.

¹CAG GmbH - Center for Animal Genetics, 72076 Tübingen, Germany

²Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634,
USA

³Department of Pathology, University of California San Diego, La Jolla, CA 92093, USA

⁴ Howard Hughes Medical Institute, Senator Paul D. Wellstone Muscular Dystrophy
Cooperative Research Center, Department of Molecular Physiology and Biophysics, and
Department of Neurology, Roy J. and Lucille A. Carver College of Medicine, The
University of Iowa, Iowa City, Iowa 52242, USA

⁵Department of Veterinary Biosciences and Research Programs Unit, University of
Helsinki, Helsinki, Finland

⁶Folkhälsan Institute of Genetics, Helsinki, Finland

*Co-corresponding authors:

G. Diane Shelton

gshelton@ucsd.edu

Leigh Anne Clark

lclark4@clemson.edu

Published – Skeletal Muscle

Copyright: © The Author(s). 2017 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Abstract

Background

Limb-girdle muscular dystrophies (LGMDs) are a heterogeneous group of inherited autosomal myopathies that preferentially affect voluntary muscles of the shoulders and hips. LGMD has been clinically described in several breeds of dogs, but the responsible mutation(s) are unknown. The clinical presentation in dogs is characterized by marked muscle weakness and atrophy in the shoulder and hips during puppyhood.

Methods

Following clinical evaluation, the identification of the dystrophic histological phenotype on muscle histology, and demonstration of the absence of sarcoglycan-sarcospan complex by immunostaining, whole exome sequencing was performed on five Boston terriers: one affected dog and three family members, and one unrelated affected dog.

Results

Within *sarcoglycan- δ* (*SGCD*), a two base pair deletion segregating with LGMD in the family was discovered, and a deletion encompassing exons 7 and 8 was found in the unrelated dog. Both mutations are predicted to cause an absence of *SGCD* protein, confirmed by immunohistochemistry. The mutations are private to each family.

Conclusions

Here we describe the first cases of canine LGMD characterized at the molecular level with the classification of LGMD2F.

Keywords

muscle, myopathy, sarcoglycanopathy, dog, LGMD

Background

Limb-girdle muscular dystrophies (LGMDs) are a heterogeneous group of Mendelian disorders affecting voluntary muscles of the shoulders and hips [1]. While proximal limb muscles are primarily affected in LGMD, other muscles may degenerate as well, such as the heart and respiratory muscles [1]. Sarcoglycanopathies are a subset of severe, recessive LGMDs (LGMD2C-F) that present in early childhood [2]. There are six known sarcoglycan genes (*SGCA*, *SGCB*, *SGCD*, *SGCG*, *SGCE*, and *SGCZ*); the first

four encode single-pass transmembrane glycoproteins (α -, β -, δ -, γ -sarcoglycans) and, along with sarcospan, make up the tetrameric sarcoglycan-sarcospan complex (SGC). As part of the dystrophin-glycoprotein complex, the SGC is critical for maintaining sarcolemmal stability [3]. Mutations in *SGCA*, *SGCB*, *SGCD*, or *SGCG* can result in non-assembly of the SGC and, therefore, the absence of all four sarcoglycans from muscle of affected patients [3,4]. There are only a handful of low frequency founder alleles in human populations responsible for sarcoglycanopathies [5]; thus they are most commonly caused by mutations in compound heterozygosity [6].

In the domestic dog (*Canis familiaris*), selective breeding practices encourage pairing of recessive alleles inherited identical by descent (IBD). Accordingly, dogs have an abundance of recessive disorders [7], including muscular dystrophies [8,9]. Most canine muscular dystrophies are associated with dystrophin deficiency, and founder alleles have been identified in several breeds [10,11]. Recently, two independent mutations causing dystrophinopathy were described in Cavalier King Charles spaniels [12,13].

The first report of LGMD associated with sarcoglycan deficiency in dogs involved three breeds: Chihuahua, Cocker spaniel, and a seven-month-old male Boston terrier from Colorado (Case 1), but mutations were not identified [8]. Four years later, sarcoglycanopathy was described again in an unrelated four-month-old male Boston terrier from Iowa [14] (Case 2). All dogs affected with sarcoglycanopathy had a clinical dystrophic phenotype including muscle wasting, gait abnormalities, enlarged tongue, dysphagia, and extremely elevated serum creatine kinase (CK) activities [8,14].

Pathologic features were consistent with dystrophy, having myofiber degeneration, regeneration, and calcific deposits [8,14]. Affected dogs lacked muscle α -, β -, and γ -sarcoglycans, confirmed by both western blotting and immunohistochemistry [8,14]. At the time of evaluation, an antibody reactive with canine δ -sarcoglycan was unavailable.

Here, we describe a sarcoglycanopathy in a third family of Boston terriers from Arkansas in which two puppies (Cases 3 and 4) from the same kennel but different litters displayed clinical signs of LGMD, pathological changes consistent with a dystrophic phenotype and immunohistochemical confirmation of absent or decreased sarcoglycans. To identify the genetic basis for LGMD in the Boston terrier breed, we performed whole exome sequencing (WES) of Cases 1 and 3 and related dogs. Evaluation of the sarcoglycan genes revealed, to our surprise, two private deletions in *SGCD*: a 2 bp deletion in exon 6 and a 19.4 kb deletion encompassing exons 7 and 8. Both cause a lack of SGCD, resulting in LGMD2F.

Materials and Methods

Animals

Clinical details of Case 1 were previously published [8]. Biological samples from Case 2 were not available. Female Boston terriers, ages 12 months and 5 months, and from the same breeder in Arkansas (Cases 3 and 4), were evaluated for a chronic history of progressive dysphagia, lack of appetite, drooling, muscle wasting, and greatly enlarged tongues. Both dogs were examined by the same veterinarian in a clinical setting.

DNA was extracted from diagnostic muscle biopsies of Cases 1 and 3 and whole blood of unaffected relatives of Cases 3 and 4 using the DNeasy extraction kit (Qiagen, Hilden, Germany). Muscle for isolation of DNA was unavailable from Case 4. Whole blood samples or buccal swabs from unrelated, healthy Boston terriers were recruited, and DNA was isolated following the Genra PureGene protocol (Qiagen, Hilden, Germany) or the MagJet Genomic DNA purification kit (ThermoFisher Scientific, Waltham, USA). Genomic DNAs from unaffected dogs from multiple breeds were available from DNA archives at Clemson University and CAG GmbH.

The dogs in this study were examined and tissues collected in a clinical practice setting with the written consent of their owners. Studies on tissue biopsies and blood samples were approved by the Institutional Animal Care and Use Committees (IACUC) of Clemson University, the University of California San Diego, the University of Iowa, and the Animal Experiment Board in Finland (ESAVI/7482/04.10.07/2015), as well as the Baden-Württemberg veterinary office at the Landratsamt Tübingen Abt. 32: Veterinärwesen und Lebensmittelüberwachung, Tübingen, Germany (Registriernummer: DE 08 416 1038 21).

Histology and Immunofluorescence

Muscle specimens from Case 1 were previously obtained as biopsies and archived at -80°C at the Comparative Neuromuscular Laboratory, University of California San Diego (CNL). Specimens from limb muscles, heart, and tongue of Cases 3 and 4 were collected by a veterinarian following humane euthanasia at 1 year of age. Muscles were

either refrigerated or immersion fixed in buffered formalin and shipped to the CNL. Cryosections from all muscle specimens were processed by a standard panel of histochemical stains and reactions [15].

Antibodies used for immunofluorescence were rabbit antibodies R98 anti- α -sarcoglycan [16], R214 anti- δ -sarcoglycan [17], I1H6 anti- α -dystroglycan [18], R256 anti-sarcospan [19]; mouse antibodies 5B1 anti- β -sarcoglycan [19], 21B5 anti- γ -sarcoglycan [19], AP83 anti- β -dystroglycan [18], anti-dystrophin (AbCam), anti-collagen VI (Fitzgerald), and anti-caveolin 3 (BD Transduction Laboratories); rat anti-perlecan (NeoMarkers).

For secondary immunofluorescence, tissues were blocked with 10% goat serum in phosphate-buffered saline, incubated in primary antibody overnight, washed, incubated in Alexa Fluor 488, 594, or 647-conjugated anti- rat, rabbit, or mouse antibodies (Life Technologies), respectively, and mounted using ProLong Gold mounting media (Life Technologies). For α -dystroglycan, sarcospan, δ -sarcoglycan, and γ -sarcoglycan staining, tissues were fixed in 2% paraformaldehyde, followed by incubations in 100 mM glycine and 0.05% SDS prior to processing as described above. Images were acquired using a VS120-S5-FL slide scanner microscope (Olympus) with VS-ASW software.

Parentage Testing

The Canine Genotypes Panel 1.1 (ThermoFisher Scientific) was used to verify parentage of the experimental dogs. Samples were amplified according to manufacturer's instructions and separated and detected on an ABI 3730XL (Applied Biosystems,

ThermoFisher Scientific). GeneMarker (Softgenetics, State College, PA, USA) was used to assign peaks and determine genotypes according to ISAG nomenclature.

Whole exome sequencing

DNA from five Boston terriers (Case 1, Case 3, and three unaffected relatives of Cases 3 and 4) was used for WES performed at CeGaT GmbH (Tübingen, Germany). Genomic DNA (1 µg) from each sample was mechanically sheared to approximately 180–250 bp fragments using a Covaris LE220 Ultrasonicator (Woburn, MA, USA). Fragment sizes were assessed for quality control purposes (Fragment Analyzer, Advanced Analytical Technologies Inc.), and the Agilent SureSelect XT Canine All Exon kit (Santa Clara, CA, USA) supplied the 120-mer biotinylated RNA bases with which the fragment library was hybridized. Magnetic streptavidin beads were used for purification according to the manufacturer's protocol (Agilent). After amplification of library DNA, adaptors and barcodes for sequencing were added (Illumina), and equimolar amounts of each sample were pooled. Both lanes of a Rapid Flowcell were used to sequence the pool on an Illumina HiSeq2500, generating 2x100bp paired-end sequences, resulting in approximately 6 GB per sample. Illumina bcl2fastq 1.8.2 was used to demultiplex sequencing data, skewer 0.1.116 was used to trim sequencing adapters, and the Burrows-Wheeler Aligner (bwa 0.7.2-r351) was used to map the sequences to the canine genome (CanFam3.1). Samtools 0.1.18 and internal software were used to remove PCR duplicates and low-quality alignments. bcftools (0.1.17) and varscan (2.3.5) and internal software

were used to call variants, and a single Variant Call Format (VCF) file was generated for each sample using internal software.

IGV [20] and Genome Browse (Golden Helix [21,22] Inc., USA) were used to visualize data, and the Ensembl dbSNP (Can Fam3.1 version) and whole genome sequences (Clemson) were used to exclude variants.

Variant characterization and genotyping

2 bp deletion

The 2 bp deletion identified in Case 3 was verified by Sanger sequencing, using primers designed to amplify *SGCD* exon 6 (Additional File 1: Table S1). The deletion disrupts a *BcoDI* restriction enzyme site, yielding a 406 bp and a 347 bp product, representing mutant and wild-type alleles respectively. Unrelated Boston terriers and dogs from other breeds were genotyped using either restriction digest or Sanger sequencing.

19.4 kb deletion

To define the break points of the micro-deletion encompassing *SGCD* exons 7 and 8 and the 3' intergenic sequence, primers were designed in flanking sequences (Additional File 1: Table S1). For genotyping, primer pairs were designed within the deletion to amplify only wild type alleles, as well as flanking the deletion for amplification of the mutant allele. Primer pairs were multiplexed for amplification using Phire Hot Start II DNA polymerase (ThermoFisher) and products were resolved by gel

electrophoresis. Products were initially verified via Sanger sequencing. The multiplex PCR was used to test unrelated Boston terriers and dogs of other breeds.

Results

Clinical findings

Muscle wasting, dysphagia, exercise intolerance, lethargy, and failure to thrive were accompanied by progressive gait abnormalities including a short, stilted gait in Cases 3 (Fig. 1a) and 4. While there was no clinical indication of cardiomyopathy, specific evaluations for heart disease by a veterinary cardiologist were not performed. Clinical chemistry included markedly elevated activities of serum alanine aminotransferase (ALT, 900IU/L, reference range 10-110 IU/L), aspartate aminotransferase (AST, 920 IU/L, reference range 16-50 IU/L), and creatine kinase (CK, >10,000 IU/L, reference range 50-275 IU/L). Progression of clinical signs necessitated euthanasia at approximately 1 year of age for Case 3 and 5 months of age for Case 4.

Histology and immunofluorescence

A dystrophic phenotype including degeneration, regeneration, and calcific deposits was evident in the skeletal muscle (Figs. 1b) and tongue (Fig 1c and d). Heart muscle was histologically normal (left ventricle, not shown) from Cases 3 and 4. Immunofluorescence staining of muscle cryosections showed markedly reduced or absent localization of α -, β -, γ -, and δ -sarcoglycans and sarcospan in Cases 3 and 4 (Fig. 2).

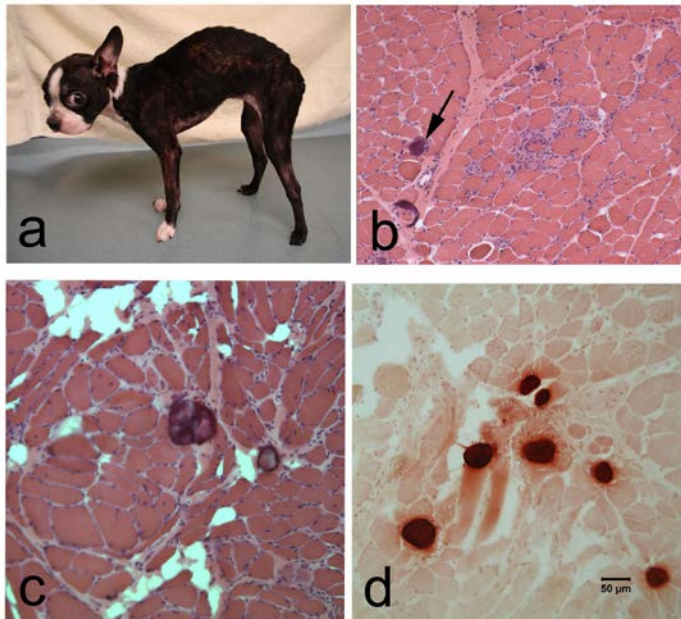


Figure 1. Histopathology of muscle biopsies from a female Boston terrier affected with sarcoglycanopathy (Case 3). A hunch back stance was evident in the dog (a). H&E stained cryosections from a representative limb muscle (b) showed degenerative changes and calcific deposits (black arrow). Similar degenerative changes and calcific deposits were observed in the tongue (c). The calcific deposits in the tongue were highlighted bright orange using the alizarin stain for calcium (d).

In contrast, staining for localization of α - and β -dystroglycans, dystrophin, caveolin 3, and perlecan was similar to control muscle (Fig. 3). Staining for collagen VI was increased in the endomysium compared to the control tissue, consistent with endomysial fibrosis. Results of histology, immunofluorescence staining, and western blotting of Case 1 were described previously [8]. Staining for localization of δ -sarcoglycan in Case 1 was performed on archived muscle cryosections and was similarly absent (not shown).

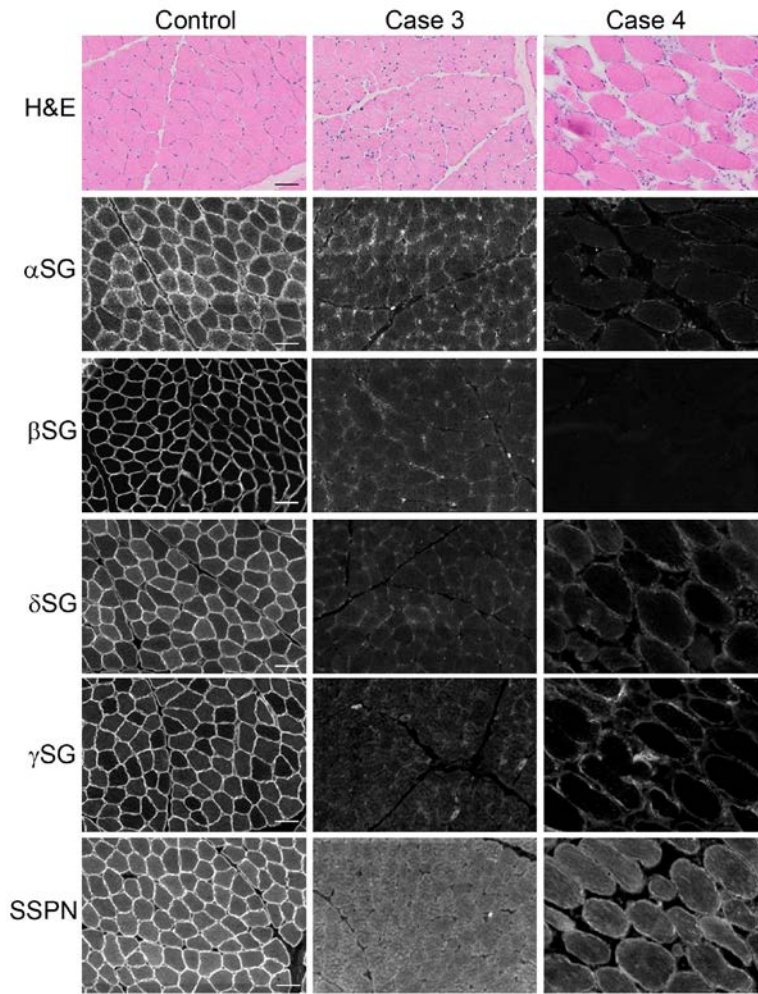


Figure 2. Loss of SGC staining in Cases 3 and 4. Representative H&E and immunofluorescence of cryosections from muscle of Cases 3 and 4, as well as control dog muscle. In control muscle, antibodies to the SGC (α -, β -, δ -, γ -sarcoglycans: α SG, β SG, δ SG, γ SG), as well as sarcospan (SSPN), localize to the sarcolemma of the muscle fibers. Staining from each of these antibodies is reduced in muscle from Cases 3 and 4.

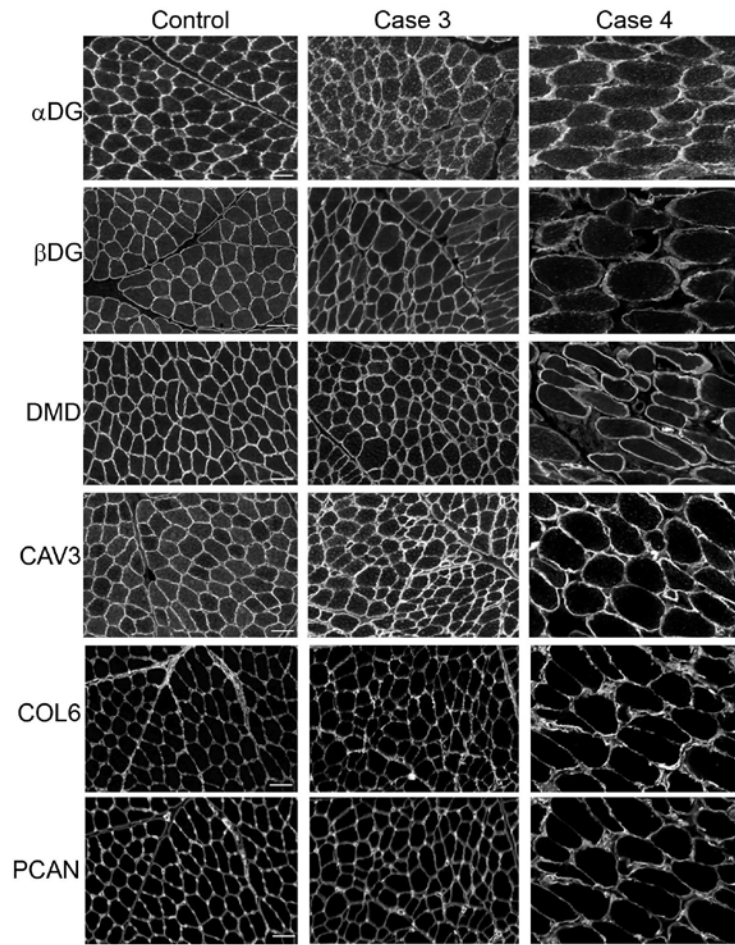


Figure 3. Representative immunofluorescence of cryosections from muscle of Cases 3 and 4, and control dog muscle. Staining of α -dystroglycan (α DG), β -dystroglycan (β DG), dystrophin (DMD), caveolin 3 (CAV3), collagen VI (COL6), and perlecan (PCAN) in Cases 3 and 4. Antibodies to α - and β -dystroglycans, dystrophin, caveolin 3, and perlecan demonstrate sarcolemmal localization and intensity that is comparable to control tissue. An antibody to collagen VI shows increased localization to the endomysium compared to the control tissue, consistent with endomysial fibrosis.

Parentage testing

Parentage testing was performed to determine relationships between Case 3 and three other dogs obtained from the same breeder. One relative was confirmed to be the dam of Case 3 and is referred to hereafter as the obligate carrier. The test excluded the

remaining two dogs from being full siblings of Case 3 or progeny of the obligate carrier. Their relationship to the other dogs or to one another could not be determined.

Variant identification from WES

Disruption of any one of the sarcoglycans results in reduced immunostaining of the entire SGC, both in LGMD patients and in animal models of sarcoglycanopathy [23,24,25]. Therefore, genetic sequencing was necessary to identify the defective sarcoglycan gene.

2 bp deletion

The candidate genes (*SGCA*, *SGCB*, *SGCD*, *SGCG*) were sequenced to an approximate depth of 30X. For each gene, we manually screened the VCF file in IGV for variants fitting a pattern of inheritance consistent with a rare recessive allele. We expected both affected dogs to have inherited the causal mutation IBD from a common ancestor; therefore, we searched for variants homozygous in Cases 1 and 3, heterozygous in the obligate carrier, and heterozygous or homozygous wild-type in the two relatives. No variants fit these criteria.

Because there was no known relationship between Cases 1 and 3, we considered that they may have different genetic causes for LGMD. Thus, we excluded Case 1 and searched again for the same pattern. Only one variant fit the pattern: a 2 bp deletion in exon 6 of *SGCD* (Figure 4). We validated the deletion through Sanger sequencing and determined that, in addition to the obligate carrier, one relative was heterozygous. The deletion predicts the substitution of an aspartate for a glutamate (E178D) and creates a

frameshift, leading to a premature stop codon two amino acids later (P180X) (Figure 4). We genotyped 199 Boston terriers and 127 dogs from 33 other breeds; none possessed the deletion.

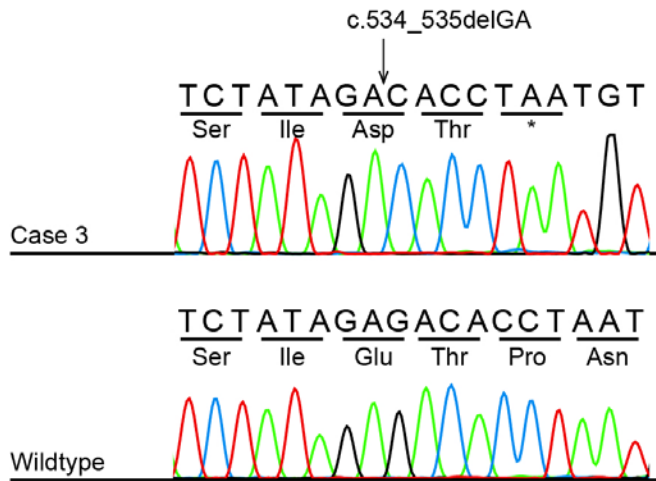


Figure 4. Electropherogram showing the 2 bp *SGCD* deletion in Case 3. The top panel shows the sequence from Case 3, while the lower panel shows the sequence from a healthy non-related Boston terrier. The *SGCD* c.534_535delGA mutation leads to a frameshift and a premature stop codon two amino acids later.

19.4 kb deletion

Using BAM files, we reexamined each candidate gene for variants homozygous in Case 1 and absent from the other Boston terriers. This approach revealed a complete absence of reads from the final two exons of *SGCD* (7 and 8) in Case 1, which was not apparent from the VCF file. No other variants fit the pattern. We hypothesized that the absence of reads represented a microdeletion and designed three primer pairs flanking exons 6, 7, and 8. PCR amplification yielded a product for exon 6 but not exons 7 or 8 in Case 1, providing further support for the presence of a deletion.

It was not possible to characterize the deletion directly from WES because intergenic and intronic sequences are minimized. Sequence coverage indicated that the deletion was between *SGCD* exon 6 and *TMD4*. Furthermore, sporadic intronic and intergenic fragments were present 5' of exon 7, beginning at chr4:53282570, and in the 3' UTR, beginning at chr4:53261359, suggesting a maximum deletion size of 21,211 bp. Primer pairs flanking this estimated deletion size yielded large products (~3-5 kb), indicating a deletion approximately 2 kb smaller than suggested by WES. Sanger sequencing of the breakpoint revealed a substitution (chr4:53262018-53262020, ATG>CC), followed by 9 bp that were unchanged before a deletion of 19,403 bp (chr4:53262030-53281432) (Figure 5). We genotyped 201 Boston terriers and 91 dogs of 19 other breeds and did not find any carriers.

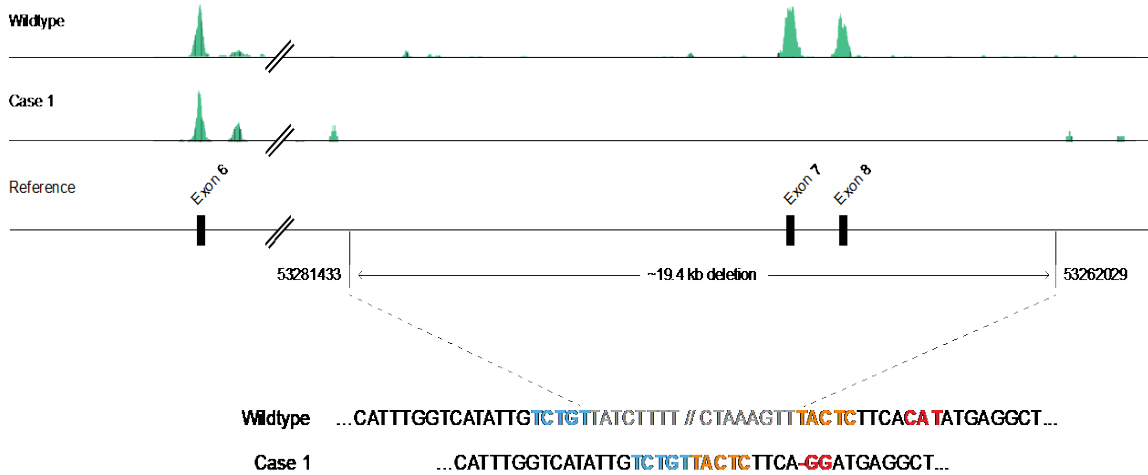


Figure 5. Schematic and sequence showing the breakpoints of the 19,403 bp *SGCD* deletion in Case 1. Note that *SGCD* is annotated on the minus strand. Whole exome sequence from a healthy dog and Case 1 are aligned to the reference genome, visualized in Golden Helix GenomeBrowse [21,22]. Case 1 has no coverage of exons 7 and 8 and flanking regions. Sequence of the wild type and Case 1 alleles show the precise breakpoints. Nucleotides 5' and 3' of the breakpoint are in bold blue and orange typeface, respectively. A substitution (chr4:53262020-53262018,CAT>GG) is found 9 bp downstream of the microdeletion and is shown in bold red typeface

Discussion

Sarcoglycanopathies in humans are rare genetic disorders, with an incidence of one in every 178,000 human births [26]. To date, only small animal models are available for study: gene targeted mouse models for α -, β -, δ -, and γ -sarcoglycanopathy [27] and a spontaneous hamster model for δ -sarcoglycanopathy [28,29]. Here, we have demonstrated that a naturally occurring muscular dystrophy in a Boston terrier family is a sarcoglycanopathy, consistent with two previously published case reports in the breed. Given that cases have been described in three Boston terrier families, we expected a single recessive allele, present at a very low frequency within the breed, to underlie all cases. Instead, we uncovered independent mutations in the two families studied herein. Unfortunately DNA from Case 2 [14] was not available to determine whether this dog shared one of the mutations described herein, a different mutation in *SGCD*, or a pathogenic variant in another gene.

Both families possessed mutations in *SGCD*, which encodes δ -sarcoglycan. Canine *SGCD* is located on CFA 4 and organized into eight exons that form a 1,297 bp mRNA transcript [30]. Human (XP_016865213.1) and dog (XP_013968526.1) amino acid sequences share 98% identity. Despite being the largest of the sarcoglycan genes, *SGCD* least commonly causes sarcoglycanopathy, with the majority of human cases attributed to changes in *SGCA* [31]. Thus, it is not only surprising that the Boston terriers had independent mutations causing sarcoglycanopathy, but that both had pathogenic alleles of *SGCD*. Curiously, the only other naturally occurring model of a sarcoglycanopathy, the Syrian hamster, also harbors an *SGCD* deletion [29].

Mutations of *SGCD* cause LGMD2F, and although clinical presentation is largely similar among the four sarcoglycanopathies, this is the only subtype not consistently characterized by concomitant cardiomyopathy [1]. The absence of heart involvement in Boston terriers is consistent with this classification; however, because the affected dogs were euthanized at an early age it is unknown if muscle degeneration would have progressed to involve the heart.

Immunohistochemistry illustrated a lack of the SGC in both cases, but provides no indication as to whether *SGCD* is abnormal or absent altogether. Due to limited sample availability, collected tissues were prioritized for histopathological analysis and genomic DNA sequencing. RNAs were thus unavailable to investigate the consequence of the deletions on *SGCD* transcripts. The 2 bp pair deletion predicts a premature stop codon in exon 6, possibly causing nonsense-mediated decay. The 19.4 kb micro-deletion eliminates the last two exons of *SGCD*; the complete loss of an *SGCD* exon is rare [5]. It is hypothesized that exon 6 would splice to one or more cryptic sites, triggering either nonsense-mediated decay and/or the production of mutant protein. It is likely that mutant *SGCD* would cause assembly of the SGC to fail, resulting in LGMD [3,32].

WES is a cost-effective method for the sequencing of multiple family members and has been used successfully to identify LGMD mutations in humans [33]. It was an advantageous choice over transcriptome sequencing in this study because *SGCD* transcripts would have been absent in Case 3 and possibly Case 1 as well, necessitating additional sequencing of *SGCD* to identify the causative mutations. In dogs, WES has led to the identification of alleles underlying progressive retinal atrophy, primary angle

closure glaucoma, and nemaline rod myopathy using small numbers of related cases [34,35,36,37,38] but is not ideal for detecting intergenic deletions or genomic rearrangements [39]. The development of improved WES enrichment kits for dogs [39,40] will facilitate future detection of disease variants in canine models.

Conclusion

The identification of canine models of disease holds promise for new advances in the understanding and treatment of analogous human diseases. For example, the well-characterized Golden retriever model of Duchenne muscular dystrophy (DMD) has proven to be an invaluable resource for gene therapy and other trials [41,42]. Here we have clinically and genetically characterized the first large animal model of sarcoglycanopathy.

Abbreviations

bp: base pair; **CAG**: Center for Animal Genetics, CAG GmbH; **CNL**: Comparative Neuromuscular Laboratory, University of California San Diego; **DMD**: Duchenne muscular dystrophy; **DNA**: deoxyribonucleic acid; **H&E**: hematoxylin and eosin; **IBD**: identical by descent; **IGV**: Integrative Genomics Viewer; **LGMD**: limb-girdle muscular dystrophy; **PCR**: polymerase chain reaction; **SGC**: sarcoglycan complex; **SGCA**: α -sarcoglycan; **SGCB**: β -sarcoglycan; **SGCD**: δ -sarcoglycan; **SGCE**: ϵ -sarcoglycan; **SGCG**: γ -sarcoglycan; **SGCZ**: ζ -sarcoglycan; **UTR**: untranslated region; **VCF**: variant call format; **WES**: Whole exome sequencing

Ethics approval

All dogs in this study were evaluated in a clinical veterinary practice by licensed veterinarians. The dogs in this study were examined and tissue biopsies collected with the written consent of their owners. Tissue studies were performed using protocols approved by the Institutional Animal Care and Use Committees (IACUC) of Clemson University, the University of California San Diego, and the University of Iowa, and the Animal Experiment Board in Finland (ESAVI/7482/04.10.07/2015), as well as the Baden-Württemberg veterinary office at the Landratsamt Tübingen Abt. 32: Veterinärwesen und Lebensmittelüberwachung, Tübingen, Germany (Registriernummer: DE 08 416 1038 21).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

MC is employed by CAG GmbH that performs canine DNA testing on a commercial basis. HL is a co-founder of Genoscooper Laboratories Oy that offers canine DNA testing on a commercial basis. All other authors declare that they have no competing interests.

Funding

Research reported in this publication was supported in part by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health

under Award Number R15AR062868 (LAC) and the Clemson University Calhoun Honors College (AGD). HL was supported by the Jane and Aatos Erkkö Foundation and the Academy of Finland. ES was supported by the Jenny and Antti Wihuri Foundation. This work was supported in part by a Paul D. Wellstone Muscular Dystrophy Cooperative Research Center grant (1U54NS053672 to KPC). KPC is an investigator of the Howard Hughes Medical Institute.

Authors' contributions

MC, JE, LAC and GDS designed the research. MC, JE, LTG, AD, LG, JL, KC, AS-M, ES, and MH performed the experiments. MC, JE, KC, ES, MH, HL, LAC, and GDS analyzed the data. MC, JE, LAC, and GDS wrote the manuscript. All authors participated in editing the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We are thankful to all dog owners, breeders, and clinicians who have contributed samples to this study. The authors also thank Florian Battke, Sally Prouty and Hammon Humphries for technical assistance.

References

1. Nigro V, Savarese M. Genetic basis of limb-girdle muscular dystrophies: the 2014 update. *Acta Myologica*. 2014;33:1-12.

2. Magri F, Brajkovic S, Brusa R, Comi GP, Govoni A. Revised genetic classification of Limb Girdle Muscular Dystrophies. *Curr Mol Med*. 2014;14:934-943.
3. Tarakci H, Berger J. The sarcoglycan complex in skeletal muscle. *Front Biosci (Landmark Ed)*. 2016;21:744-756.
4. Shi W, Chen Z, Schottenfeld J, Stahl RC, Chan YM, Kunkel LM. Specific assembly pathway of sarcoglycans is dependent on beta- and delta-sarcoglycan. *Muscle Nerve*. 2004;29:409-419.
5. Trabelsi M, Kavian N, Daoud F, Commere V, Deburgrave N, Beugnet C, et al. Revised spectrum of mutations in sarcoglycanopathies. *Eur J Hum Genet*. 2008;16:793-803.
6. Pegoraro E, Hoffman EP. Limb-girdle muscular dystrophy overview. In: Pagon RA, Adam MP, Ardinger HH, et al., editors. *Gene Reviews*. Seattle: University of Washington; 2012.
7. Summers J, Diesel G, Asher L, McGreevy PD, Collins LM. Inherited defects in pedigree dogs. Part 2: Disorders that are not related to breed standards. *Vet J*. 2010;183:39-45.
8. Schatzberg SJ, Shelton GD. Newly identified neuromuscular disorders. *Vet Clin N Am Small Anim Pract*. 2004;34:1497-1520.
9. Shelton GD, Engvall E. Canine and feline models of human inherited muscle diseases. *Neuromuscular Disord*. 2005;15:127-138.

10. Kornegay JN, Bogan JR, Bogan DJ, Childers MK, Nghiem P, Xiao X, et al. Canine models of Duchenne muscular dystrophy and their use in therapeutic strategies. *Mamm Genome*. 2012;23:85-108.
11. Atencia-Fernandez S, Shiel RE, Mooney CT, Nolan CM. Muscular dystrophy in the Japanese Spitz: an inversion disrupts the DMD and PGR genes. *Anim Genet*. 2015;46:175-184.
12. Walmsley GL, Arechavala-Gomez V, Fernandez-Fuente M, Burke MM, Nagel N, Holder A, et al. A Duchenne muscular dystrophy gene hot spot in dystrophin-deficient Cavalier King Charles Spaniels is amenable to exon 51 skipping. *PLoS ONE*. 2010;13:e8647. doi: 10.1371/journal.pone.0008647.
13. Nghiem PP, Bello L, Balog-Alvarez C, Lopez SM, Bettis A, Barnett H, et al. Whole genome sequencing reveals a 7 base-pair deletion in DMD exon 42 in a dog with muscular dystrophy. *Mamm Genome*. 2016;28:106-113.
14. Deitz K, Morrison JA, Kline K, Guo L, Shelton G. Sarcoglycan-deficient muscular dystrophy in a Boston Terrier. *J Vet Intern Med*. 2008;22:476-480.
15. Dubowitz V, Sewry CA, Oldfors A. *Muscle biopsy: a practical approach*. 4th ed. Elsevier Health Sciences; 2013.
16. Roberds SL, Anderson RD, Ibraghimov-Beskrovnaya O, Campbell KP. Primary structure and muscle-specific expression of the 50-KDa dystrophin-associated glycoprotein (adhelin). *J Biol Chem*. 1993;268:23739-23742.

17. Duclos F, Straub V, Moore SA, Venzke DP, Hrstka RF, Crosbie RH, et al. Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *J Cell Biol.* 1998;142:1461-1471.
18. Ervasti JM, Campbell KP. Membrane stabilization of the dystrophin-glycoprotein complex. *Cell.* 1991;66:1121-1131.
19. Durbeej M, Cohn RD, Hrstka RF, Moore SA, Allamand V, Davidson BL, et al. Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. *Mol Cell.* 2000;5:141-151.
20. Broad Institute. IGV Integrative Genomics Viewer. [Online]. Available from: <http://software.broadinstitute.org/software/igv/>.
21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078-2079.
22. Bozeman MT. Golden Helix GenomeBrowse ® visualization tool (Version 2.1) [Software]. [Online]. Available from: <http://www.goldenhelix.com>.
23. Barresi R, Moore SA, Stolle CA, Mendell JR, Campbell KP. Expression of gamma-sarcoglycan in smooth muscle and its interaction with the smooth muscle sarcoglycan-sarcospan complex. *J Biol Chem.* 2000;275:38554-38560.
24. Crosbie RH, Lebakken CS, Holt KH, Venzke DP, Straub V, Lee JC, et al. Membrane targeting and stabilization of sarcospan is mediated by the sarcoglycan subcomplex. *J Cell Biol.* 1999;145:153-165.
25. Holt KH, Campbell KP. Assembly of the sarcoglycan complex. Insights for muscular dystrophy. *J Biol Chem.* 1998;273:34667-34670.

26. Fanin M, Duggan DJ, Mostacciuolo ML, Martinello F, Freda MP, Sorarù G, et al. Genetic epidemiology of muscular dystrophies resulting from sarcoglycan gene mutations. *J Med Genet.* 1997;34:973–977.
27. Ng R, Banks GB, Hall JK, Muir LA, Ramos JN, Wicki J, et al. Animal models of muscular dystrophy. *Prog Mol Biol Transl Sci.* 2012;105:83-111.
28. Straub V, Duclos F, Venzke DP, Lee JC, Cutshall S, Leveille CJ, et al. Molecular pathogenesis of muscle degeneration in the δ -sarcoglycan-deficient hamster. *Am J Pathol.* 1998;153:1623-1630.
29. Nigro V, Okazaki Y, Belsito A, Piluso G, Matsuda Y, Politano L, et al. Identification of the Syrian hamster cardiomyopathy gene. *Hum Mol Genet.* 1997;6:601-607.
30. Ensembl.org. [Online]. [cited 2017 May 28. Available from:
http://www.ensembl.org/Canis_familiaris/Transcript/Summary?db=core;g=ENSCAFG00000017592;r=4:53267777-53642810;t=ENSCAFT000000027892.
31. Duggan DJ, Gorospe JR, Fanin M, Hoffman EP, Ozawa E, Angelini C. Mutations in the sarcoglycan genes in patients with myopathy. *New Engl J Med.* 1997;336:618-624.
32. Moreira ES, Vainzof M, Marie S, Nigro V, Zatz M, Passos-Bueno MR. A first missense mutation in the delta sarcoglycan gene associated with a severe phenotype and frequency of limb-girdle muscular dystrophy type 2F (LGMD2F) in Brazilian sarcoglycanopathies. *J Med Genet.* 1998;35:951–953.

33. Ghaoui R, Cooper ST, Lek M, Jones K, Corbett A, Reddel SW, et al. Use of whole-exome sequencing for diagnosis of limb-girdle muscular dystrophy: outcomes and lessons learned. *JAMA Neurol.* 2015;72:1424-1432.
34. Ahonen S, Arumilli, Lohi H. A CNGB1 frameshift mutation in Papillon and Phalene dogs with progressive retinal atrophy. *PLoS ONE.* 2013;8:e72122. doi:10.1371/journal.pone.0072122.
35. Ahram DF, Collin RWJ, Kuehn MH, Grozdanic SD, Henkes A, Kecova H. Variants in nebulin (NEB) are linked to the development of familial primary angle closure glaucoma in Basset Hounds. *PLoS ONE.* 2015;10:e0126660. doi:10.1371/journal.pone.0126660.
36. Evans JM, Cox ML, Huska , Li F, Gaitero L, Guo L, et al. Exome sequencing reveals a nebulin nonsense mutation in a dog model of nemaline myopathy. *Mamm Genome.* 2016;27:495-502.
37. Kropatsch R, Akkad DA, Frank M, Rosenhagen C, Altmüller J, Nürnberg P, et al. A large deletion in RPGR causes XLPPRA in Weimaraner dogs. *Canine Genet Epidemiol.* 2016;3. doi:10.1186/s40575-016-0037
38. Forman OP, Pettitt L, Komáromy AM, Bedford P, Mellersh C. A novel genome-wide association study approach using genotyping by exome sequencing leads to the identification of a primary open angle glaucoma associated inversion disrupting ADAMTS17. *PLoS ONE.* 2015;10:e0143546. doi:10.1371/journal.pone.0143546.

39. Broeckx BJ, Coopman F, Verhoeven GE, Bavegems V, De Keulenaer S, De Meester E, et al. Development and performance of a targeted whole exome sequencing enrichment kit for the dog (*Canis Familiaris* Build 3.1). *Sci Rep.* 2014;4:5597.
40. Broeckx BJ, Hitte C, Coopman F, Verhoeven GEC, De Keulenaer S, De Meester E, et al. Improved canine exome designs, featuring ncRNAs and increased coverage of protein coding genes. *Sci Rep.* 2015;5:12810.
41. Duan D. Duchenne muscular dystrophy gene therapy in the canine model. *Hum Gene Ther Clin Dev.* 2015;26:57-69.
42. Yu X, Bao B, Echigoya Y, Yokota T. Dystrophin-deficient large animal models: translational research and exon skipping. *Am J Transl Res.* 2015;7:1314-1331.

CHAPTER FOUR

BEYOND THE MHC: A CANINE MODEL OF DERMATOMYOSITIS SHOWS A COMPLEX PATTERN OF GENETIC RISK INVOLVING NOVEL LOCI

Jacquelyn M. Evans¹, Rooksana E. Noorai^{1,2}, Kate L. Tsai¹, Alison N. Starr-Moss¹, Cody M. Hill¹, Kendall J. Anderson¹, Thomas R. Famula³, Leigh Anne Clark^{1*}

¹Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina, United States of America

²Genomics and Computational Laboratory, Clemson University, Clemson, South Carolina, United States of America

³Department of Animal Science, University of California, Davis, California, United States of America

*Corresponding author

E-mail: lclark4@clemson.edu (LAC)

Published – PLOS Genetics

Copyright: © 2017 Evans et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Abstract

Juvenile dermatomyositis (JDM) is a chronic inflammatory myopathy and vasculopathy driven by genetic and environmental influences. Here, we investigated the genetic underpinnings of an analogous, spontaneous disease of dogs also termed dermatomyositis (DMS). As in JDM, we observed a significant association with a haplotype of the major histocompatibility complex (MHC) (*DLA-DRB1*002:01/-DQA1*009:01/-DQB1*001:01*), particularly in homozygosity (P -val=0.0001). However, the high incidence of the haplotype among healthy dogs indicated that additional genetic risk factors are likely involved in disease progression. We conducted genome-wide association studies in two modern breeds having common ancestry and detected strong associations with novel loci on canine chromosomes 10 (P -val= 2.3×10^{-12}) and 31 (P -val= 3.95×10^{-8}). Through whole genome resequencing, we identified primary candidate polymorphisms in conserved regions of *PAN2* (encoding p.Arg492Cys) and *MAP3K7CL* (c.383_392ACTCCACAAA>GACT) on chromosomes 10 and 31, respectively. Analyses of these polymorphisms and the MHC haplotypes revealed that nine of 27 genotypic combinations confer high or moderate probability of disease and explain 93% of cases studied. The pattern of disease risk across *PAN2* and *MAP3K7CL* genotypes provided clear evidence for a significant epistatic foundation for this disease, a risk further impacted by MHC haplotypes. We also observed a genotype-phenotype correlation wherein an earlier age of onset is correlated with an increased number of risk alleles at *PAN2* and *MAP3K7CL*. High frequencies of multiple genetic risk factors are unique to affected breeds and likely arose coincident with artificial selection for desirable

phenotypes. Described herein is the first three-locus association with a complex canine disease and two novel loci that provide targets for exploration in JDM and related immunological dysfunction.

Author summary

Juvenile dermatomyositis (JDM) is an autoimmune disease of the skin and muscle influenced by both genetic and environmental components. Although genes independent of the MHC are thought to contribute to disease pathogenesis, their identification has been complicated by a paucity of biological samples, disease heterogeneity, and genetically diverse subjects. A naturally occurring inflammatory disease of domestic dogs, also termed dermatomyositis (DMS), is analogous to JDM and is the only animal model available for genetic study. We first determined that, as in JDM, a particular MHC haplotype confers susceptibility to DMS. Capitalizing on the genetic isolation of dog breeds and extremely low MHC diversity within affected breeds, we used multibreed genome-wide association studies to identify two novel loci. Through genome resequencing and additional genotyping, we uncovered highly associated polymorphisms in conserved positions of *PAN2* and *MAP3K7CL*. Analysis of three-locus genotypes revealed uniquely high frequencies among affected breeds and nine allelic combinations that confer moderate or high risk for DMS. The pattern of disease probability illustrates the presence of gene-gene interaction, as well as an inverse correlation between age of onset and number of risk alleles. This study highlights the utility of canine models for

mapping susceptibility loci in complex diseases and detecting patterns of genetic interactions obscured in diverse human populations.

Introduction

Juvenile dermatomyositis (JDM) is an autoimmune vasculopathy that causes a characteristic skin rash (heliotrope rash across the eyelids and Gottron's papules on the bony prominences) and proximal muscle weakness [1]. It is the most frequently diagnosed childhood inflammatory myopathy, comprising 80% of all cases [1] and affecting 3.2 in every million children between the ages of 2 and 17 within the USA [2]. Prognosis is positively correlated with early diagnosis and swift treatment with corticosteroids and/or immunosuppressants [1,3]. While treatment of JDM is much improved overall, many children still suffer from chronic flare-ups [1].

Though the etiology is unknown, JDM is thought to be triggered by exposure to a virus or other environmental agent. Manlhiot et al. [4] reported that 71% of JDM patients had a clinical history consistent with infection prior to symptoms. Investigations into the class II major histocompatibility loci (MHC), *TNF*, and *IL1* identified several susceptibility and protective alleles, but their collective contribution to pathogenesis is poorly understood [5-8]. Recent genome-wide association studies (GWASs) to identify additional susceptibility loci in JDM confirmed a strong association with the MHC but failed to detect novel major risk factors, likely because of a paucity of biological samples and genetically heterogeneous populations [9,10].

In domestic dogs, an inflammatory vasculopathy of the skin and muscle, also termed dermatomyositis (DMS), is clinically, histologically, and immunologically similar to JDM and provides the only animal model available to study genetic risk factors [11-16]. The earliest clinical signs of DMS are crusting and scaling on the face, ears, tail tip, and/or across the bony prominences of the limbs and feet [17-19] (S1 Fig). Alopecia and more extensive skin lesions may develop over time, resulting in dermal scarring associated with erythema and mottled pigmentation [17,19]. Lesions persist for weeks to months, and may or may not chronically recur [17]. Muscle wasting manifests as atrophy of the head musculature; difficulty eating, drinking, and swallowing; and an atypical, high-stepping gait [17,19].

Similar to JDM, DMS is an immune-mediated disease [13,18,20] that typically develops following an environmental trigger, such as vaccination or viral infection, and is exacerbated by subsequent stressors like exposure to UV light [13,17,21,22]. Anecdotal reports indicate that rabies vaccination, parvovirus infection, owner surrender, or mistreatment commonly precede disease onset. Consistent with an environmental trigger, age at onset is variable with many cases occurring between seven weeks and six months of age, but others not developing until well into adulthood [17-19,23].

DMS is diagnosed almost exclusively in the genetically [24] and phenotypically similar collie and Shetland sheepdog breeds, suggesting the presence of a strong heritable component(s) arising from ancestors common to both breeds. A 1980s study of disease transmission in the collie eliminated simple Mendelian modes of inheritance [14]. In two test matings, an affected male collie sired litters from an affected collie and a healthy

Labrador retriever. All six collie puppies were affected with variable degrees of severity, while three of the four mixed breed puppies developed milder forms of DMS.

Retrospective pedigree analyses of the collie sire and dam showed a complete absence of affected ancestors [14].

The availability of a naturally-occurring canine model provides a new opportunity for the identification of genetic risk factors of JDM. The conserved genomic backgrounds of genetically isolated dog breeds have enabled detection of risk loci in complex diseases that are often obscured by heterogeneity within human cohorts [25-30]. Here, we evaluated class II MHC haplotypes, performed multibreed GWASs, and generated whole genome and transcriptome sequencing data to dissect the genetic basis of DMS. We uncovered common polymorphisms of the MHC and two novel loci that are strongly associated with DMS, as well as patterns of allelic inheritance that explain 93% of cases studied. A genetic test is now available to determine the likelihood of a dog developing DMS and to facilitate breeding decisions that avoid progeny having high-risk genotypes.

Results and Discussion

Association of a major histocompatibility complex haplotype

Given the involvement of MHC genes in JDM, we first determined alleles of the highly polymorphic canine MHC class II dog leukocyte antigen (DLA) genes: *DLA-DRB1*, *-DQAI*, and *-DQBI*. Two locus (*DLA-DRB1* and *-DQBI*) and three locus (*DLA-DRB1*, *-DQAI*, and *-DQBI*) haplotypes were first generated for 50 collies and 117 Shetland sheepdogs, respectively. Because all observed haplotypes contained a unique

DLA-DRB1 allele, the 355 remaining dogs were genotyped for this locus only and the haplotype was inferred (Table 1).

COLLIE					
<i>DLA-DRB1</i> */- <i>DQA1</i> */- <i>DQB1</i> *	Cases (2n=80)	Controls (2n=370)	OR	95% CI	P-value
<i>002:01/009:01/001:01</i>	74	353	0.59	0.23-1.56	0.40
<i>006:01/050:11/007:01</i>	0	3	-	-	-
<i>015:01/009:01/001:01</i>	6	14	2.06	0.77-5.54	0.23
SHETLAND SHEEPDOG					
<i>DLA-DRB1</i> */- <i>DQA1</i> */- <i>DQB1</i> *	Cases (2n=184)	Controls (2n=410)	OR	95% CI	P-value
<i>002:01/009:01/001:01</i> [†]	158	301	2.20	1.38-3.52	0.0010
<i>015:01/009:01/001:01</i>	1	6	0.37	0.04-3.08	0.4454
<i>023:01/003:01/005:01</i>	24	102	0.45	0.28-0.73	0.0011
Other haplotypes	1	1	-	-	-

Table 1. Frequency of *DLA-DRB1*-*DQA1*-*DQB1* haplotypes in collies and Shetland sheepdogs. The number of times each haplotype was observed is reported for cases and controls with odds ratios (OR), 95% confidence intervals, and Fisher's exact two-tailed *P*-values. Significant statistics are bolded.

[†]71 out of 92 cases and 109 out of 205 controls were homozygous for the *002:01/009:01/001:01* haplotype (OR=2.98, 95% CI=1.70-5.21, *P*-val=0.0001).

We observed remarkably low *DLA* diversity among collies, with only three haplotypes present in 225 collies worldwide. This lack of heterogeneity precluded detection of associations with DMS, as 91% of collies were homozygous for the haplotype *DLA-DRB1***002:01*-*DQA1***009:01*-*DQB***001:01*. In 297 Shetland sheepdogs, we identified two predominant haplotypes, *002:01/009:01/001:01* and *023:01/003:01/005:01*. The former was over-represented among cases (*P*-val=0.0010, OR=2.20), primarily because of increased homozygosity (*P*-val=0.0001, OR=2.98). We therefore conclude that *002:01/009:01/001:01* is a risk factor for DMS and that homozygosity confers increased susceptibility. Under the assumption that the causal

alleles derive from an ancestor common to both breeds, we extrapolate the observed DLA risk to collies. The high frequency of the DLA risk haplotype in both populations indicates that additional loci must influence disease progression.

Genome-wide association studies reveal signals on chromosomes 10 & 31

We conducted an independent GWAS for each breed, using a total of 97 cases (31 collies, 66 Shetland sheepdogs), 68 controls (23 collies, 45 Shetland sheepdogs), and 98,520 SNPs after filtering. In collies, a single signal ($P\text{-val}=1.47\times 10^{-8}$) composed of 17 SNPs at the centromeric end of chromosome 10 exceeded Bonferroni significance (Fig 1A). In Shetland sheepdogs, this association was replicated ($P\text{-val}=2.56\times 10^{-7}$), and a second signal ($P\text{-val}=1.83\times 10^{-9}$) composed of 11 SNPs surpassing Bonferroni significance was detected on chromosome 31 (Fig 1B). Both associations persisted in a combined breed analysis (chr10: $P\text{-val}=2.3\times 10^{-12}$, chr31: $P\text{-val}=3.95\times 10^{-8}$) (S1 Table, S2 Fig); although the breeds possessed a common haplotype on chromosome 31, the Shetland sheepdogs appeared to drive this association. No associated SNPs were detected near the MHC loci on chromosome 12, likely a result of high homogeneity in our cohort and poor SNP coverage on the array [25].

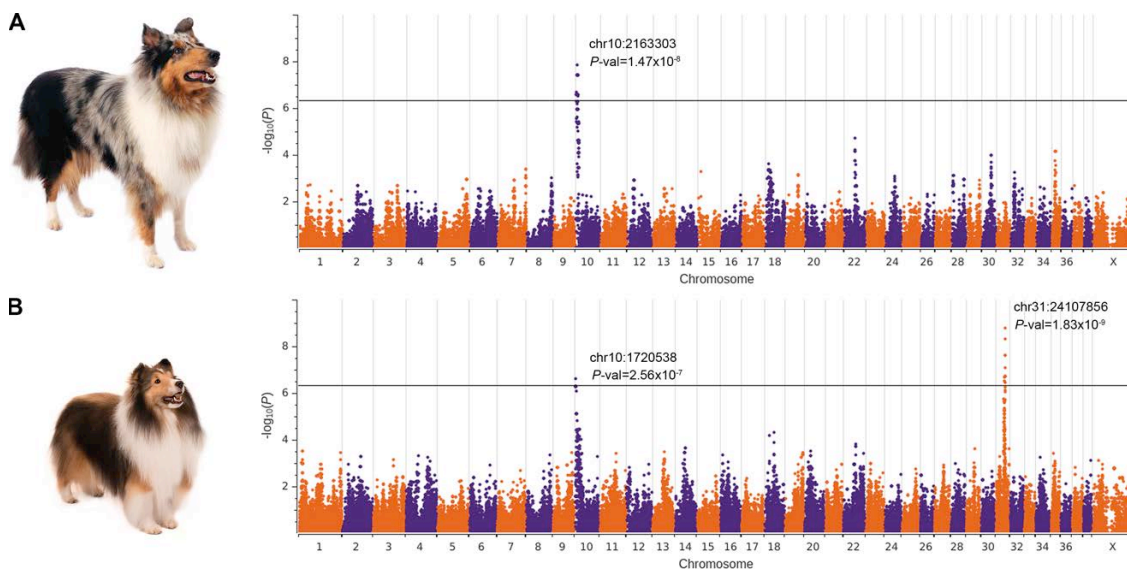


Figure 1. Manhattan plots of DMS GWASs in two breeds. (a) Collies: 31 cases and 23 controls. (b) Shetland sheepdogs: 66 cases and 45 controls. The $-\log_{10}P$ -values (y-axis) for 98,520 Illumina SNPs are plotted against chromosome position (x-axis). The threshold for Bonferroni significance is shown as a black horizontal line. The P -value and position (canFam3) of the lead SNPs are reported.

On chromosome 10, 97% of all affected dogs were homozygous or heterozygous for the risk alleles of the lead SNPs. On chromosome 31, 88% of affected Shetland sheepdogs, but only 39% of affected collies, shared the risk alleles of the lead SNPs. As neither locus appeared to be independently necessary for disease development, we surveyed the extent of regional linkage disequilibrium (LD) to demarcate candidate intervals of ~ 1.33 Mb on chromosome 10 (Fig 2A) and ~ 696 kb on chromosome 31 (Fig 2B), harboring ~ 65 and 6 genes, respectively. The large size of the chromosome 10 region is attributed to lower recombination rates near the centromere and a dearth of informative SNPs.

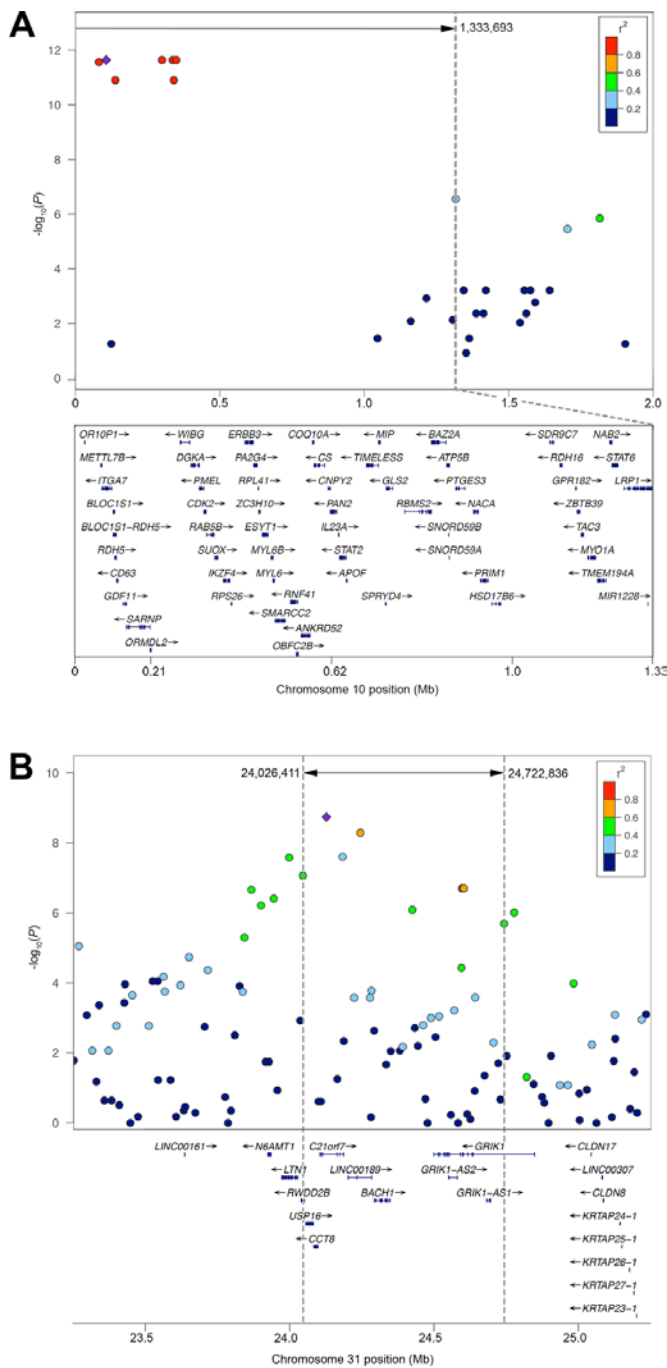


Figure 2. Regional plots depicting LD at two associated loci. Association results ($-\log_{10}P$) for (a) combined breed GWAS on chromosome 10 and (b) Shetland sheepdog GWAS on chromosome 31 are color-coded based on pairwise LD (r^2) with the lead SNPs (purple diamond). Dotted lines demarcate the candidate interval. Genes within the region (based on hg18) are shown below the plot with chromosome positions reported in canFam3. On chromosome 10, the gene track is zoomed in to represent the candidate interval only; 4 genes were omitted from the figure (*DNAJC14*, *LOC440104*, *MMP19*, and *SLC39A5*).

Identification of candidate variants in PAN2 and MAP3K7CL

Whole genome resequencing was performed for four affected dogs (three collies and one Shetland sheepdog) and two unaffected collies, resulting in 17X to 41X coverage. Variants were filtered for those lying within our delineated intervals (chr10:1-1,333,693; chr31:24,026,411-24,722,836) and following the inheritance pattern of the most significantly associated SNPs in the affected dogs (S2 and S3 Tables). Five intergenic and two intronic variants were unique to these breeds (*i.e.*, not present in the Boxer reference genome, dbSNP, or 27 whole genome sequences from 16 other breeds); however, most were in repetitive regions and none were in conserved positions. Thus, the pathogenic variants were likely to be common polymorphisms, so we next prioritized variants within predicted exons and splice sites of genes expressed in skin for further study.

To confirm exon/intron boundaries predicted by Ensembl 79 and establish expression of variants in affected tissue, we generated RNAseq data. We obtained a minimum of 89 million reads per tissue from active skin lesions of two affected dogs (one collie and one Shetland sheepdog) and skin from one unaffected Australian shepherd dog. All genes expressed in skin within the candidate regions were also expressed in affected tissues. Seventeen exonic variants were expressed; seven of these were nonsynonymous and evaluated using *in silico* programs [31-33] (Table 2).

Position	Gene	A1/A2	AA change	Poly-Phen2	PANTHER	SIFT	P-value
CHROMOSOME 10							Cases n=93 Controls n=63
95042	<i>ITGA7</i>	T/C	Val/Val	-	-	-	-
121835	<i>RDH5</i>	G/A	Ala/Ala	-	-	-	-
222809	<i>MMP19</i>	T/G	Gln/Pro	0	91	0.43	-
331135	<i>SUOX</i>	A/G	His/Arg	0	85	N/A	-
565958	<i>ANKRD52</i>	G/C	Ser/Cys	0.92	361	0.02	1.93X10⁻¹⁷
627760	<i>PAN2</i>	G/A	Arg/Cys	1	1628	0	1.93X10⁻¹⁷
1127082	<i>RDH16</i>	C/T	Ala/Ala	-	-	-	-
1228277	<i>NAB2</i>	G/C	Gly/Arg	0	85	N/A	-
1239562	<i>STAT6</i>	G/A	Thr/Met	0.98	361	0.05	1.52X10 ⁻⁴
1286150	<i>LRP1</i>	C/T	Asn/Asn	-	-	-	-
1333693	<i>LRP1</i>	G/A	Gln/Gln	-	-	-	-
CHROMOSOME 31							Cases n=63 Controls n=45
24068039	<i>CCT8</i>	T/C	Ala/Ala	-	-	-	-
24132273	<i>MAP3K7CL</i>	Indel/-	5' UTR	N/A	N/A	N/A	2.09X10⁻¹¹
24132343	<i>MAP3K7CL</i>	A/C	5' UTR	N/A	N/A	N/A	1
24292521	<i>BACH1</i>	A/G	Asn/Asp	0.007	324	0.21	-
24294659	<i>BACH1</i>	A/G	Glu/Glu	-	-	-	-
24295208	<i>BACH1</i>	A/G	Pro/Pro	-	-	-	-

Table 2. Expressed exonic changes on chromosomes 10 and 31. Alleles 1 (A1=minor allele) and 2 (A2=reference allele) are reported. Amino acid changes are based on the dog reference genome. P-values are reported for matched populations. Variants with PolyPhen2 scores ranging from 0.85-1, PANTHER preservation time (in millions of years) >450my, and SIFT scores ranging from 0-0.05 are considered deleterious/probably damaging. Variants more strongly associated with DMS than the most associated SNPs from the array are bolded.

We genotyped a subset of our mapping population for three nonsynonymous SNPs on chromosome 10 (*ANKRD52* g.565958G>C, *PAN2* g.627760G>A, and *STAT6* g.1239562G>A) that were predicted to be deleterious or probably damaging by more than one *in silico* program. The *ANKRD52* and *PAN2* variants were more strongly associated with DMS than the lead SNP. These variants were in perfect linkage disequilibrium with each other; however, *PAN2* g.627760G>A was assigned damaging scores with higher confidence by *in silico* programs (Table 2). We therefore focused further studies on

PAN2 g.627760G>A, encoding p.Arg492Cys (XP_531635.3), although *ANKRD52* cannot be excluded. On chromosome 31, we genotyped Shetland sheepdogs for a SNP (g.24132343A>C) and an indel (c.383_392ACTCCACAAA>GACT, XM_846337.4), both located in a 5' non-coding exon of *MAP3K7CL*. Only the indel was associated with DMS (Table 2, S3 Fig). In an expanded, combined population (132 affected and 390 unaffected collies and Shetland sheepdogs), both the *PAN2* ($P\text{-val}=2.08\times 10^{-35}$) and *MAP3K7CL* ($P\text{-val}=1.45\times 10^{-33}$) variants were highly associated with DMS (S4 Table).

PAN2 (or *USP52*) encodes the catalytic subunit of the poly(A) nuclease deadenylation complex (PAN2-PAN3) and is one of two exonucleases involved in mRNA degradation in eukaryotes [34,35]. Deadenylation plays a role in translational regulation of inflammatory response [36]. Independent of this function, PAN2 also stabilizes *HIF1A* transcripts via their 3'-UTR, which contain AU-rich elements (AREs), and may be involved in regulating other transcripts having AREs [37]. *HIF1A*, a key regulator of inflammation [38], and other ARE-containing transcripts, such as *IL-6* [39], are upregulated in JDM [40, 41]. *PAN2* is widely expressed and highly evolutionarily conserved [35]; human (NP_001120932.1) and dog (XP_013972628.1) amino acid sequences share 98% identity.

MAP3K7CL (also known as *TAK1L* or *C21orf7*) is a poorly studied kinase gene that is transcriptionally active in immunological tissues and expressed primarily in peripheral blood leukocytes [42,43]. Human (NP_001273546.1) and dog (XP_013965340.1) *MAP3K7CL* protein sequences share >90% identity. The transcription factors RUNX3 and EP300 bind the 5' non-coding exon of human

MAP3K7CL (UCSC Genome Browser ENCODE Transcription Factor ChIP-seq track).

In this exon, the c.383_392ACTCCACAAA>GACT indel causes the loss of six conserved base pairs, omitting a RUNX3 binding motif ($P\text{-val}=2.07\times 10^{-3}$ from TOMTOM [44]) (S3 Fig). RUNX3 has known roles in inflammatory response (*e.g.*, thymopoiesis [45,46] and the TGF- β signaling cascade [47]), and it has been directly implicated in a number of immune-related diseases [48-50]. Furthermore, SNPs disrupting RUNX binding motifs in target genes confer susceptibility to autoimmune rheumatic diseases, including psoriasis [51] and systemic lupus erythematosus [52].

Analysis of three-locus genotypes reveal gene-gene interactions

We next considered three-locus genotypes in our expanded, combined population (132 affected and 390 unaffected dogs) where A=the *PAN2* variant encoding p.Arg492Cys, B=*MAP3K7CL* c.383_392ACTCCACAAA>GACT, and C=*DLA-DRB1*002:01*, lower-case letters denote wild type alleles (*c* represents any alternate allele of *DLA-DRB1*). Only 4% of dogs possessed a three-locus genotype with *cc*, barring further analysis of those nine genotypes. We considered nine of the remaining genotypes to be low-risk, as less than 6% of dogs with these allelic combinations had DMS (Table 3). Among healthy dogs (Fig 3A), the most frequently observed genotypes were *AabbCC* (24%) and *aabbCC* (15%). Based on penetrance, we classified five genotypes as conferring moderate risk (33-50%) and four as high risk (90-100%) for DMS. All cases possessed at least two risk alleles and all but one were homozygous for at least one risk

allele. The most common genotypes of affected dogs (Fig 3B) were *AaBBCC* (20% of cases), followed by *AAbbCC*, *AABbCC*, and *AABBCC* (17% of cases each).

Interestingly, only affected dogs possessed *AABBCc* or *AABBCC* (n=29), indicating that DMS is fully penetrant in dogs having these combinations. Epistasis plots illustrated that genotypes with at least one *a* or *b* allele confer a lower probability of disease when a *c* allele is present, compared to their *CC* counterparts (Fig 3, compare C and D). The plots also depicted a greater probability of disease than expected under a strictly additive model, providing evidence for additive-by-additive epistasis between the chromosome 10 and 31 loci [53,54]. We noted at least one ARE in *MAP3K7CL*, presenting a mechanism for interaction with *PAN2*. No difference in gene interactions was observed between the sexes (S4 Fig).

Information regarding age at onset or diagnosis was available for 91 dogs. We compared dogs having two, three, or four risk alleles across *PAN2* and *MAP3K7CL* and observed an inverse correlation between age of onset and number of risk alleles (S5 Fig). Dogs having four risk alleles developed DMS at a significantly younger median age (5 months) than did dogs with only two risk alleles (18.5 months). The complete penetrance of *AABB* genotypes, combined with an early age of onset, suggest that these dogs may be hypersensitive to commonplace environmental stimuli (*e.g.*, routine puppy vaccinations).

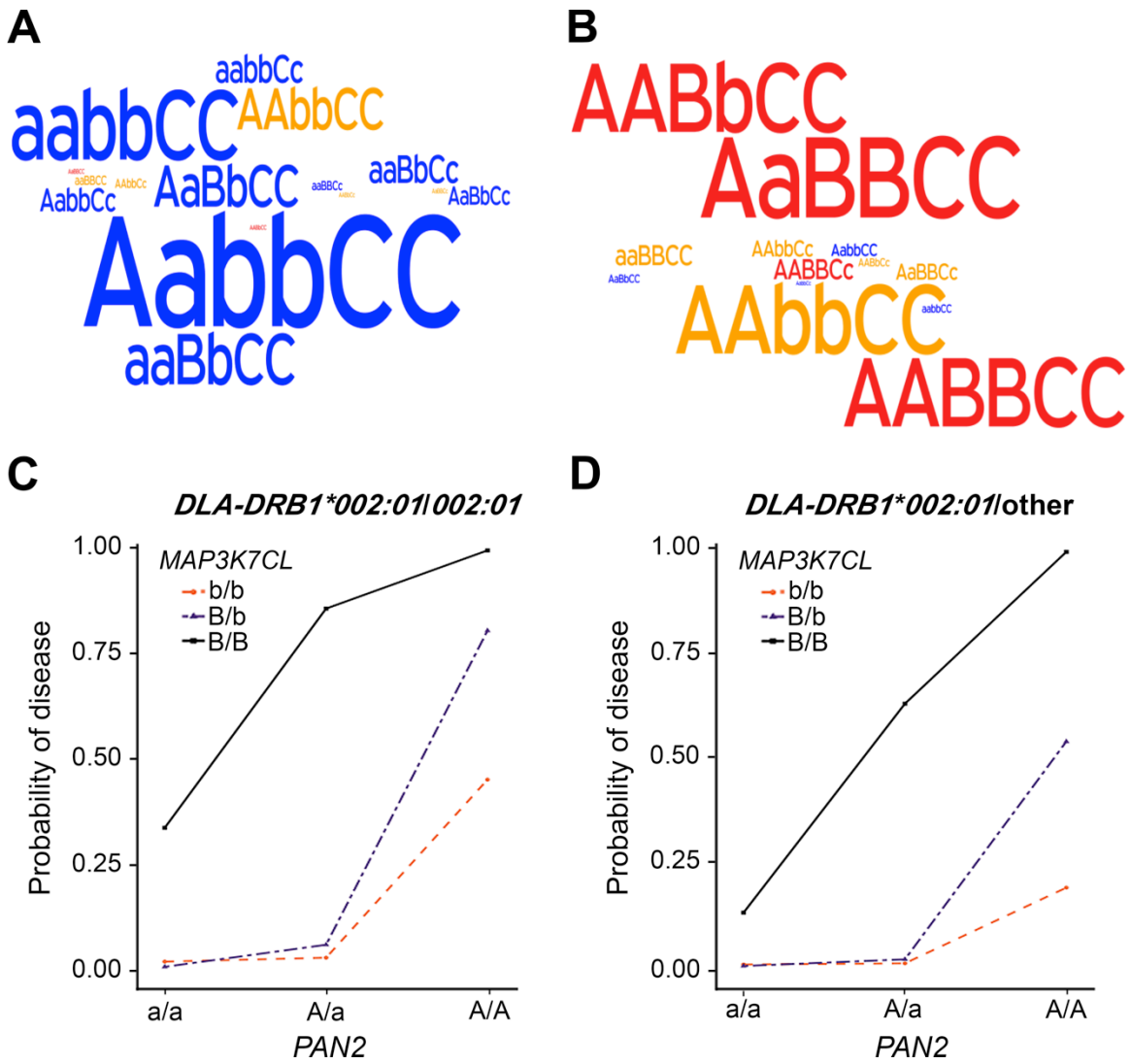


Figure 3. Three-locus genotype disease probabilities and frequencies (132 cases, 390 controls). Word clouds of three-locus genotypes illustrate frequency differences between (A) controls and (B) cases with high-risk genotypes in red, moderate-risk in yellow, and low-risk in blue. Probability of disease (y-axis) for all combinations of *PAN2* (x-axis) and *MAP3K7CL* genotypes are plotted in dogs (C) homozygous and (D) heterozygous for *DLA-DRB1*002:01*.

Genotype	Collies		Shetland sheepdogs		Combined			
	Cases (n=40)	Controls (n=185)	Cases (n=92)	Controls (n=205)	Penetrance (%)	Risk*	95% Confidence Interval	P-value **
<i>aabbcc</i>	0	1	0	1	-	-	0-0.842	1
<i>aabbCc</i>	0	5	0	18	-	low	0-0.148	0.0023
<i>aabbCC</i>	1	43	1	17	-	low	0.004-0.112	0
<i>Aabbcc</i>	0	0	0	3	-	-	0-0.708	0.577
<i>AabbCc</i>	0	3	1	15	-	low	0.001-0.260	0.060
<i>AabbCC</i>	2	71	2	23	-	low	0.011-0.101	0
<i>aaBbcc</i>	0	0	0	5	-	low	0-0.522	0.399
<i>aaBbCc</i>	0	1	0	21	-	low	0-0.154	0.002
<i>aaBbCC</i>	0	10	0	32	-	low	0.084	0
<i>AaBbcc</i>	0	0	0	1	-	-	0-0.975	1
<i>AaBbCc</i>	0	0	0	14	-	low	0-0.232	0.028
<i>AaBbCC</i>	0	13	2	23	-	low	0.006-0.177	0.002
<i>AABbcc</i>	1	0	0	0	-	-	0.025-1	0.253
<i>AABbCc</i>	4	6	1	1	42	moderate	0.152-0.723	0.194
<i>AABbCC</i>	15	31	7	3	39	moderate	0.265-0.532	0.021
<i>aaBBcc</i>	0	0	0	1	-	-	0-0.975	1
<i>aaBBCc</i>	0	0	0	6	-	low	0.0459	0.348
<i>aaBBCC</i>	0	0	6	7	46	moderate	0.192-0.749	0.107
<i>AaBBcc</i>	0	0	2	1	-	-	0.094-0.992	0.160
<i>AaBBCc</i>	0	0	5	5	50	moderate	0.187-0.813	0.136
<i>AaBBCC</i>	3	0	23	3	90	high	0.762-0.978	0
<i>AABbcc</i>	0	0	3	0	-	-	0.292-1	0.016
<i>AABbCc</i>	0	0	2	4	33	moderate	0.043-0.777	0.647
<i>AABbCC</i>	11	1	11	1	92	high	0.730-0.990	0
<i>AABBcc</i>	0	0	0	0	-	-	-	-
<i>AABBCC</i>	0	0	7	0	100	high	0.590-1	0.0001
<i>AABBCC</i>	3	0	19	0	100	high	0.846-1	0

Table 3. Distribution of three-locus genotypes in 132 cases and 390 controls with penetrance, 95% confidence intervals, and P-values. Significant values in bold.

*Risk interpretations were only made for three-locus genotypes observed at least five times.

**P-values are calculated for each genotype compared to the population as a whole (0.253, 132 affected dogs out of 522 total dogs).

Collies and Shetland sheepdogs have uniquely high frequencies of associated alleles

All three identified variants associated with DMS are polymorphisms present in several breeds, raising the question: why are other breeds rarely, if ever, affected by DMS? We genotyped five or more unrelated individuals from each of 30 diverse breeds for all three loci (Fig 4). The only other breeds to possess all three risk alleles were Cardigan Welsh corgis and Cairn terriers. Three Jack Russell terriers had moderate-risk genotypes ($AabbCc$), as did one Cardigan Welsh corgi ($AABbCc$); both breeds are occasionally diagnosed with dermatomyositis-like disease [56,57]. None of the 229 individuals possessed a high-risk genotype (S5 Table). Interestingly, Labrador retrievers had both B and C , which could have enabled moderate or high risk genotypes ($AaBBCc$ or $AaBBCC$) in puppies from the outcross mating described by Haupt et al. [14].

Combined frequencies of risk alleles in other breeds were dramatically lower than those observed among collie and Shetland sheepdog populations, and homozygosity for a risk allele (a characteristic of all moderate- to high-risk genotypes) was rare. Additionally, breeds having a high frequency of one risk allele had few or no risk alleles at the other loci. For example, Cairn terriers had a high frequency of A (75%) but low frequencies of B (19%) and C (3%), and no high- or moderate-risk genotypes were observed among 18 individuals. These findings suggest that independently the polymorphisms are neither deleterious nor selected against.

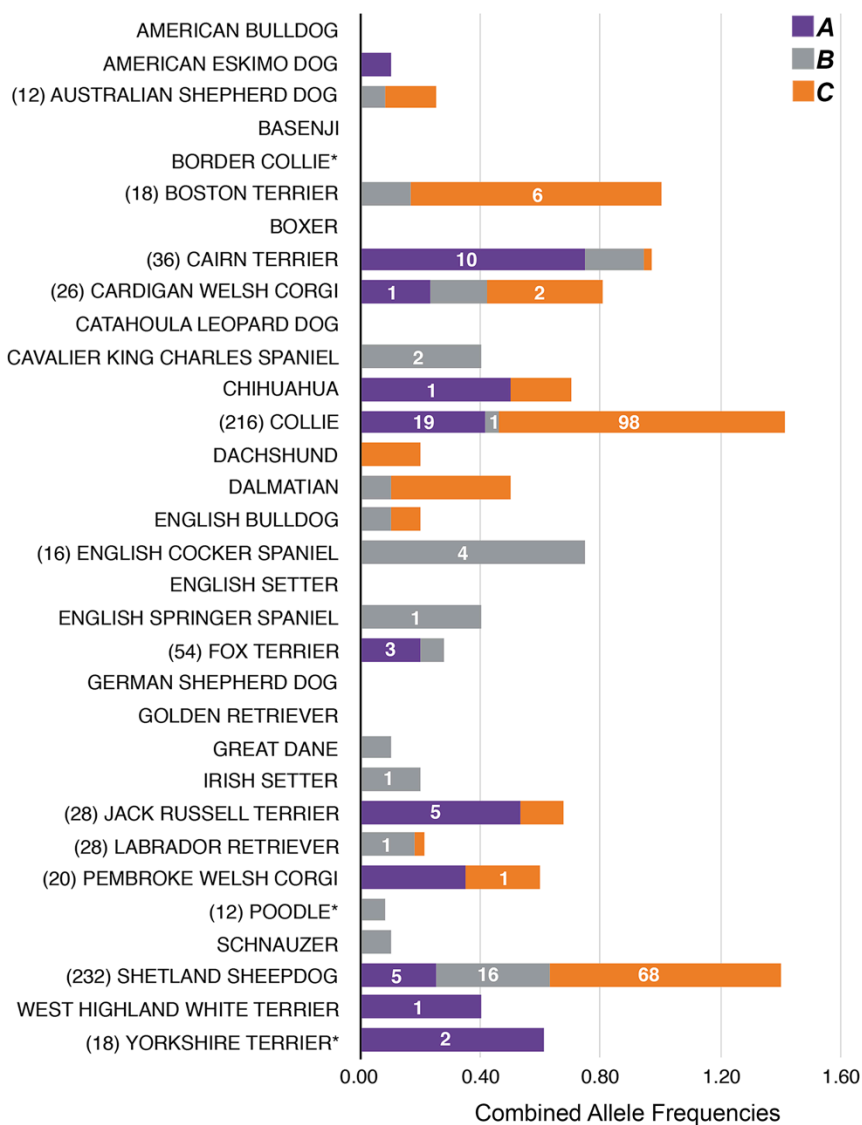


Figure 4. Frequencies of DMS-associated polymorphisms across breeds. The number of alleles used to calculate the frequencies is shown to the left of the breed name, if > 10. Risk allele frequencies (x-axis) for each gene are represented by a single, segmented colored bar: the *PAN2* variant encoding p.Arg492Cys is represented in purple, *MAP3K7CL* c.383_392ACTCCACAAA>GACT in gray, and *DLA-DRB1*002:01* in orange. The number of homozygous dogs observed is reported on the corresponding color bar. Asterisks denote breeds lacking *DLA-DRB1*002:01* in this study, but which were previously reported to possess the allele [55].

It is likely that recent artificial selection for phenotypes shared by collies and Shetland sheepdogs led to increased frequencies of *A*. We propose that persistence of *A* in these two breeds is attributed to linkage disequilibrium ($D'=0.998$) between wildtype *PAN2* (*a*) and another chromosome 10 allele, *Merle* of *PMEL*. In heterozygosity, *Merle* causes a popular pigmentation pattern (see collie in Fig 1A), but homozygosity for the allele results in severe hypopigmentation often with auditory and ocular defects [58]. Wildtype *PMEL* occurred on chromosomes with either *A* or *a*, whereas *Merle* was only found in conjunction with *a*. Accordingly, the *Merle* phenotype was underrepresented in affected dogs (P -value=0.0018), 64% of which were homozygous for *A*. Consistent selection for heterozygosity (but not homozygosity) for *Merle* would simultaneously encourage maintenance of both *PAN2* alleles. To our knowledge, there are no loci on chromosome 31 under positive selection for maintenance of a characteristic phenotype of collies and/or Shetland sheepdogs.

Across five collie genomes, we observed a 1.2 Mb selective sweep on chromosome 12 (S6 Fig) encompassing the MHC class II loci and leading to near fixation of the 002:01/009:01/001:01 haplotype. We suggest that essentially all purebred collies have increased susceptibility for DMS, conferred by homozygosity for allele *C*. The Shetland sheepdog population has retained a less common second haplotype that permits heterozygosity at the DLA loci, associated with a lower risk for developing DMS. Ironically, this reduced risk may have masked the presence of otherwise high-risk genotypes (*i.e.*, *AABb* and *AaBB*) and hindered selection against *A* and *B* alleles.

We observed striking allele frequency differences between the two affected breeds at *PAN2* and *MAP3K7CL*: collies had a higher frequency of *A*, 42% (25% in Shetland sheepdogs), whereas Shetland sheepdogs had a higher frequency of *B*, 38% (5% in collies) (Fig 4). Consequently, the frequency of observed allelic combinations varied between the breeds. The most common genotypes in healthy dogs were *aaBbCC* in Shetland sheepdogs, 16% (5% in collies), and *AabbCC* in collies, 38% (11% in Shetland sheepdogs). Among affected dogs, *AaBBCC* predominated in Shetland sheepdogs, 25% (8% in collies), whereas *AAbbCC* was the most frequent genotype in diseased collies, 38% (8% in Shetland sheepdogs). The latter finding is interesting given that *AAbbCC* is only a moderate-risk genotype. Among collies having this genotype, 67% were unaffected by age 8, whereas only 30% of Shetland sheepdogs with this genotype were disease-free. This discrepancy in disease probabilities between breeds was unique to this genotype. Further studies will be necessary to determine if other loci confer additional risk for or protection from DMS.

The contribution of alleles from multiple loci explains the spontaneous appearance of the disease in lines with no prior history [14] and has hindered elimination of DMS in the absence of a genetic test. For example, a mating between two healthy dogs having low risk genotypes (*e.g.*, *AaBbCC* x *AaBbCC*) can produce puppies with low, moderate, or high risk for DMS. This study has led to the first three-locus genetic test for a complex disease of dogs, which will allow breeders to carefully reduce the frequency of *A* and *B* among collie and Shetland sheepdog populations while preserving genetic diversity and desirable breed characteristics.

In a canine model of JDM, we have identified a complex pattern of causation involving three independent loci, two of which offer new targets for exploration in JDM. Furthermore, these data provide support for the involvement of genetic risk factors independent of the MHC in human inflammatory myopathies. While further experiments are necessary to determine the exact contribution of the chromosome 10 and 31 loci, our findings suggest that DMS may result from an inability to properly regulate inflammatory response. This work highlights the utility of the canine model for unraveling genetic susceptibility conferred by common polymorphisms and/or gene-gene interactions in complex diseases.

Methods

Ethics statement

All samples were obtained with informed consent according to protocols approved by the Clemson University Institutional Review Board (IBC2008-17) and IACUC (2012-039).

Study population.

Three populations were assembled: DMS-affected dogs (92 Shetland sheepdogs, 40 collies), control dogs for GWAS (45 Shetland sheepdogs, 23 collies), and unaffected dogs (160 Shetland sheepdogs, 162 collies). Affected dogs were diagnosed either through histopathology (75 Shetland sheepdogs, 30 collies) or by a veterinarian based on clinical signs and elimination of demodectic mange, a differential diagnosis. Pedigrees were

collected when available; some samples were obtained from affected dogs surrendered to rescue organizations without paperwork. Twenty percent of affected dogs were collected internationally, and in each case a regionally-matched control sample was obtained. Control dogs for GWAS were eight years of age or older at time of collection, had no known family history of DMS, had no personal history of skin disease, and were unrelated to other study participants within two (most often three) generations. Pedigrees were obtained from all control dogs. The population of unaffected dogs had no clinical signs of DMS and were eight years of age or older at time of collection. This subset was collected without regard to family history of DMS, presence of other skin disorders, or relationship to other study participants. Archival samples from 229 dogs of 30 other breeds were not closely related to each other to our knowledge.

Whole blood or buccal cells were obtained from each dog, and genomic DNA was isolated according to the Puregene DNA Isolation protocol (Gentra). Skin punch biopsies from active lesions or healthy tissue were also obtained from one Shetland sheepdog, one collie, and one Australian shepherd dog and preserved in RNAlater (Ambion).

DLA class II genotyping.

The hypervariable regions of *DLA-DRB1*, *-DQA1* and *-DQB1* were sequenced and genotyped according to protocols previously described [59-61]. *DLA-DQA1* was largely uninformative in collies and was inferred. Association of a haplotype or allele state with DMS was assessed through Fisher's exact tests, using VassarStats (Web Site for Statistical Computation, Vassar College, Poughkeepsie, NY).

Genome-wide association and LD analyses.

Genotyping was performed for 166 dogs (85 males and 81 females) using the Illumina CanineHD BeadChip, containing 173,662 SNPs. One sample having call rates <95% was excluded. SNPs having call rates <95%, minor allele frequencies <5%, and/or significant deviation from Hardy-Weinberg equilibrium (P -value <0.0001) in the control dogs were excluded. No evidence of genomic inflation was observed in the combined Shetland sheepdog and collie analysis ($\lambda=1.04$). Fisher's exact P -values were calculated under a dominant model.

LD pairwise analysis was performed to calculate r^2 values, which were plotted using LocusZoom [62]. We calculated r^2 values between SNPs on chromosome 10 using all controls. For chromosome 31, r^2 values were calculated using only control Shetland sheepdogs because no association was detected in collies alone. Assuming that the pathogenic variant would be in high LD with the lead SNP, candidate regions were characterized by SNPs with pairwise $r^2 \geq 0.6$ and defined by the first flanking associated SNPs in lower LD. Filtering and statistical analyses were conducted with SNP & Variation Suite v8 (SVS, Golden Helix). All chromosome positions throughout the text are reported in CanFam3.1.

Whole genome resequencing.

Three affected and two control collies were selected for whole genome resequencing along with one affected Shetland sheepdog. Genomic DNA fragments of

approximately 500bp were gel size selected for each sample and sequenced on two lanes of an Illumina HiSeq 2000, generating 2x100 (collies) and 2x125 bp (Shetland sheepdog) paired-end reads. 531 to 997 million total reads were generated for the five collies. Over one billion total reads were generated for the Shetland sheepdog. Paired reads were aligned to the indexed reference (CanFam3.1) with Bowtie2 [63] under sensitive parameters. The alignments were sorted and indexed with SAMtools [64,65] to be visualized in the Interactive Genomics Viewer [66].

RNA isolation and sequencing.

Total RNA was extracted from 30-40mg of skin punch biopsy tissue from active lesions (collie and Shetland sheepdog) or healthy skin (Australian shepherd) using the ToTALLY RNA kit (Ambion), according to the manufacturer's protocols. RNA samples were treated with DNase to remove DNA contamination, using the DNA-free kit (Ambion). Samples were quantitated on a NanoDrop 1000 spectrophotometer (Fisher Scientific).

Three RNAseq libraries were constructed per dog using normalized total RNA and the manufacturer's protocol for one of the following: TruSeq RNA library prep kit v2.0 (Illumina) or TruSeq stranded total RNA library prep kit (Illumina). An Agilent Bioanalyzer 2100 was used for size validation. Each sample was sequenced at 2 x 125bp paired-end on an Illumina HiSeq 2500 to a depth of at least 22 million reads.

FastQC from the Babraham Institute was used to assess read quality before and after preprocessing by Trimmomatic [67], which removed low quality bases and

remaining sequence adapters. Trimmed reads for each sample were aligned to CanFam3.1 using gsnap [68], and SAMtools [64,65] was used to generate sorted and indexed bam files.

Variant filtering and genotyping.

SAMtools and BCFtools [64,65] were used to generate variant call files for each sample, which were analyzed in SVS. On chromosome 10 (1-1,333,693 bp) variants that were heterozygous in affected collies 1 & 2 and homozygous in affected collie 3 and the affected Shetland sheepdog (*i.e.*, reference/reference or alternate/alternate), consistent with the allele states at the lead SNPs, were selected. On chromosome 31 (24,026,411-24,722,836 bp), all homozygous variants in the Shetland sheepdog were selected. Ensembl 79 was used to identify variants lying within predicted exons and 10 bp flanking sequences to capture splice sites. RNAseq data were manually inspected in IGV to determine whether predicted exonic variants were expressed in the affected dogs. Alternate variants were investigated using dbSNP and 27 genomes of 16 other breeds (either sequenced as part of ongoing studies or shared by other research groups) to determine whether any were unique to collies and Shetland sheepdogs. SIFT [31], PolyPhen [32], and PANTHER [33] were used to predict the impact of the amino acid substitutions. Genotyping of variants in additional dogs was accomplished by restriction digest assays or Sanger sequencing (S6 Table).

Age of onset.

Because age is not distributed as a normal random variable, we made use of the Weibull distribution, a density commonly used for time-to-event data [69]. Our goal was to estimate the mean and median age of onset, along with their confidence intervals, to facilitate comparisons across the three genotypic groups.

Taking advantage of a Bayesian strategy, we assume $y_{ij} \sim \text{Weibull}(r, e^{-b_i})$, where y_{ij} is the observed age at diagnosis for the j -th ($j = 1, 2, \dots, n_i$) dog in the i -th ($i = 1, 2, \dots, 3$) genotypic class and b_i is the effect of the i -th genotypic class. With this representation of the scale (r) and shape (e^{-b_i}) parameters of the Weibull density, the median of the i -th genotypic class is $e^{-b_i} \log(2)^{1/r}$ and the mean of the i -th genotypic class is $e^{-b_i} \Gamma(1 + 1/r)$.

In addition, we assumed a prior density for the unknown genetic parameters (b_i) to have a null mean and common variance σ (*i.e.*, $b_i \sim N(0, \sigma)$). We consider the weakly informative prior for $\sigma \sim \text{Cauchy}(0, 25)$ [70]. For the scale parameter, r , we assume the weakly informative prior of $r \sim \text{exponential}(0.001)$. The Monte Carlo Markov Chain (MCMC) sampling process was run in 4 chains through the public-domain software Stan [71], with each chain based on 50,000 total samples, and the first 20,000 were removed as part of the warm-up process, then thinned to every 20th sample, resulting in a MCMC sample of 6,000 values [71]. Convergence to the posterior density was evaluated by the Gelman-Rubin test statistic, where values less than 1.05 indicate that the MCMC sampling process was adequate to the data evaluated [72].

Genetic interaction analyses.

All affected (n=132) and healthy (n=390) dogs were used to investigate possible interactions between the *PAN2* variant encoding p.Arg492Cys and *MAP3K7CL* c.383_392ACTCCACAAA>GACT in dogs homozygous for 002:01/009:01/001:01 vs. dogs heterozygous for the haplotype. Three-locus genotypes observed in fewer than five dogs were excluded. For a given disease state (cases vs. controls) and a given pair of genotypic classes i ($i = A/A, A/a, \text{ or } a/a$) and j ($j = B/B, B/b, \text{ or } b/b$), $n_{\text{cases } ij} \sim \text{Binomial}(n_{\text{cases } ij} + n_{\text{controls } ij}, p_{ij})$ where $n_{\text{cases } ij}$ is the number of observed cases in the combination of the genotypic class of locus i and that of locus j , $n_{\text{controls } ij}$ is the number of unaffected dogs in the combination of genotypic class of locus i and that of locus j . Finally, p_{ij} is the probability of disease for the combined genotypic classes of i and j .

Estimation of the unknown elements of our model must ensure estimates of p_{ij} within the interval [0,1], recognizing that several genotypic classes have zero or few cases. We utilized a hierarchical Bayesian framework facilitated through the public-domain software Stan [71] to address this problem [73]. Stan can be accessed through the public-domain language R [74]. Log-odds was used to estimate p_{ij} , *i.e.*, $\log(p_{ij}/(1-p_{ij})) = \text{intercept} + \text{add}A_i + \text{dom}A_i + \text{add}B_j + \text{dom}B_j + \text{add } x \text{ add}_{ij} + \text{add } x \text{ dom}_{ij} + \text{dom } x \text{ add}_{ij} + \text{dom } x \text{ dom}_{ij}$, where intercept represents a term common to all genotypic classes, $\text{add}A_i$, $\text{add}B_j$, $\text{dom}A_i$ and $\text{dom}B_j$ represent the additive and dominance terms respectively for loci A and B , and $\text{add } x \text{ add}_{ij} + \text{add } x \text{ dom}_{ij} + \text{dom } x \text{ add}_{ij} + \text{dom } x \text{ dom}_{ij}$ represent the four possible epistatic interaction terms for all possible additive and dominance combinations [53].

We assumed a prior density for the intercept and unknown genetic parameters, with null mean and common variance σ (*i.e.*, $N(0, \sigma)$). Subsequently, we consider the weakly informative prior for $\sigma \sim \text{Cauchy}(0,25)$ [70]. We used the same MCMC parameters here as described above.

Chromosome 12 selective sweeps.

We identified all SNPs on chromosome 12 present in five collie genomes (3 cases and 2 controls) and used a creeping window size ≤ 1 Mb to identify runs of homozygosity [75]. Windows containing fewer than 50 SNPs were excluded and gaps >10 kb between SNPs were ignored. The heterozygosity (H_p) statistic was calculated for all windows and Z-transformed, making the average H_p value equal to zero and the standard deviation equal to 1. The $-ZH_p$ distribution was plotted in R to show putatively selected regions greater than 3.4 standard deviations from the mean.

PMEL

D' between *PMEL* and *PAN2* was calculated as a measure of LD using all dogs for which coat pattern phenotypes were available (112 cases, 385 controls). A total of 98 dogs (10 cases and 88 controls) were described as merle-patterned, a semi-dominant trait caused by the *Merle* (*M*) allele of *PMEL*, and assumed to possess the *Mm* genotype. Computation of D' , along with a test of significance from zero, was facilitated through the package *genetics* [76], a public domain program that is part of the R language [74].

Accession numbers

SNP data are available in dbSNP under BioProject number PRJNA338128. All whole genome and transcriptome data generated for this study were deposited in SRA genomes under accession number SRP081080. Accession numbers for eight of the 27 other breeds used in variant filtering are SRX1360633, SRX1360635, SRX1360637, SRX1360639, SRX1022256, SRX1022262, SRX1022286, and SRP081080.

Acknowledgments

The authors wish to thank Dr. Christine Rees and Sherry Lindsey for contributing samples, as well as the many dog owners and veterinarians. We thank Drs. Gary Johnson and Kathryn Meurs for providing genomes from unaffected dogs. We thank Drs. Mike Vaughan, Michael Sehorn, and Stephen Kresovich for critical reading of the manuscript.

Author contributions

Conceptualization: LAC, KLT, JME

Validation: JME, LAC, REN, TRF, KLT, ASM

Formal analysis: JME, TRF, ASM

Investigation: JME, REN, CMH, KJA

Data curation: REN, KLT, JME

Writing – Original draft preparation: JME, LAC, TRF, REN

Writing – Review & editing: JME, LAC, ASM, KJA, KLT

Visualization: LAC, JME, REN, TRF

Supervision: JME, LAC

Project administration: LAC

Funding acquisition: LAC, KLT

References

1. Rider LG, Nistala K. The juvenile idiopathic inflammatory myopathies: pathogenesis, clinical and autoantibody phenotypes, and outcomes. *J Intern Med.* 2016;280:24-38.
2. Mendez EP, Lipton R, Ramsey-Goldman R, Roettcher P, Bowyer S, Dyer A, et al. US incidence of juvenile dermatomyositis, 1995-1998: results from the National Institute of Arthritis and Musculoskeletal and Skin Diseases Registry. *Arthritis Rheum.* 2003;49:300-305.
3. Bowyer SL, Blane CE, Sullivan DB, Cassidy JT. Childhood dermatomyositis: factors predicting functional outcome and development of dystrophic calcification. *J Pediatr.* 1983;103:882-888.
4. Manlhiot C, Liang L, Tran D, Bitnun A, Tyrrell PN, Feldman BM. Assessment of an infectious disease history preceding juvenile dermatomyositis symptom onset. *Rheumatology* 2008;47:526-529.
5. Reed AM, Stirling JD. Association of the HLA-DQA1*0501 allele in multiple racial groups with juvenile dermatomyositis. *Hum Immunol.* 1995;44:131-135.
6. Pachman LM, Liotta-Davis MR, Hong DK, Kinsella TR, Mendez EP, Kinder JM, et al. TNF α -308A allele in juvenile dermatomyositis: Association with increased

- production of tumor necrosis factor α , disease duration, and pathologic calcifications. *Arthritis Rheum.* 2000;43:2368-2377.
7. Mamyrova G, O'Hanlon TP, Monroe JB, Carrick DM, Malley JD, Adams S, et al. Immunogenetic risk and protective factors for juvenile dermatomyositis in Caucasians. *Arthritis Rheum.* 2006;54:3979-3987.
 8. Mamyrova G, O'Hanlon TP, Sillers L, Malley K, James-Newton L, Parks CG, et al. Cytokine gene polymorphisms as risk and severity factors for juvenile dermatomyositis. *Arthritis Rheum.* 2008;58:3941-3950.
 9. Miller FW, Cooper RG, Vencovský J, Rider LG, Danko K, Wedderburn LR, et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis Rheum.* 2013;65:3239-3247.
 10. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes Immun.* 2015;16:470-480.
 11. Hargis AM, Haupt KH, Hegreberg GA, Prieur DJ, Moore MP. Familial canine dermatomyositis. Initial characterization of the cutaneous and muscular lesions. *Am J Pathol.* 1984;116:234-244.
 12. Hargis AM, Haupt KH, Prieur DJ, Moore MP. Animal model of human disease. Dermatomyositis. Familial canine dermatomyositis. *Am J Pathol.* 1985;120:323-325.

13. Haupt KH, Prieur DJ, Hargis AM, Cowell RL, McDonald TL, Werner LL, et al. Familial canine dermatomyositis: clinicopathologic, immunologic, and serologic studies. *Am J Vet Res.* 1985;46:1870-1875.
14. Haupt KH, Prieur DJ, Moore MP, Hargis AM, Hegreberg GA, Gavin PR, et al. Familial canine dermatomyositis: clinical, electrodiagnostic, and genetic studies. *Am J Vet Res.* 1985;46:1861-1869.
15. Nagaraju K, Plotz PH. Animal models of myositis. *Rheum Dis Clin N Am.* 2002;28:917-933.
16. Avci P, Sadasivam M, Gupta A, De Melo WC, Huang YY, Yin R, et al. Animal models of skin disease for drug discovery. *Expert Opin Drug Discov.* 2013;8:331-355.
17. Kunkle GA, Chrisman CL, Gross TL, Fadok V, Werner LL. Dermatomyositis in collie dogs. *Comp Contin Educ Pract Vet.* 1985;7:185-192.
18. Hargis A, Mundell A. Familial canine dermatomyositis. *Compend Contin Educ Vet.* 1992;4:855-864.
19. Scott DW, Miller WH, Griffin CE. Congenital and hereditary defects. In: Muller and Kirk's *Small Animal Dermatology*, 6th ed. Philadelphia: Saunders, 2000 pp. 940-946.
20. Wahl JM, Clark LA, Skalli O, Ambrus A, Rees CA, Mansell JL, et al. Analysis of gene transcript profiling and immunobiology in Shetland sheepdogs with dermatomyositis. *Vet Dermatol.* 2008;19:52-58.

21. Hargis AM, Prieur DJ, Haupt KH, McDonald TL, Moore MP. Prospective study of familial canine dermatomyositis: correlation of the severity of dermatomyositis and circulating immune complex levels. *Am J Pathol.* 1986;123:465-479.
22. Rider LG, Wu L, Mamyrova G, Targoff IN, Miller FW. Environmental factors preceding illness onset differ in phenotypes of the juvenile idiopathic inflammatory myopathies. *Rheumatology* 2010;49:2381–2390.
23. Dourmishev LA, Dourmishev AL. *Dermatomyositis: Advances in Recognition, Understanding and Management.* New York: Springer; 2009.
24. vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 2010;464:898-902.
25. Wilbe M, Jokinen P, Truvé K, Seppala EH, Karlsson EK, Biagi T et al. Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. *Nat Genet.* 2010;42:250-254.
26. Karlsson EK, Sigurdsson S, Ivansson E, Thomas R, Elvers I, Wright J et al. Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome Biol.* 2013;14:R132.
27. Karyadi DM, Karlins E, Decker B, vonHoldt BM, Carpintero-Ramirez G, Parker HG et al. A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit. *PLoS Genet.* 2013;9:e1003409.

28. Tang R, Noh HJ, Wang D, Sigurdsson S, Swofford R, Perloski M, et al.
Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biol.* 2014;15:R25.
29. Ivansson EL, Megquier K, Kozyrev SV, Murén E, Körberg IB, Swofford R, et al.
Variants within the SP110 nuclear body protein modify risk of canine degenerative myelopathy. *Proc Natl Acad Sci.* 2016;113:E3091-100.
30. Melin M, Rivera P, Arendt M, Elvers I, Murén E, Gustafson U, et al. Genome-wide analysis identifies germ-line risk factors associated with canine mammary tumours. *PLoS Genet.* 2016;12:e1006029.
31. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4:1073-1081.
32. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-249.
33. Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics.* 2016;32:2230-2232.
34. Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* 2001;104:377-386.

35. Wolf J, Passmore LA. mRNA deadenylation by Pan2-Pan3. *Biochem Soc Trans.* 2014;42:184-187.
36. Weill L, Belloc E, Bava FA, Méndez R. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat Struct Mol Biol.* 2012;19:577-585.
37. Bett JS, Ibrahim AF, Garg AK, Kelly V, Pedrioli P, Rocha S, et al. The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability. *Biochem J.* 2013;451:185-194.
38. Palazon A, Goldrath AW, Nizet V, Johnson RS. HIF transcription factors, inflammation, and immunity. *Immunity* 2014;41:518-528.
39. Bakheet T, Frevel M, Williams BR, Greer W, Khabar KS. ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. *Nucleic Acids Res.* 2001;29:246-254.
40. Preuße C, Allenbach Y, Hoffmann O, Goebel HH, Pehl D, Radke J, et al. Differential roles of hypoxia and innate immunity in juvenile and adult dermatomyositis. *Acta Neuropathol Commun.* 2016;4:45.
41. Bilgic H, Ytterberg SR, Amin S, McNallan KT, Wilson JC, Koeuth T, et al. Interleukin-6 and type I interferon-regulated genes and chemokines mark disease activity in dermatomyositis. *Arthritis Rheum.* 2009;60:3436-3446.
42. Li J, Ji C, Yang Q, Chen J, Gu S, Ying K, et al. Cloning and characterization of a novel human TGF-beta activated kinase-like gene. *Biochem Genet.* 2004;42:129-137.

43. Kilpinen S, Ojala K, Kallioniemi O. Analysis of kinase gene expression patterns across 5681 human tissue samples reveals functional genomic taxonomy of the kinome. *PLoS One* 2010;5:e15068.
44. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol.* 2007;8:R24.
45. Woolf E, Xiao C, Fainaru O, Lotem J, Rosen D, Negreanu V, et al. Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proc Natl Acad Sci USA.* 2003;100:7731-7736.
46. Djuretic IM1, Cruz-Guilloty F, Rao A. Regulation of gene expression in peripheral T cells by Runx transcription factors. *Adv Immunol.* 2009;104:1-23.
47. Fainaru O, Woolf E, Lotem J, Yarmus M, Brenner O, Goldenberg D, et al. Runx3 regulates mouse TGF-beta-mediated dendritic cell function and its absence results in airway inflammation. *EMBO J.* 2004;23:969-979.
48. Brenner O, Levanon D, Negreanu V, Golubkov O, Fainaru O, Woolf E, et al. Loss of Runx3 function in leukocytes is associated with spontaneously developed colitis and gastric mucosal hyperplasia. *Proc Natl Acad Sci USA.* 2004;101:16016-16021.
49. Guo C, Yao F, Wu K, Yang L, Zhang X, Ding J. Chromatin immunoprecipitation and association study revealed a possible role of Runt-related transcription factor 3 in the ulcerative colitis of Chinese population. *Clin Immunol.* 2010;135:483-489.

50. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet.* 2011;43:761-767.
51. Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, Gordon D, et al. A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nat Genet.* 2003;35:349-356.
52. Prokunina L, Castillejo-López C, Oberg F, Gunnarsson I, Berg L, Magnusson V, et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet.* 2002;32:666-669.
53. Hallgrímsson IB, Yuster DS. A complete classification of epistatic two-locus models. *BMC Genet.* 2008;9:17.
54. Mackay TR. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet.* 2014;15:22-33.
55. Kennedy LJ, Barnes A, Short A, Brown JJ, Lester S, Seddon J, et al. Canine DLA diversity: 1. New alleles and haplotypes. *Tissue Antigens* 2007;69:272–288.
56. Parker WM, Foster RA. Cutaneous vasculitis in five Jack Russell Terriers. *Vet Dermatol.* 1996;7:109-115.
57. Ihrke PJ. Ischemic skin disease in the dog. 31st World Small Animal Veterinary Congress (2006).

58. Clark LA, Wahl JM, Rees CA, Murphy KE. Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. *Proc Natl Acad Sci USA* 2006;103:1376-1381.
59. Evans JM, Cox ML, Huska J, Li F, Gaitero L, Guo LT, et al. Exome sequencing reveals a nebulin nonsense mutation in a dog model of nemaline myopathy. *Mamm Genome* 2016;27:495-502.
60. Kennedy LJ, Davison LJ, Barnes A, Short AD, Fretwell N, Jones CA, et al. Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. *Tissue Antigens* 2006;68:467–476.
61. Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SGE. IPD - the immuno polymorphism database. *Nucleic Acids Res.* 2013;41:1234–1240.
62. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP et al. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336-7.
63. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357-359.
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 2009;25:2078-2079.
65. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987-2993.

66. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. *Nat Biotechnol.* 2011;29:24–26.
67. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114-2120.
68. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 2010;26:873-881.
69. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. New York: Springer; 2005.
70. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 2006;1:515-534.
71. Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, et al. Stan: A probabilistic programming language. *J Stat Softw.* 2016;Forthcoming
72. Gelman A, Rubin D. Inference from iterative simulation using multiple sequences. *Statist Sci.* 1992;7:457-472.
73. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*, 3rd ed. Broken Sound Parkway: CRC Press Taylor & Francis Group; 2014.
74. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria; 2015.
75. Qanbari S, Strom TM, Haberer G, Weigend S, Gheyas AA, Turner F, et al. A high resolution genome-wide scan for significant selective sweeps: an application to pooled sequence data in laying chickens. *PLoS One* 2012;7:e49525.

76. Warnes G, Gorjanc G, Leisch F, Man M. Genetics: Population Genetics. R package version 1.3.8.1 2013.

CHAPTER FIVE

CONCLUSION

The outcome of this work was the identification of mutations underlying three inherited diseases of dogs: nemaline rod myopathy in American bulldogs, limb girdle muscular dystrophy in Boston terriers, and dermatomyositis in collies and Shetland sheepdogs. Genetic tests are now available to facilitate breeding decisions and reduce the risk of disease in the affected breeds.

Nemaline rod myopathy in American bulldogs

Genome-wide SNP and whole-exome resequencing (WES) data were generated for a nuclear family of American bulldogs (ABDs) with two autosomal recessive cases of nemaline rod myopathy. Candidate intervals were determined based on allele frequency in the cases vs. the unaffected family members. SNP genotypes consistent with an autosomal recessive inheritance pattern were identified proximal to *NEB* and *TNNT1*, two known candidate genes. WES revealed a nonsense mutation in exon 176 of 183 in *NEB*. This variant was absent in unrelated ABDs as well as dogs of other breeds, indicating that it is private to the family. A genetic test now exists for nemaline rod myopathy in ABDs, and this work demonstrated the utility of WES in detecting autosomal recessive variants when few cases are available for study.

Limb girdle muscular dystrophy in Boston terriers

We identified independent mutations in *SGCD* causing limb girdle muscular dystrophy (LGMD) in two unrelated Boston terriers. LGMD has been described previously in dogs but never molecularly characterized until now. WES of the cases and unaffected family members of one case revealed two *SGCD* deletions: a 2 bp deletion segregating in the family and a 19.4 kb deletion encompassing exons 7 and 8 of *SGCD* in the unrelated case. Given that LGMD is rare in dogs and that both affected dogs were of the same breed, we expected to identify a single mutation present at a very low frequency in the Boston terrier breed. This work again demonstrates the efficacy of WES for identifying causative mutations of simple recessive disorders, particularly when candidate genes are available. Should future cases of LGMD be identified in Boston terriers, a test is now available for each of the identified mutations, and there is now potential for a large animal model for LGMD2F gene therapy.

Dermatomyositis in collies and Shetland sheepdogs

Dermatomyositis (DMS) is a complex autoimmune disease with environmental and genetic influences and is a model for human juvenile dermatomyositis (JDM). To date, major associations with JDM have not been identified outside of the major histocompatibility complex (MHC), although other risk factors of strong effect are believed to exist. Investigation of the dog leukocyte antigen loci in affected and healthy collies and Shetland sheepdogs revealed a haplotype overrepresented among affected dogs (*DLA-DRB1*002:01/-DQA1*009:01/-DQB1*001:01*). Through a combination of

genome-wide association studies (GWAS), whole genome resequencing (WGS), and transcriptome sequencing, we identified strong associations with variants in two genes never before implicated in disease, *PAN2* and *MAP3K7CL*. Furthermore, we found evidence for additive-by-additive epistasis between the novel risk loci.

None of the risk alleles were private to collies and Shetland sheepdogs, but these breeds were unique in their high frequencies of risk alleles at the three loci. We determined that linkage disequilibrium between the *M* allele of *SILV* and the wildtype *PAN2* allele has impeded elimination of the risk allele in the absence of a genetic test. Furthermore, we found evidence for a selective sweep including the MHC on chromosome 12 in collies, explaining their fixation for the DMS risk haplotype. It is likely that breeders fixed this locus inadvertently while selecting for a desirable phenotype. The presence of a second MHC haplotype in Shetland sheepdogs contributed to the higher incidence of the *MAP3K7CL* variant in this breed.

This work is the first in dogs to identify disease risk associated with combinations of polymorphisms across three independent loci, and we can explain 93% of cases with these three-locus genotypes. It has also yielded novel genes for investigation in human cases of dermatomyositis. Because DMS is fully penetrant in dogs homozygous for risk alleles at all three loci develop DMS, there is potential to study changes to the immune system prior to disease onset. Such research has not previously been possible and may allow for the development of biomarkers for diagnosis of JDM before clinical symptoms ensue.

For dogs, this work has resulted in a genetic test that will allow breeders to preserve genetic diversity, while reducing the incidence of the risk alleles in these breeds. Prior to the development of a test, affected dogs and parents of affected dogs would be culled from the breeding population. Now all dogs can be safely bred to produce puppies with low-risk genotypes, and Shetland sheepdog breeders can maintain the second DLA haplotype found in that breed.

APPENDICES

Appendix A

Association of DLA-DQB1 alleles with exocrine pancreatic insufficiency in Pembroke

Welsh corgis

J.M. Evans¹, K.L. Tsai¹, A.N. Starr-Moss¹, J.M. Steiner², L.A. Clark¹

¹Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634,
USA

²Department of Small Animal Clinical Sciences, Texas A&M University, College
Station, TX 77843, USA

Published – Animal Genetics

Reprinted with permission from John Wiley and Sons, see Appendix D.

Summary

Exocrine pancreatic insufficiency is a digestive disorder resulting from insufficient secretion of enzymes from the pancreas. In dogs, this condition is often attributed to pancreatic acinar atrophy, wherein the enzyme-producing acinar cells are believed to be destroyed through an autoimmune process. Although exocrine pancreatic insufficiency affects many diverse breeds, to date, molecular studies have been limited to the German shepherd dog. A recent study of major histocompatibility genes in diseased and healthy German shepherd dogs identified both risk and protective haplotypes. Herein, we genotyped *DLA-DQB1* in Pembroke Welsh corgis to determine if dog leukocyte

antigen alleles contribute to the pathogenesis of exocrine pancreatic insufficiency across dog breeds. We evaluated 14 affected and 43 control Pembroke Welsh corgis, which were selected based on an age of onset similar to German shepherd dogs. We identified one protective allele (odds ratio=0.13, *P*-value=0.044) and one risk allele (odds ratio=3.8, *P*-value=0.047). As in German shepherd dogs, the risk allele is a duplication of *DLA-DQB1* (alleles *DQB1*013:03* and *017:01*); however, Pembroke Welsh corgis have acquired a single polymorphism on *DQB1*017:01*. Thus, the *DLA-DQB1* duplication is a risk allele for exocrine pancreatic insufficiency in at least two breeds.

Exocrine pancreatic insufficiency (EPI) results from the failure of pancreatic acinar cells to produce enough enzymes for proper digestion of macromolecular nutrients. Clinical signs include increased appetite, weight loss, and loose stools. It has been suggested that clinical signs occur when <10% of pancreatic enzyme reserves remain [1]. EPI is typically managed with supplemental enzymes. Veterinarians screen for EPI through measurement of serum canine trypsin-like immunoreactivity (cTLI) concentration; however, the underlying cause is usually undetermined. In dogs, EPI is most commonly caused by pancreatic acinar atrophy (PAA), an autoimmune disorder resulting from selective degradation of acinar cells after lymphocytic infiltration [2]. PAA is responsible for hereditary forms of EPI and has a complex mode of inheritance [3].

EPI occurs across many dog breeds but is especially prevalent among German shepherd dogs (GSDs). Tsai et al. [4] recently identified associations between EPI and

four-locus dog leukocyte antigen (DLA) haplotypes in GSDs. The DLA locus is part of the canine major histocompatibility complex, a superlocus of genes involved in immune function. The identified risk haplotype (*DLA-88*045:02/DLA-DRB1*012:01/DLA-DQA1*004:01/DLA-DQB1*013:03+017:01*; for nomenclature, see Kennedy et al. [5]) contains a *DLA-DQB1* allele that is not found on other haplotypes. This allele is a duplication composed of two alleles (*DQB1*013:03* and *017:01*) and has been reported in at least 15 breeds [6], including several reported to have EPI. As in GSDs, Pembroke Welsh corgis (PWCs) have a young average age of onset (~3 years), which may imply a similar pathogenesis [7]. To determine whether *DQB1*013:03+017:01* confers risk for development of EPI in breeds besides GSDs, we genotyped *DLA-DQB1* in diseased and healthy PWCs.

Whole blood or buccal cells were obtained from PWCs across 22 states recruited through www.epi4dogs.com and the Pembroke Welsh Corgi Club of America. All samples and informed owner consent were collected according to protocols approved by the Clemson University Institutional Review Board (IBC2008-17). Pedigree information was obtained for 54% of participants. Among these, all dogs were unrelated at the parental level with most unrelated through the grandparent level. All 14 PWC cases were diagnosed based on clinical assessment by their private veterinarians and fasting serum cTLI concentrations ≤ 2.5 $\mu\text{g/L}$ (reference interval: 5.7-45 $\mu\text{g/L}$). PWC cases had an average age at diagnosis of 2.9 years.

All 43 dogs selected as controls were ≥ 8 years of age (average=11 years), exhibited no clinical signs of EPI or digestive disorders, and had no immediate family

members diagnosed with EPI to the owner's knowledge. Although rare, subclinical cases have been reported in GSDs [8]. Fasting serum samples were collected from six controls having the *DLA-DQB1* duplication and submitted to the Gastrointestinal Laboratory at Texas A&M University for cTLI measurement. All dogs had cTLI concentrations within the normal range.

The hypervariable regions of *DLA-DQB1* and *DLA-DRB1* were amplified from genomic DNA using primers and thermal cycling parameters reported by Kennedy et al. [6] or modified from Massey et al. [9] (Table S1). Products were purified according to Massey et al. [9] and sequenced on an ABI 3730xl Genetic Analyzer (Life Technologies). Results were compared with published nucleotide sequences to assign genotypes [10], and statistical analyses were performed using VassarStats (Vassar College).

Gel extracted PCR product from a dog having a novel *DLA-DQB1* allele was subcloned into pGEM[®]-T Easy vector according to the manufacturer's protocols (Promega). Colonies containing the insert were selected, and the vector and insert were sequenced.

In our cohort, we identified 11 *DLA-DQB1* alleles, for which 96% of PWCs were heterozygous (Table 1). The most commonly observed alleles were *DQB1**020:02 and *001:01* (18% each), followed by *003:01* (16%).

<i>DLA-DQB1</i>	Cases		Controls		OR	95% CI	P-value
	n=14	(%)	n=43	(%)			
<i>001:01</i>	7	50	13	30	2.3	0.67–7.9	0.21
<i>002:01</i>	2	14	13	30	0.38	0.075–1.97	0.31
<i>003:01</i>	3	21	15	35	0.51	0.12–2.1	0.51
<i>007:01</i>	0	0	1	2.3			-
<i>008:02</i>	1	7.1	16	37	0.13	0.02–1.1	0.044
<i>013:02</i>	0	0	1	2.3			-
<i>013:03</i>	0	0	1	2.3			-
<i>013:03+</i> <i>017:02</i>	7	50	9	21	3.8	1.1–14	0.047
<i>020:02</i>	6	43	15	35	1.4	0.41–4.8	0.75
<i>023:01</i>	1	7.1	1	2.3			-
<i>035:01</i>	0	0	1	2.3			-

Table 1. Frequencies and statistical values for each *DLA-DQB1* allele in PWC cases and controls under a dominant model. Statistical values are shown for alleles present in more than two PWCs and are calculated based on numbers of dogs with and without each allele. Significant values are highlighted in bold font. OR=odds ratio, CI=confidence interval

We initially genotyped seven cases and nine controls as heterozygous for *DQB1*013:03+017:01*. Eight of these were in combination with three alleles that did not match reported alleles but most closely resembled *DQB1*001:01*, *003:01*, and *035:01*. To better characterize these alleles, *DLA-DRB1* genotypes were generated for five PWCs, representing each of the three unknown alleles. All dogs were heterozygous for *DRB1*012:01*, which is on a haplotype with *DQB1*013:03+017:01* [11]. The alternate alleles were *DRB1*002:01*, *015:01*, or *025:01*, which are reported on the same haplotypes as *DQB1*001:01*, *003:01*, and *035:01*, respectively [11]. Further examination of *DQB1* sequence data from all dogs having the duplication revealed a shared mutation

(c.236A>T) at a known polymorphic site. These data suggest that the mutation is present on the common allele, *DQB1*013:03+017:01*, rather than on the alternative alleles. Subcloning and sequencing of the *DLA-DQB1* duplication revealed that the c.236A>T mutation is found in *DQB1*017:01*; no changes were present in *DQB1*013:03* (Figure 1A). The curator of the canine immunopolymorphism-MHC database termed the new allele *DLA-DQB1*017:02* and the duplicated allele present in PWCs *DLA-DQB1*013:03+017:02*.

We subsequently developed a method of distinguishing dogs having *DQB1*013:03+017:01* from those having *DQB1*013:03+017:02* in direct sequences. Initially, c.236A>T was only apparent when the *DLA-DQB1* duplication was paired with an allele having an A at this position (*DQB1*001:01*, *003:01*, *023:01*, and *035:01*). Chromatograms from this allelic combination have a stronger A signal resulting from a 2:1 ratio of As to Ts: one T from *DQB1*017:02*, one A from *DQB1*013:03*, and a second A from the third allele (Figure 1B).

The remaining eight *DQB1*013:03+017:01* alleles were paired with alleles having c.236A>T (*DQB1*002:01*, *008:02*, and *020:02*). Chromatograms from these allelic combinations showed a stronger T signal, suggesting a 1:2 ratio of As to Ts. In chromatograms from GSDs having *DQB1*013:03+017:01* and an allele with c.236A>T, the A signal is greater (Figure 1B). Therefore, we determined that the remaining PWCs have the *DQB1*013:03+017:02* duplication, which is contributing the second T in the 1:2 ratio, and that no dogs in our cohort have *DQB1*013:03+017:01*.

To determine if *DLA-DQB1* alleles are associated with EPI in PWCs, a Fisher's exact test was conducted using numbers of dogs possessing a given allele. This approach assumes a dominant model, wherein only one allele is necessary to increase risk. *DQB1*013:03+017:02* was determined to confer risk (OR=3.8, *P*-value=0.047). Another allele, *DQB1*001:01*, had an odds ratio suggestive of risk (OR>2.0), but its *P*-value was not significant (Table 1).

A single allele, *DQB1*008:02*, was significantly associated with PWC controls, indicating protection. This allele was not observed in GSDs [4]. In GSDs, EPI-protective *DLA* haplotypes included *DQB1* alleles *002:01* and *003:01* [4]. In the PWCs, these alleles had odds ratios suggestive of a protective effect (OR<0.5); however, the *P*-values were not significant (Table 1).

In GSDs, the protective haplotype was dominant, such that the single dog having both protective and risk haplotypes did not develop EPI [4]. In this study, we observed two PWC controls that possessed both the risk and protective alleles. This genotype was not observed among PWC cases, indicating that the protective allele may be dominant to the risk allele. One case had *DQB1*008:02* paired with a non-risk allele, suggesting that the protective allele does not fully prevent risk for EPI.

The duplication of *DQB1*013:03* and *017:01* occurs across several breeds and represents an ancestral event. Genotyping of *DLA-DRB1* showed that the PWC duplication occurs with *DRB1*012:01*, as does *DQB1*013:03+017:01*. Taken together, these data indicate that the duplication in PWCs is not a novel event. Rather, we suggest

that PWCs have acquired a polymorphism on *DQBI*017:01* within the duplication, which has not yet been reported in other breeds.

In conclusion, we have identified both risk and protective alleles contributing to EPI in PWCs. Our findings combined with those of Tsai et al. [4] show that this *DLA-DQBI* duplication is associated with EPI in multiple dog breeds. Further studies are necessary to determine if the duplication itself or a variant in linkage disequilibrium is contributing to the pathogenesis of EPI.

Acknowledgements

The authors would like to thank the dog owners and veterinarians for assisting with sample collection. We thank Lorna Kennedy, the curator of canine immunopolymorphism-MHC database, for assisting with allele identification. Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number R15AR062868. Parts of this work were also funded by the Clemson University Calhoun Honors College and donations from EPI dog owners.

References

1. Dimagno E, Go VL, Summerskill W. Relations between pancreatic enzyme outputs and malabsorption in severe pancreatic insufficiency. *N Engl J Med.* 1973;288:813-815.

2. Wiberg ME, Saari SAM, Westermarck E. Exocrine pancreatic atrophy in German shepherd dogs and rough-coated collies: an end result of lymphocytic pancreatitis. *Vet Pathol.* 1999;36:530-541.
3. Westermarck E, Saari SAM, Wiberg ME. Heritability of exocrine pancreatic insufficiency in German shepherd dogs. *J Vet Intern Med.* 2010;24:450-452.
4. Tsai KL, Starr-Moss AN, Venkataraman GM, Robinson C, Kennedy LJ, Steiner JM, et al. Alleles of the major histocompatibility complex play a role in the pathogenesis of pancreatic acinar atrophy in dogs. *Immunogenetics.* 2013;65:501-509.
5. Kennedy LJ, Altet L, Angles JM, Barnes A, Carter SD, Francino O, et al. Nomenclature for factors of the Dog Major Histocompatibility System (DLA), 1998: first report of the ISAG DLA Nomenclature Committee. *Anim Genet.* 2000;31:52-61.
6. Kennedy LJ, Davidson LJ, Barnes A, Short AD, Fretwell N, Jones CA, et al. Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. *Tissue Antigens.* 2006;68:467-476.
7. German AJ. Exocrine pancreatic insufficiency in the dog: Breed associations, nutritional considerations, and long-term outcome. *Top Companion Anim Med.* 2012;27:104-108.
8. Wiberg ME, Nurmi AK, Westermarck E. Serum trypsinlike immunoreactivity measurement for the diagnosis of subclinical exocrine pancreatic insufficiency. *J Vet Intern Med.* 1999;13:426-432.

9. Massey J, Rothewell S, Rusbridge C, Tauro A, Addicott D, Chinoy H, et al. Association of an MHC class II haplotype with increased risk of polymyositis in Hungarian Vizsla dogs. PLoS One. 2013;8:e56490. doi: 10.1371/journal.pone.0056490.
10. Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SGE. IPD - the Immuno Polymorphism Database. Nucleic Acids Res. 2013;41(Database issue):D1234-1240.
11. Kennedy LJ, Barnes A, Short A, Brown JJ, Lester S, Seddon J, et al. Canine DLA diversity: 1. New alleles and haplotypes. Tissue Antigens. 2007;69:272-288.

Reagents	Temperature	Time	Cycles
2X Reddy Mix	95°C	5 minutes	1
Sterile water	95°C	30 seconds	
0.4 µM Primer	73°C	1 minute	30
50 ng DNA	72°C	1 minute	
	72°C	5 minutes	1
Total reaction volume: 25 µL	4°C	hold	hold

S1 Table. Reagents and thermal cycling conditions for *DLA-DQB1* amplification are shown.

Appendix B

Supplemental information for Chapter 3

Primer Set	F primer sequence	R primer sequence	Expected size (bp)	Purpose
Exon 6	CGCTGGCAAAAT CCGTGCTTTGGG	GGCAATCTCCT CCTCCAGACCC CC	420	Case 3 & 4 mutation genotyping
Exon 7	ACCTTCTGATTG GCTCCTCT	TTACTTCCTCCC TCATAAACTTG	427	Not present in Case 1
Exon 8	CCACCCACACAC ATACACACA	TGCTCACTTAG GCTGAACACA	441	Not present in Case 1
SGCD Case 1 mut	GTGTCGAGGGTG AAATAAGTGA	ATTAGGGAGGC CAGTCTATTCA	678	Case 1 mutation genotyping
SGCD Case 1 wt	CACATTATGGAG TCTGAGCTAACT	CCATCACTAAA TAGTCTCAGCT GT	434	Case 1 mutation genotyping
Case 1 breakpoint	GGCTTCATAGAT AGTGGGCTG	ATTTCCATCTCA AACTCATTCAA	Wt = 20681 (no amplification) Case 1 affected: 1275	Case 1 breakpoint determination

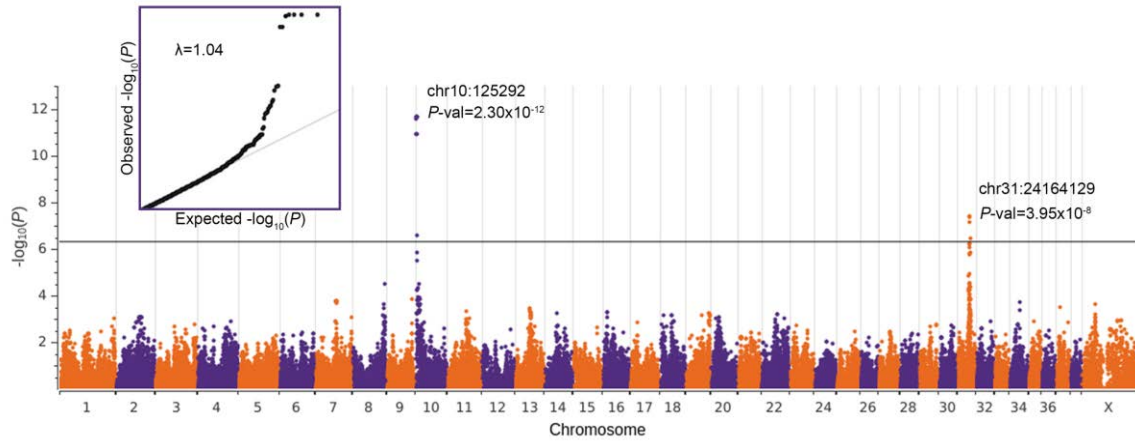
S1 Table. Primers used to define and genotype *SGCD* mutations.

Appendix C

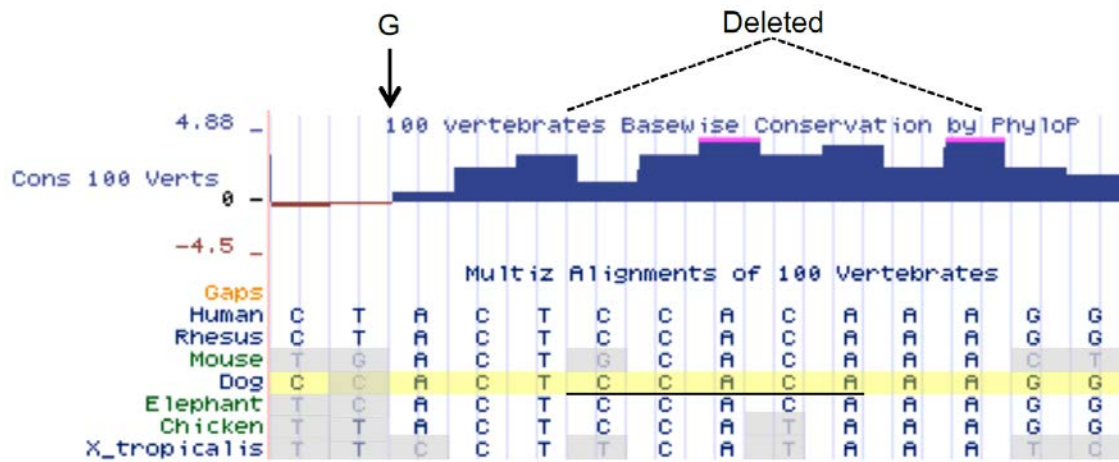
Supplemental information for Chapter 4



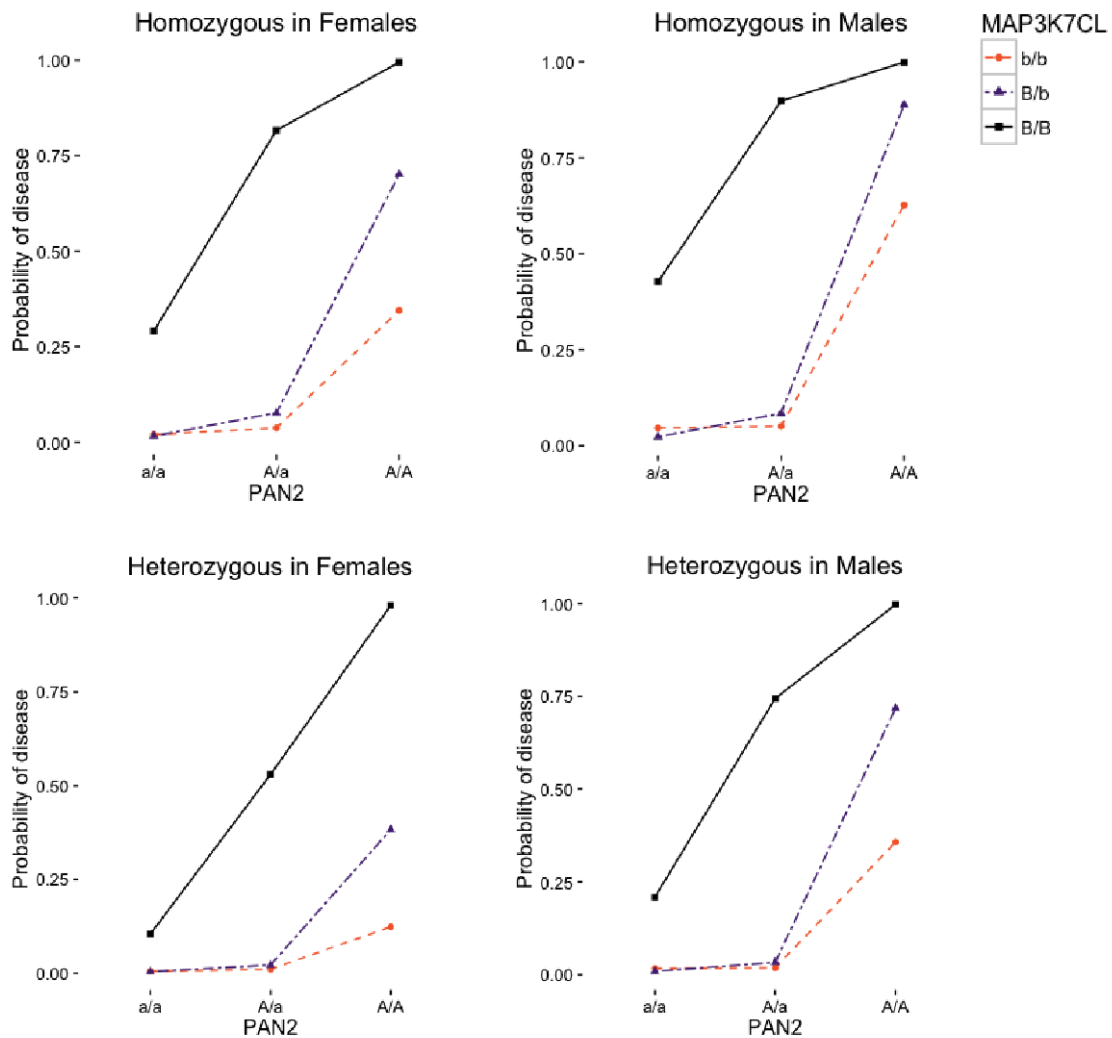
S1 Figure. Clinical presentation of dermatomyositis. Canine dermatomyositis is a vasculopathy that initially manifests as cutaneous lesions across the bony prominences of the face, tail tip, limbs, and feet, shown here. Some dogs develop alopecia and more extensive lesions over time, resulting in dermal scarring associated with erythema and mottled pigmentation.



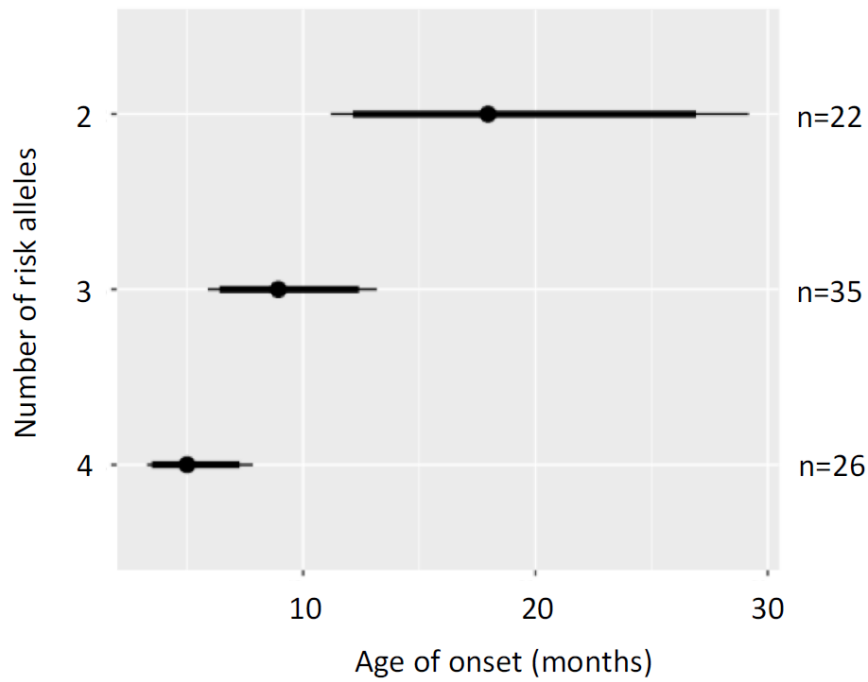
S2 Figure. Manhattan and quantile-quantile plots of combined DMS GWAS (97 cases vs. 68 controls). The $-\log_{10}P$ -values (y-axis) for 98,520 Illumina SNPs are plotted against chromosome position (x-axis). The threshold for Bonferroni significance is shown as a black horizontal line. The P -value and position (canFam3) of the lead SNPs are reported. The Q-Q plot is boxed in purple and plots observed vs. expected Fisher's exact $-\log_{10}P$ -values. The inflation factor (λ) is shown on the Q-Q plot.



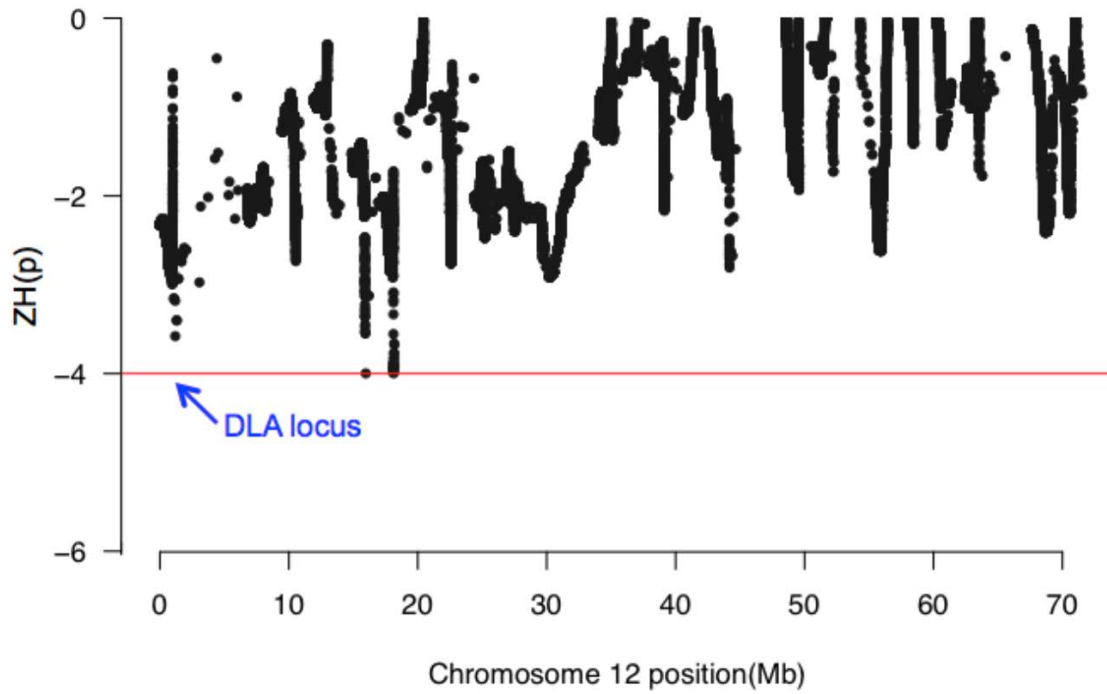
S3 Figure. *MAP3K7CL* indel with conservation and RUNX3 binding motif. UCSC 100 Vertebrates track for human chr21:29,130,846-29,130,860 showing the G insertion and seven base pair deletion created by *MAP3K7CL* indel (Dog c.383_392ACTCCACAAA>GACT). Bases in gray differ from the dog reference sequence. The canine sequence is highlighted in yellow. The RUNX3 binding motif is underlined.



S4 Figure. Three-locus genotype disease probabilities by sex. Probability of disease (y-axis) for all combinations of *PAN2* (x-axis) and *MAP3K7CL* genotypes are plotted in dogs (73 affected and 145 unaffected males; 59 affected and 245 unaffected females) (top) homozygous and (bottom) heterozygous for *DLA-DRB1*002:01*.



S5 Figure. Median age of onset for combinations of *PAN2* and *MAP3K7CL* genotypes. Median age of onset is plotted for genotypes consisting of 2 (AAbb, aaBB, AaBb), 3 (AABb, AaBB), and 4 (AABB) risk alleles at *PAN2* and *MAP3K7CL*. CC and Cc genotypes were combined for analyses. Number of individuals is shown to right.



S6 Figure. Collie selective sweeps on chromosome 12. $ZH(p)$ values for all creeping windows containing 50 or more SNPs are plotted against chromosome position. Creeping windows are ≤ 1 Mb.

COLLIE GWAS			SHETLAND SHEEPDOG GWAS			COMBINED GWAS		
CHR	BASE PAIR	P-VALUE	CHR	BASE PAIR	P-VALUE	CHR	BASE PAIR	P-VALUE
10	2163303	1.47X10 ⁻⁸	31	24107856	1.83X10 ⁻⁹	10	125292	2.30X10 ⁻¹²
10	2125046	4.24X10 ⁻⁸	31	24225299	5.19X10 ⁻⁹	10	319019	2.30X10 ⁻¹²
10	2170946	4.24X10 ⁻⁸	31	24164129	2.52X10 ⁻⁸	10	355241	2.30X10 ⁻¹²
10	2453907	4.24X10 ⁻⁸	31	23978661	2.63X10 ⁻⁸	10	368251	2.30X10 ⁻¹²
10	3757306	4.24X10 ⁻⁸	31	24026411	8.63X10 ⁻⁸	10	100589	2.77X10 ⁻¹²
10	4413901	4.24X10 ⁻⁸	31	24577081	2.01X10 ⁻⁷	10	157505	1.25X10 ⁻¹¹
10	4460808	4.24X10 ⁻⁸	31	24581581	2.01X10 ⁻⁷	10	359293	1.27X10 ⁻¹¹
10	4482172	4.24X10 ⁻⁸	31	24584079	2.01X10 ⁻⁷	31	24164129	3.95X10 ⁻⁸
10	1333693	2.20X10 ⁻⁷	31	23847695	2.20X10 ⁻⁷	31	24026411	4.41X10 ⁻⁸
10	125292	2.71X10 ⁻⁷	10	1720538	2.56X10 ⁻⁷			
10	319019	2.71X10 ⁻⁷	31	22033805	3.33X10 ⁻⁷			
10	355241	2.71X10 ⁻⁷	31	23925418	3.88X10 ⁻⁷			
10	368251	2.71X10 ⁻⁷						
10	2908436	2.71X10 ⁻⁷						
10	4038486	2.71X10 ⁻⁷						
10	4054430	2.71X10 ⁻⁷						
10	4047342	3.21X10 ⁻⁷						

S1 Table. SNPs exceeding Bonferroni significance (5.08X10⁻⁷).

Position	Reference	Alternate
983	C	T
1455	C	G
4302	A	G
4668	T	C
4860	A	G
4940	GATTCTTTCAATTCGATTC	-
5145	C	T
5194	A	G
19674	A	G
19704	G	C
19913	C	T
20162	G	T
20220	T	G
20727	A	C
21565	T	A
24602	G	A
25645	G	A
25646	T	A
25650	G	A
28481	G	A
29441	C	T
29593	C	T
29747	A	G
30885	G	A
31306	G	A
31669	-	G
31689	A	C
31838	T	G
32200	G	C
33320	A	G
33783	-	TGTGTGTG
34721	G	A
35524	AAATAAAT	-AAAT
35632	-	GATT
35919	A	G
35948	A	T
36416	G	A
36580	T	C
38490	A	G
38994	G	A
39186	T	G
39583	T	G
40824	-	G
41681	A	G
42722	-	TTTCTTTCTTTC
46324	G	A
46566	-	ACAT
47636	T	C
48860	G	A
51938	G	A
52373	A	T
53042	A	G
53231	A	T

53630	A	G
53669	A	T
54586	T	A
55154	C	G
55497	C	G
58074	A	G
58163	A	C
58193	C	T
58464	A	G
58750	C	G
58838	-	CC
61325	A	C
61350	C	G
62250	-	AAAT
62307	G	T
62348	C	A
62513	C	T
62583	A	G
65849	TTATTATATCCATTTCTTTTTTTTTTTTTTTTATTA	-
67255	T	C
69315	G	T
70504	A	G
71227	-	A
71367	A	-
71494	G	A
72187	C	G
72799	T	G
73717	C	T
74999	G	A
75146	C	T
76869	A	C
77666	C	T
77841	C	T
77870	-	A
80439	-	GA
80594	T	A
81349	C	T
81424	T	G
82860	T	C
83789	T	C,G
84785	-	AA,AAATAA
84996	A	G
85168	A	-
86938	T	G
87675	CTCTCTCT	-,CTCTCTCTCTCT
87983	T	C
89056	T	G
89486	C	T
90648	-	A
91287	T	C
91288	G	C
92589	T	C
92627	C	T
95042	T	C
95875	A	G
95901	C	T

96173	-	TGTGTG
96654	-	T
97689	C	T
97690	C	T
100589	G	A
103427	T	G
105473	T	C
106427	GT	-
106429	-	ACACACACAC
108574	C	G
108782	A	G
109616	A	G
109776	G	C
114443	T	C
115789	T	C
115831	T	C
116599	A	G
117970	T	C
118024	T	C
119310	A	C
119614	T	C
121835	G	A
121966	-	G
122074	A	G
123135	TTT	-,T
123155	-	T
123784	C	T
124773	G	C
124785	G	T
125199	T	C
125292	A	G
125609	C	T
126124	A	T
126271	T	C
127268	-	C
127461	T	C
128048	G	A
130085	-	C
130214	C	A
130972	C	T
131270	A	G
132736	G	A
133034	-	CT
133275	G	A
133484	C	T
134896	TAAAT	-
135031	G	T
136172	A	G
137970	C	A
138038	T	A
138609	T	C
138649	G	A
138764	A	G
141322	-	C
141332	C	A
143362	-	A

144474	T	A
144542	T	G
146163	G	-
147307	TTTC	-
154392	C	T
155755	A	G
155779	C	A
155975	-	T
157505	A	G
158846	T	C
159056	C	-
159624	C	G
162915	-	AAA
163496	-	AA
163852	G	A
163991	C	T
164263	T	-
164266	-	ATCC
164267	TCC	-
164272	C	T
164273	C	T
164533	A	T
164534	A	C
164794	T	C
165665	C	A
165898	C	T
167713	A	-
168931	-	TAAA
168934	G	A
171502	T	G
171667	G	T
171742	C	T
172944	T	C
173571	-	AAC
174218	A	G
174553	-	A
175817	C	T
177122	A	G
178947	A	C
182833	G	T
183310	CT	-
187227	-	AA
188595	A	-
189225	-	CT
189269	T	A
189284	CC	-
189286	A	T
189791	T	A
190251	T	C
190571	TAAGAG	-
	TTATATCAGGGGATCCCTGGGTGGCGCAGCGGT	
190578	TTGGTGCCTGCCTTTGGCCCAGGGCGCGATCC	-
	TGGAGACCCGGGATCGAATCCC	
191464	T	C
191530	-	TT
191590	TT	-

192779	-	CC,CCCC
192804	C	T
192869	A	-
193366	AAAAGTTTCTAAATGCTAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA	-
194353	C	A
196673	-	GAGA
197178	C	G
197187	T	C
197579	TT	-
198004	CTTTTCTTTTCTTTTCTTT	-
198064	G	A
198073	C	T
198367	T	C
198717	A	C
198929	-	CTCTCTCTCTCTCTCT
199840	T	C
199868	C	T
203558	G	A
204436	T	C
205671	G	A
206024	C	T
206156	T	-
208223	G	A
209994	A	G
210826	A	-
211260	A	T
213933	G	T
219112	C	T
222809	T	G
223827	A	G
223963	T	C
227962	T	A
228189	T	C
230921	T	-
231747	C	G
232975	TAAAAATTT	-
233908	T	C
236429	T	C
236812	-	A
238792	AAG	-
238925	A	T
238934	C	T
239258	C	T
239760	G	A
240699	T	A
240737	-	CAGAGAGG
242049	G	A
244444	CTCT	-
245248	A	-
248834	G	T
249441	AAATAA	-
249645	G	C
251126	TG	-
251348	-	AAAT,AAATAAAT
251368	-	AATA

255125	A	C
255884	T	C
256328	G	C
256801	-	A
258954	-	A
258968	A	G
260700	-	TCTCTC
260805	G	A
263272	A	C
264609	A	-
264826	T	C
264891	CTCT	-
264966	-	A,AA
265620	-	ACATATGCT
265976	C	T
267300	-	GGCTAC
268234	G	C
272766	-	T
273866	A	-
277244	G	A
277609	T	C
279483	T	C
279819	-	T
283554	G	T,GT
284398	CT	-
288586	G	A
289277	C	A
289502	C	T
290259	C	T
292084	G	T
293323	T	G
293324	C	A
294584	A	G
297298	-	TTTTGT
297302	-	G
297305	-	TG
302341	-	ACACAC,ACACACAC
304059	-	ATTT,ATTTATTT
309651	G	C
310721	G	C
313460	T	G
314549	A	-
315259	T	G
315772	G	A
316103	C	G
316273	C	G
317274	TTTA	-,TTTTATTTATTTATTTA
317694	T	C
319019	G	A
319508	A	G
319770	CC	-
319770	C	G
319771	C	A
321002	C	T
326105	G	A
327065	T	C

328251
328758
331135
331768
334964
335222
337919
338118
338323
339097
342843
344853
344919
345210
345919
346090
350369
351679
352840
355241
355949
355953
355969
355973
356507
358754
359293
363034
366583
368251
369609
370205
370561
370976
371592
371755
372114
376253
376823
376826
378943
380460
381048
382635
383033
384401
384806
385406
386021
387325
388036
388893
390674
391084
391262
392287

T
T
A
C
-
-
T
-
-
C
G
A
C
-
A
C
-
-
A
T
G
G
T
T
T
C
A
A
T
A
T
A
T
T
A
-
-
G
C
AA
A
-
T
G
-
-
A
-
G
T
G
A
AAA
-
A
TTA
-
C
C

G
C
G
T
TCTC
A
-
T
T
T
A
G
G
G
CA
G
G
AAAATAAAAT,AAAATAAAATAAAATAAAAT
A
G
G
T
T
T
G
G
A
G
G
C
G
C
C
T
AGAG,AGAGAGAGAG
AA
A
T
-
-
T
C
T
TATTTC
CT
C
T
A
C
C
T
-
T
T
-
A
T
G

393819	A	-
394188	C	T
395522	AGTA	-
397303	C	T
397443	-	A
397839	T	G
397865	C	T
399997	AA	-
400111	T	C
400234	T	A
401025	-	AA
		delTTTCTTTC,TTTCTTTCTTTCTTTCTTTCTTTCT
		TTC,TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTTC
401521	TTTCTTTCTTTCTTTTC	TTTC
401970	A	C
403085	-	T
403733	A	T
407714	GA	-
407821	T	G
408176	G	C
408277	C	T
409137	-	CA
410451	G	C
411186	T	-
411648	A	C
425456	T	A
426511	T	G
435789	G	A
446629	G	A
450399	G	A
463902	T	A
464745	CA	-
468814	-	TTTC
473244	GT	-
479004	G	A
480891	-	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
527236	A	-,AAA
529504	A	T
531522	G	A
539979	A	-
541546	G	T
545297	G	A
553356	C	-
559530	AGA	-
565958	G	C
569295	G	A
572024	G	A
584808	-	A
592283	A	-
592859	G	A
592863	T	A
596255	A	C
600296	ATTTTT	-
603725	-	A
606457	A	G
607719	-	ATTCTTTTT

614684	G	C
621366	G	A
627760	G	A
628057	T	C
644479	C	T
679901	A	-
681353	T	A
684541	AAAGAAAAAGAAAAAGAA	-
685467	-	G
691560	AG	-
705625	ACAC	-,ACACACAC
708173	C	T
709584	T	A
713468	-	A
718727	A	G
721159	A	G
737228	G	-
746264	A	-
753324	G	A
754312	AG	-
754369	C	T
754507	-	ATTTATTT
759008	A	T
766773	C	T
770370	TTTATTTA	-
773874	-	T
773944	AC	-
774278	G	A
774287	-	TA
781273	TAAATAAA	-
791538	ATTT	-
792750	T	-
795887	T	C
800064	G	-
800906	TTGTT	-
801063	-	GA
808263	A	G
810666	C	T
811286	-	T
812076	CT	-,CTCTCT
812789	T	C
816416	A	T
834949	G	A
835850	T	C
838641	A	-
850906	CC	-,C
851959	C	-
855683	T	A
856040	C	T
863353	G	A
867832	G	A
871478	A	-
879196	AATAAATAAATA	-,AATA
883486	C	T
888939	T	-
890576	-	AATGAATAAATAAATAAATA

891964	A	-
893954	T	G
896459	-	AA
906950	-	A
910309	T	-
910992	T	C
917310	-	T
917884	T	A
919616	TTTTAT	-
923362	AAAT	-
926622	T	-
927308	A	G
929039	-	TTAT
933041	TT	-
940501	-	A
940996	A	G
941544	G	A
949170	A	G
962343	TATTTATT	-
964052	G	A
965072	-	TTTA
965128	AGAG	-
971007	TGTTT	-
971895	T	-
974559	A	T
975201	T	C
975203	C	T
975216	C	A
978940	C	T
993657	T	C
1010309	-	CT
1020038	TTC	-
1024911	-	T
1025010	-	AGAG
1031837	C	T
1042098	TATG	-
1051909	T	C
1061145	C	-
1074035	T	C
1090400	-	TTTA,TTTATTTA
1096770	G	A
1096919	A	T,AT
1099689	ATAA	-
1108799	C	T
1111745	-	ATTT
1121749	G	A
1127082	C	T
1128658	-	T
1130331	-	AC
1131360	C	T
1131903	-	C
1135206	A	G
1150305	T	C
1164726	C	T
1165045	-	AG
1168367	T	C

1175038	T	C
1182874	TCTC	-
1195393	-	TC
1207836	-	A
1217285	-	CA
1222163	T	-
1228277	G	C
1239562	G	A
1259335	T	A
1272949	G	A
1286150	C	T
1290104	C	T
1291016	C	T
1296029	G	A
1299356	C	T
1315005	C	A
1319925	G	C
1322218	G	A
1325422	G	A
1333693	G	A

S2 Table. Chromosome 10 variants segregating with the lead SNPs in the affected dogs.

Position	Reference	Alternate
24026411	T	C
24026652	A	C
24026717	-	TAAGAGGAG
24026740	T	C
24026756	G	A
24026797	A	T
24026835	G	A
24026878	C	T
24026886	C	T
24027079	G	T
24027126	T	C
24027219	G	A
24027267	G	A
24027324	T	C
24027430	G	A
24027565	AACT	-
24027640	C	T
24027679	-	CAA
24027791	G	T
24027957	G	T
24028241	A	G
24028305	G	-
24028334	G	A
24028339	C	A
24028560	-	AAA
24028649	-	GA
24028659	T	C
24028900	A	G
24028918	C	A
24029003	C	T
24029289	C	T
24029465	A	C
24029536	G	T
24029568	T	-
24029571	-	ATAAATAAATAAATAAAT
24029606	AA	-
24029785	T	C
24029985	C	T
24030534	C	T
24030579	C	-
24030609	T	A
24030637	C	G

S3 Table. Chromosome 31 variants segregating with the lead SNPs in the affected Shetland sheepdog. The first page of the table is shown; the remaining 92 pages can be found at <http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006604#sec021>

Genotype	Cases (n=132)	Controls (n=390)	OR	95%CI	P-value
AA*	84	47	12.77	8.001-20.39	1.81X10 ⁻²⁹
Aa	40	175	0.53	0.35-0.81	0.0041
aa	8	168	0.085	0.041-0.18	5.40X10 ⁻¹⁵
No. of A alleles (2n)	208	269	7.042	5.025-10	2.08X10 ⁻³⁵
BB†	68	23	16.95	9.86-29.16	8.26X10 ⁻²⁹
Bb	29	126	0.59	0.37-0.94	0.0274
bb	35	241	0.22	0.14-0.35	1.61X10 ⁻¹²
No. of B alleles (2n)	165	172	5.88	4.31-8.065	1.45X10 ⁻³³

*A=the *PAN2* variant encoding p.Arg492Cys; a=wild type

†B=*MAP3K7CL* c.383_392ACTCCACAAA>GACT; b=wild type

S4 Table. Frequency of *PAN2* and *MAP3K7CL* genotypes.

	BREED	GENOTYPE		
		PAN2 R492C	MAP3K7CL indel	DLA-DRB1
1	AMERICAN BULLDOG	GG	wt	001:01/015:01
2	AMERICAN BULLDOG	GG	wt	001:01/uncharacterized
3	AMERICAN BULLDOG	GG	wt	015:02/015:02
4	AMERICAN BULLDOG	GG	wt	018:01/018:01
5	AMERICAN BULLDOG	GG	wt	018:01/018:01
6	AMERICAN ESKIMO	GG	wt	015:01/015:02
7	AMERICAN ESKIMO	GG	wt	015:01/040:01
8	AMERICAN ESKIMO	AG	wt	015:01/040:01
9	AMERICAN ESKIMO	GG	wt	015:01/040:01
10	AMERICAN ESKIMO	GG	wt	013:01/040:01
11	AUSTRALIAN SHEPHERD	GG	wt	002:01/046:01
12	AUSTRALIAN SHEPHERD	GG	indel/wt	026:01/030:01
13	AUSTRALIAN SHEPHERD	GG	wt	006:01/011:01
14	AUSTRALIAN SHEPHERD	GG	wt	002:01/011:01
15	AUSTRALIAN SHEPHERD	GG	wt	011:01/015:01
16	AUSTRALIAN SHEPHERD	GG	wt	001:01/015:01
17	BASENJI	GG	wt	009:01/030:01
18	BASENJI	GG	wt	009:01/027:01
19	BASENJI	GG	wt	004:01/027:01
20	BASENJI	GG	wt	009:01/022:01
21	BASENJI	GG	wt	uncharacterized
22	BORDER COLLIE	GG	wt	013:01/023:01
23	BORDER COLLIE	GG	wt	013:01/013:01
24	BORDER COLLIE	GG	wt	015:01/023:01
25	BORDER COLLIE	GG	wt	013:01/013:01
26	BORDER COLLIE	GG	wt	013:01/018:01
27	BOSTON TERRIER	GG	wt	002:0*/015:0*
28	BOSTON TERRIER	GG	wt	002:01/002:01
29	BOSTON TERRIER	GG	indel/wt	002:01/002:01
30	BOSTON TERRIER	GG	indel/wt	002:01/002:01
31	BOSTON TERRIER	GG	indel/wt	002:01/002:01
32	BOSTON TERRIER	GG	wt	002:01/002:01
33	BOSTON TERRIER	GG	wt	002:01/002:01
34	BOSTON TERRIER	GG	wt	001:01/002:01
35	BOSTON TERRIER	GG	wt	002:01/023:01
36	BOXER	GG	wt	004:01/015:02
37	BOXER	GG	wt	001:01/001:01
38	BOXER	GG	wt	015:01/uncharacterized
39	BOXER	GG	wt	001:01/020:01
40	BOXER	GG	wt	012:01/012:01

41	CAIRN TERRIER	AA	wt	009:01/015:02
42	CAIRN TERRIER	AG	indel/wt	015:01/015:01
43	CAIRN TERRIER	AG	wt	009:01/012:01
44	CAIRN TERRIER	GG	wt	015:02/015:02
45	CAIRN TERRIER	AA	wt	015:01/015:01
46	CAIRN TERRIER	AG	wt	009:01/015:02
47	CAIRN TERRIER	AA	indel/wt	015:01/015:02
48	CAIRN TERRIER	AG	indel/wt	009:01/015:02
49	CAIRN TERRIER	AG	wt	009:01/015:02
50	CAIRN TERRIER	AA	wt	015:02/018:01
51	CAIRN TERRIER	AA	wt	015:01/015:02
52	CAIRN TERRIER	AA	indel/wt	015:01/015:02
53	CAIRN TERRIER	AG	wt	001:01/001:01
54	CAIRN TERRIER	AG	wt	002:01/015:02
55	CAIRN TERRIER	AA	indel/wt	009:01/015:01
56	CAIRN TERRIER	AA	indel/wt	001:01/009:01
57	CAIRN TERRIER	AA	indel/wt	015:01/015:02
58	CAIRN TERRIER	AA	wt	001:01/015:02
59	CARDIGAN WELSH CORGI	AG	indel/wt	015:02/015:02
60	CARDIGAN WELSH CORGI	GG	wt	015:02/015:02
61	CARDIGAN WELSH CORGI	GG	indel/wt	020:01/020:01
62	CARDIGAN WELSH CORGI	GG	wt	015:02/015:02
63	CARDIGAN WELSH CORGI	AG	wt	002:01/015:02
64	CARDIGAN WELSH CORGI	GG	wt	002:01/002:01
65	CARDIGAN WELSH CORGI	AG	wt	002:01/015:02
66	CARDIGAN WELSH CORGI	GG	indel/wt	002:01/018:01
67	CARDIGAN WELSH CORGI	AA	indel/wt	002:01/002:01
68	CARDIGAN WELSH CORGI	GG	wt	002:01/015:02
69	CARDIGAN WELSH CORGI	GG	indel/wt	002:01/002:01
70	CARDIGAN WELSH CORGI	GG	wt	015:02/020:01
71	CARDIGAN WELSH CORGI	AG	wt	002:01/002:01
72	CATAHOULA	GG	wt	001:01/015:01
73	CATAHOULA	GG	wt	001:01/015:01
74	CATAHOULA	GG	wt	001:01/001:01
75	CATAHOULA	GG	wt	026:01/030:01
76	CATAHOULA	GG	wt	026:01/030:01
77	CAVALIER KING CHARLES SPANIEL	GG	wt	011:01/011:01
78	CAVALIER KING CHARLES SPANIEL	GG	wt	009:01/011:01
79	CAVALIER KING CHARLES SPANIEL	GG	indel/indel	011:01/011:01
80	CAVALIER KING CHARLES SPANIEL	GG	wt	006:01/006:01
81	CAVALIER KING CHARLES SPANIEL	GG	indel/indel	006:01/020:01
82	CHIHUAHUA	AA	wt	uncharacterized
83	CHIHUAHUA	AG	wt	002:01/009:01
84	CHIHUAHUA	AG	wt	001:01/002:01
85	CHIHUAHUA	AG	wt	015:01/015:01

86	CHIHUAHUA	GG	wt	001:01/009:01
87	DACHSHUND	GG	wt	009:01/015:02
88	DACHSHUND	GG	wt	001:01/015:01
89	DACHSHUND	GG	wt	002:01/009:01
90	DACHSHUND	GG	wt	002:0*015:0*
91	DACHSHUND	GG	wt	001:01/009:01
92	DALMATIAN	GG	wt	002:01/006:01
93	DALMATIAN	GG	wt	006:01/020:01
94	DALMATIAN	GG	indel/wt	002:01/020:01
95	DALMATIAN	GG	wt	002:01/037:01
96	DALMATIAN	GG	wt	002:01/020:01
97	ENGLISH BULLDOG	GG	wt	015:01/015:02
98	ENGLISH BULLDOG	GG	wt	018:01/018:01
99	ENGLISH BULLDOG	GG	indel/wt	013:01/015:02
100	ENGLISH BULLDOG	GG	wt	013:01/018:01
101	ENGLISH BULLDOG	GG	wt	002:01/015:02
102	ENGLISH COCKER SPANIEL	GG	indel/indel	001:01/006:01
103	ENGLISH COCKER SPANIEL	GG	indel/wt	001:01/001:01
104	ENGLISH COCKER SPANIEL	GG	indel/wt	006:01/006:01
105	ENGLISH COCKER SPANIEL	GG	indel/indel	006:01/006:01
106	ENGLISH COCKER SPANIEL	GG	indel/wt	006:01/006:01
107	ENGLISH COCKER SPANIEL	GG	indel/indel	006:01/006:01
108	ENGLISH COCKER SPANIEL	GG	indel/wt	006:01/006:01
109	ENGLISH COCKER SPANIEL	GG	indel/indel	006:01/006:01
110	ENGLISH SETTER	GG	wt	001:01/001:01
111	ENGLISH SETTER	GG	wt	001:01/001:01
112	ENGLISH SETTER	GG	wt	001:01/001:01
113	ENGLISH SETTER	GG	wt	001:01/001:01
114	ENGLISH SETTER	GG	wt	001:01/001:01
115	ENGLISH SPRINGER SPANIEL	GG	indel/wt	015:01/015:01
116	ENGLISH SPRINGER SPANIEL	GG	indel/indel	012:01/015:01
117	ENGLISH SPRINGER SPANIEL	GG	indel/wt	015:01/015:01
118	ENGLISH SPRINGER SPANIEL	GG	wt	015:01/015:01
119	ENGLISH SPRINGER SPANIEL	GG	wt	012:01/012:01
120	FOX TERRIER	GG	wt	001:01/015:01
121	FOX TERRIER	AG	wt	013:01/013:01
122	FOX TERRIER	GG	wt	001:01/001:01
123	FOX TERRIER	AG	indel/wt	013:01/013:01
124	FOX TERRIER	GG	wt	001:01/001:01
125	FOX TERRIER	GG	wt	001:01/001:01
126	FOX TERRIER	GG	wt	001:01/013:01
127	FOX TERRIER	AG	wt	001:01/001:01
128	FOX TERRIER	AG	indel/wt	013:01/013:01
129	FOX TERRIER	GG	wt	001:01/001:01
130	FOX TERRIER	AA	wt	026:01/030:01

131	FOX TERRIER	GG	wt	001:01/001:01
132	FOX TERRIER	AA	indel/wt	013:01/013:01
133	FOX TERRIER	GG	wt	001:01/018:01
134	FOX TERRIER	AG	wt	013:01/013:01
135	FOX TERRIER	GG	indel/wt	013:01/013:01
136	FOX TERRIER	GG	wt	001:01/001:01
137	FOX TERRIER	GG	wt	001:01/001:01
138	FOX TERRIER	AG	wt	008:01/013:01
139	FOX TERRIER	GG	wt	001:01/001:01
140	FOX TERRIER	GG	wt	001:01/001:01
141	FOX TERRIER	GG	wt	001:01/001:01
142	FOX TERRIER	AG	wt	001:01/001:01
143	FOX TERRIER	GG	wt	001:01/015:02
144	FOX TERRIER	GG	wt	001:01/018:01
145	FOX TERRIER	GG	wt	001:01/001:01
146	FOX TERRIER	GG	wt	001:01/013:01
147	GERMAN SHEPHERD DOG	GG	wt	001:01/002:01
148	GERMAN SHEPHERD DOG	GG	wt	001:02/015:01
149	GERMAN SHEPHERD DOG	GG	wt	001:01/011:01
150	GERMAN SHEPHERD DOG	GG	wt	001:02/015:01
151	GERMAN SHEPHERD DOG	GG	wt	011:01/015:01
152	GOLDEN RETRIEVER	GG	wt	006:01/012:01
153	GOLDEN RETRIEVER	GG	wt	012:01/012:01
154	GOLDEN RETRIEVER	GG	wt	006:01/015:01
155	GOLDEN RETRIEVER	GG	wt	006:01/012:01
156	GOLDEN RETRIEVER	GG	wt	012:01/012:01
157	GREAT DANE	GG	indel/wt	001:01/006:01
158	GREAT DANE	GG	wt	001:01/001:01
159	GREAT DANE	GG	wt	001:01/012:01
160	GREAT DANE	GG	wt	001:01/012:01
161	GREAT DANE	GG	wt	001:01/001:01
162	IRISH SETTER	GG	wt	001:01/008:01
163	IRISH SETTER	GG	wt	001:01/001:01
164	IRISH SETTER	GG	wt	001:01/005:01
165	IRISH SETTER	GG	indel/indel	001:01/005:01
166	IRISH SETTER	GG	wt	005:01/006:01
167	JACK RUSSELL TERRIER	AA	wt	015:01/015:02
168	JACK RUSSELL TERRIER	AG	wt	002:01/008:01
169	JACK RUSSELL TERRIER	AA	wt	001:01/002:01
170	JACK RUSSELL TERRIER	AA	wt	011:01/023:01
171	JACK RUSSELL TERRIER	AG	wt	009:01/009:01
172	JACK RUSSELL TERRIER	AA	wt	002:01/015:02
173	JACK RUSSELL TERRIER	GG	wt	013:01/020:01
174	JACK RUSSELL TERRIER	AA	wt	002:01/009:01
175	JACK RUSSELL TERRIER	AG	wt	011:01/038:01

176	JACK RUSSELL TERRIER	AG	wt	013:01/015:01
177	JACK RUSSELL TERRIER	GG	wt	006:01/011:01
178	JACK RUSSELL TERRIER	GG	wt	015:01/020:01
179	JACK RUSSELL TERRIER	GG	wt	009:01/009:01
180	JACK RUSSELL TERRIER	AG	wt	013:01/015:01
181	LABRADOR RETRIEVER	GG	wt	006:01/012:01
182	LABRADOR RETRIEVER	GG	wt	008:02/015:02
183	LABRADOR RETRIEVER	GG	wt	001:01/012:01
184	LABRADOR RETRIEVER	GG	indel/indel	008:02/012:01
185	LABRADOR RETRIEVER	GG	wt	012:01/015:02
186	LABRADOR RETRIEVER	GG	indel/wt	012:01/012:01
187	LABRADOR RETRIEVER	GG	indel/wt	012:01/012:01
188	LABRADOR RETRIEVER	GG	indel/wt	002:01/046:01
189	LABRADOR RETRIEVER	GG	wt	001:01/012:01
190	LABRADOR RETRIEVER	GG	wt	008:02/020:01
191	LABRADOR RETRIEVER	GG	wt	015:02/015:02
192	LABRADOR RETRIEVER	GG	wt	012:01/015:02
193	LABRADOR RETRIEVER	GG	wt	008:02/020:01
194	LABRADOR RETRIEVER	GG	wt	012:01/015:02
195	PEMBROKE WELSH CORGI	AG	wt	012:01/015:01
196	PEMBROKE WELSH CORGI	AG	wt	012:01/015:01
197	PEMBROKE WELSH CORGI	AG	wt	002:01/002:01
198	PEMBROKE WELSH CORGI	AG	wt	012:01/015:01
199	PEMBROKE WELSH CORGI	GG	wt	012:01/018:01
200	PEMBROKE WELSH CORGI	GG	wt	002:01/015:01
201	PEMBROKE WELSH CORGI	GG	wt	002:01/015:01
202	PEMBROKE WELSH CORGI	AG	wt	002:01/015:01
203	PEMBROKE WELSH CORGI	AG	wt	012:01/015:01
204	PEMBROKE WELSH CORGI	AG	wt	012:01/015:01
205	POODLE	GG	wt	001:01/015:01
206	POODLE	GG	wt	015:01/015:01
207	POODLE	GG	wt	015:01/015:01
208	POODLE	GG	wt	015:01/015:01
209	POODLE	GG	indel/wt	015:01/020:01
210	POODLE	GG	wt	015:01/015:01
211	SCHNAUZER	GG	wt	001:01/013:01
212	SCHNAUZER	GG	wt	009:01/009:01
213	SCHNAUZER	GG	wt	001:01/009:01
214	SCHNAUZER	GG	indel/wt	009:01/015:02
215	SCHNAUZER	GG	wt	009:01/009:01
216	WEST HIGHLAND WHITE TERRIER	AG	wt	001:01/015:02
217	WEST HIGHLAND WHITE TERRIER	AA	wt	001:01/015:01
218	WEST HIGHLAND WHITE TERRIER	AG	wt	001:01/001:01
219	WEST HIGHLAND WHITE TERRIER	GG	wt	001:01/015:01
220	WEST HIGHLAND WHITE TERRIER	GG	wt	001:01/015:01

221	YORKSHIRE TERRIER	AA	wt	006:01/006:01
222	YORKSHIRE TERRIER	AG	wt	006:01/015:02
223	YORKSHIRE TERRIER	AG	wt	006:01/006:01
224	YORKSHIRE TERRIER	AA	wt	006:01/006:01
225	YORKSHIRE TERRIER	GG	wt	006:01/015:02
226	YORKSHIRE TERRIER	AG	wt	015:02/015:02
227	YORKSHIRE TERRIER	AG	wt	006:01/006:01
228	YORKSHIRE TERRIER	AG	wt	015:02/015:02
229	YORKSHIRE TERRIER	GG	wt	015:02/015:02

S5 Table. Three-locus genotypes for 229 individuals of 30 breeds.

Chr	Position	Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Genotyping method
10	565958	ANKRD52	GGAGTCTGGGTAGTGAGCCT	GATGCCCTATGGACAGTGAGTG	Restriction digest: <i>Bam</i> HI
10	627760	PAN2	ANGMAGGCCAMGGCAGTAGG	ACCGACTCAGGGAATCAGACAG	Restriction digest: <i>Acl</i> I
10	1239562	STAT6	CCTCTTCACCCACCCTGCTC	CGTAGTTCCCTGCCCTCTCACC	Sanger sequencing
31	24132273	MAP3K7C1	TGGAGAGGCCAMGGAMAGGA	CAACCAAGAACATCTGTCCGG	Restriction digest: <i>Hpy</i> I 88I
31	24132343	MAP3K7C1	CTGGAMTAGACACAGACACCGA	CCTAACCCGCCCTTGCCTT	Sanger sequencing

S6 Table. Primers and genotyping method for variants.

Appendix D

Permission to reprint published work

**SPRINGER LICENSE
TERMS AND CONDITIONS**

May 31, 2017

This Agreement between Jacquelyn Evans ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4119570665583
License date	May 31, 2017
Licensed Content Publisher	Springer
Licensed Content Publication	Mammalian Genome
Licensed Content Title	Exome sequencing reveals a nebulin nonsense mutation in a dog model of nemaline myopathy
Licensed Content Author	Jacquelyn M. Evans
Licensed Content Date	Jan 1, 2016
Licensed Content Volume	27
Licensed Content Issue	9
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of your thesis / dissertation	Identification of genetic variants underlying simple and complex canine myopathies
Expected completion date	Jun 2017
Estimated size(pages)	150
Requestor Location	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Billing Type	Invoice
Billing Address	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Total	0.00 USD
Terms and Conditions	

Introduction

The publisher for this copyrighted material is Springer. By clicking "accept" in connection

with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

Limited License

With reference to your request to reuse material on which Springer controls the copyright, permission is granted for the use indicated in your enquiry under the following conditions:

- Licenses are for one-time use only with a maximum distribution equal to the number stated in your request.

- Springer material represents original material which does not carry references to other sources. If the material in question appears with a credit to another source, this permission is not valid and authorization has to be obtained from the original copyright holder.

- This permission

- is non-exclusive

- is only valid if no personal rights, trademarks, or competitive products are infringed.

- explicitly excludes the right for derivatives.

- Springer does not supply original artwork or content.

- According to the format which you have selected, the following conditions apply accordingly:

- **Print and Electronic:** This License include use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.

- **Print:** This License excludes use in electronic form.

- **Electronic:** This License only pertains to use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.

For any electronic use not mentioned, please contact Springer at permissions.springer@spi-global.com.

- Although Springer controls the copyright to the material and is entitled to negotiate on rights, this license is only valid subject to courtesy information to the author (address is given in the article/chapter).

- If you are an STM Signatory or your work will be published by an STM Signatory and you are requesting to reuse figures/tables/illustrations or single text extracts, permission is granted according to STM Permissions Guidelines: <http://www.stm-assoc.org/permissions-guidelines/>

For any electronic use not mentioned in the Guidelines, please contact Springer at permissions.springer@spi-global.com. If you request to reuse more content than stipulated in the STM Permissions Guidelines, you will be charged a permission fee for the excess content.

Permission is valid upon payment of the fee as indicated in the licensing process. If permission is granted free of charge on this occasion, that does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

-If your request is for reuse in a Thesis, permission is granted free of charge under the following conditions:

This license is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published, permission needs to be reobtained.

- includes use in an electronic form, provided it is an author-created version of the thesis on his/her own website and his/her university's repository, including UMI (according to the definition on the Sherpa website: <http://www.sherpa.ac.uk/romeo/>);

- is subject to courtesy information to the co-author or corresponding author.

Geographic Rights: Scope

Licenses may be exercised anywhere in the world.

Altering/Modifying Material: Not Permitted

Figures, tables, and illustrations may be altered minimally to serve your work. You may not alter or modify text in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s).

Reservation of Rights

Springer reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction and (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

License Contingent on Payment

While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Springer or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received by the date due, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Springer reserves the right to take any and all action to protect its copyright in the materials.

Copyright Notice: Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material:

"Springer book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), (original copyright notice as given in the publication in which the material was originally published) "With permission of Springer"

In case of use of a graph or illustration, the caption of the graph or illustration must be included, as it is indicated in the original publication.

Warranties: None

Springer makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity

You hereby indemnify and agree to hold harmless Springer and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you without Springer's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer, by CCC on Springer's behalf).

Objection to Contrary Terms

Springer hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other conditions:

V 12AUG2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

**SPRINGER LICENSE
TERMS AND CONDITIONS**

May 31, 2017

This Agreement between Jacquelyn Evans ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4119570853439
License date	May 31, 2017
Licensed Content Publisher	Springer
Licensed Content Publication	Mammalian Genome
Licensed Content Title	Exome sequencing reveals a nebulin nonsense mutation in a dog model of nemaline myopathy
Licensed Content Author	Jacquelyn M. Evans
Licensed Content Date	Jan 1, 2016
Licensed Content Volume	27
Licensed Content Issue	9
Type of Use	Thesis/Dissertation
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	4
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Original figure numbers	Figures 1,2,3 and Table 1
Title of your thesis / dissertation	Identification of genetic variants underlying simple and complex canine myopathies
Expected completion date	Jun 2017
Estimated size(pages)	150
Requestor Location	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Billing Type	Invoice
Billing Address	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Total	0.00 USD
Terms and Conditions	

Introduction

The publisher for this copyrighted material is Springer. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

Limited License

With reference to your request to reuse material on which Springer controls the copyright, permission is granted for the use indicated in your enquiry under the following conditions:

- Licenses are for one-time use only with a maximum distribution equal to the number stated in your request.

- Springer material represents original material which does not carry references to other sources. If the material in question appears with a credit to another source, this permission is not valid and authorization has to be obtained from the original copyright holder.

- This permission

- is non-exclusive

- is only valid if no personal rights, trademarks, or competitive products are infringed.

- explicitly excludes the right for derivatives.

- Springer does not supply original artwork or content.

- According to the format which you have selected, the following conditions apply accordingly:

- **Print and Electronic:** This License include use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.

- **Print:** This License excludes use in electronic form.

- **Electronic:** This License only pertains to use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.

For any electronic use not mentioned, please contact Springer at permissions.springer@spi-global.com.

- Although Springer controls the copyright to the material and is entitled to negotiate on rights, this license is only valid subject to courtesy information to the author (address is given in the article/chapter).

- If you are an STM Signatory or your work will be published by an STM Signatory and you are requesting to reuse figures/tables/illustrations or single text extracts, permission is granted according to STM Permissions Guidelines: <http://www.stm-assoc.org/permissions-guidelines/>

For any electronic use not mentioned in the Guidelines, please contact Springer at permissions.springer@spi-global.com. If you request to reuse more content than stipulated in the STM Permissions Guidelines, you will be charged a permission fee for the excess content.

Permission is valid upon payment of the fee as indicated in the licensing process. If permission is granted free of charge on this occasion, that does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

- If your request is for reuse in a Thesis, permission is granted free of charge under the following conditions:

This license is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published, permission needs to be reobtained.

- includes use in an electronic form, provided it is an author-created version of the thesis on his/her own website and his/her university's repository, including UMI (according to the definition on the Sherpa website: <http://www.sherpa.ac.uk/romeo/>);

- is subject to courtesy information to the co-author or corresponding author.

Geographic Rights: Scope

Licenses may be exercised anywhere in the world.

Altering/Modifying Material: Not Permitted

Figures, tables, and illustrations may be altered minimally to serve your work. You may not alter or modify text in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s).

Reservation of Rights

Springer reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction and (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

License Contingent on Payment

While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Springer or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received by the date due, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Springer reserves the right to take any and all action to protect its copyright in the materials.

Copyright Notice: Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material:

"Springer book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), (original copyright notice as given in the publication in which the material was originally published) "With permission of Springer"

In case of use of a graph or illustration, the caption of the graph or illustration must be included, as it is indicated in the original publication.

Warranties: None

Springer makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity

You hereby indemnify and agree to hold harmless Springer and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you without Springer's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer, by CCC on Springer's behalf).

Objection to Contrary Terms

Springer hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations

5/31/2017

RightsLink Printable License

established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other conditions:

V 12AUG2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

<https://s100.copyright.com/AppDispatchServlet>

4/4

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

May 31, 2017

This Agreement between Jacquelyn Evans ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4119570450249
License date	May 31, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Animal Genetics
Licensed Content Title	Association of DLA-DQB1 alleles with exocrine pancreatic insufficiency in Pembroke Welsh Corgis
Licensed Content Author	J. M. Evans, K. L. Tsai, A. N. Starr-Moss, J. M. Steiner, L. A. Clark
Licensed Content Date	Jun 19, 2015
Licensed Content Pages	4
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Identification of genetic variants underlying simple and complex canine myopathies
Expected completion date	Jun 2017
Expected size (number of pages)	150
Requestor Location	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Jacquelyn Evans 131 Freedom Dr. Apt. C CLEMSON, SC 29631 United States Attn: Jacquelyn Evans
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order**, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. **For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts**, You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.
- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding

("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.
- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.
- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License \(CC-BY-NC-ND\)](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library

<http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.