G3: Genes|Genomes|Genetics Early Online, published on October 7, 2015 as doi:10.1534/g3.115.021923

A nonsense variant in *COL6A1* in Landseer dogs with muscular dystrophy

Frank Steffen[1¶], Thomas Bilzer[2¶], Jan Brands[2], Lorenzo Golini[1], Vidhya Jagannathan[3], Michaela Wiedmer[3], Michaela Drögemüller[3], Cord Drögemüller[3], Tosso Leeb[3,*]

[1] Neurology Service, Department of Small Animals, Vetsuisse Faculty, University of Zurich, Winterthurerstrasse 260, 8057 Zurich, Switzerland

[2] Institute of Neuropathology, University Hospital Düsseldorf, Moorenstraße 5, 40225 Düsseldorf, Germany

[3] Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland

*Corresponding author

E-mail: Tosso.Leeb@vetsuisse.unibe.ch (TL)

¶These authors contributed equally to this work.

source: https://doi.org/10.7892/boris.72278 | downloaded: 12.10.2024

© The Author(s) 2013. Published by the Genetics Society of America.

Abstract

A novel canine muscular dystrophy in Landseer dogs was observed. We had access to five affected dogs from two litters. The clinical signs started at a few weeks of age and the severe progressive muscle weakness led to euthanasia between 5 and 15 months of age. The pedigrees of the affected dogs suggested a monogenic autosomal recessive inheritance of the trait. Linkage and homozygosity mapping indicated two potential genome segments for the causative variant on chromosomes 10 and 31 harboring a total of 4.8 Mb of DNA or 0.2% of the canine genome. Using the illumina sequencing technology we obtained a whole genome sequence from one affected Landseer. Variants were called with respect to the dog reference genome and compared to the genetic variants of 170 control dogs from other breeds. The affected Landseer dog was homozygous for a single private non-synonymous variant in the critical intervals, a nonsense variant in the *COL6A1* gene (Chr31:39,303,964G>T; *COL6A1:c.289G>T*; p.E97*). Genotypes at this variant showed perfect concordance with the muscular dystrophy phenotype in all five cases and more than one thousand control dogs. Variants in the human *COL6A1* gene cause Bethlem myopathy or Ullrich congenital muscular dystrophy. We therefore conclude that the identified canine *COL6A1* variant is most likely causative for the observed muscular dystrophy in Landseer dogs. Based on the nature of the genetic variant in Landseer dogs and their severe clinical phenotype these dogs represent a model for human Ullrich congenital muscular dystrophy.

Introduction

Muscular dystrophies are a heterogeneous group of inherited, degenerative disorders characterized by progressive muscular dysfunction (Flanigan 2012). Duchenne muscular dystrophy caused by variants in the X-chromosomal *DMD* gene and a complete loss of function of dystrophin is the text book representative for these entities (Koenig *et al.* 1987). However, although many of the muscular dystrophies are due to single gene defects, there is considerable variance in the phenotype (age of onset, severity and localization of clinical symptoms, rate of progression, etc.). This complexity arises from the fact that many genes are required for the structural integrity and proper function of skeletal muscle and that different genetic variants within a given gene can lead to quite distinct phenotypes. This again can be seen with the *DMD* gene: Variants that lead to a complete loss of function cause the severe Duchenne muscular dystrophy whereas variants that only partially affect dystrophin function result in the milder Becker muscular dystrophy (Hart *et al.* 1987; Koenig *et al.* 1989). The dystrophin protein interacts with many other skeletal muscle proteins and variants in genes encoding these other dystrophin-associated proteins can lead to very similar clinical phenotypes.

Another group of muscular dystrophies with a much less well understood pathophysiology are the collagen VI myopathies. Collagen VI is a ubiquitously expressed extracellular matrix protein, composed of two chains of 140-150 kDa each, named alpha1(VI) and alpha2(VI), and one larger chain of 240-300 kDa named alpha3(VI) (Furthmayr *et al.* 1983). The three chains fold together into a triple-helical collagen VI molecule (monomer) that further assembles into dimers and tetramers. The tetramers are secreted into the extracellular space where they form collagen VI beaded microfibrils. In muscle, the microfibrillar network of collagen VI surrounds the basement membrane of fibers, binding components of the extracellular matrix and transferring mechanical and biochemical signals from the extracellular matrix to the muscle cell (De Palma *et al.* 2014). The three subunits of collagen VI are encoded by the *COL6A1*, *COL6A2*, and *COL6A3* genes. Variants in any of these genes can cause the severe Ullrich muscular dystrophy or the much milder Bethlem myopathy (Jöbsis *et al.* 1996; Camacho Vanegas *et al.* 2001). Many different variants have been identified in human patients and it is becoming increasingly clear that the phenotypic spectrum of these collagen VI myopathies is almost continuous depending on the specific genetic variants (Briñas *et al.* 2010). A complete list of all known functional human variants in the collagen VI genes is compiled at the Leiden Muscular Dystrophy pages (http://www.dmd.nl/).

Dogs have been recognized as valuable models for many human hereditary diseases and the OMIA database currently lists ~350 potential canine models for human diseases. In roughly 200 of these, the causative genetic defect is known (Nicholas and Hobbs 2014; http://omia.angis.org.au/home/). Different muscular dystrophies have also been described in dogs. Golden Retrievers with a Duchenne-like muscular dystrophy due to a *DMD* variant are an extremely valuable animal model as *Dmd* mutant mice do not recapitulate the severe human (and canine) phenotypes (Cooper *et al.* 1988; Valentine *et al.* 1988; Valentine *et al.* 1992). Many other cases of muscular dystrophies in dogs and other companion animals have been reported.

Recently, the case of a myopathic Labrador Retriever was reported that showed greatly reduced or absent sarcolemmal collagen VI expression (Marioni-Henry *et al.* 2014). Collagen VI expression in this dog was still detectable by antibody staining in the endomysium. The clinical phenotype of this Labrador Retriever resembled an intermediate severity when compared with human Bethlem myopathy or Ullrich congenital muscular dystrophy. As the genetic defect in this Labrador Retriever is not known, it remains to be determined whether this represents a primary collagen VI defect or whether the reduction in collagen VI is a secondary effect of some other genetic lesion. Therefore, without the conclusive identification of the underlying causative genetic defect the value of these reports for our basic understanding of muscle function and a comparative analysis with respect to the human diseases remains limited.

In this report, we describe a novel canine muscular dystrophy, which occurred in Landseer dogs. We also provide an initial limited characterization of the clinical and histopathological phenotype. The main focus of this study is on the identification of the presumed causative genetic defect by a positional cloning approach in combination with whole genome sequencing of an affected dog.

Results

Clinical presentation and laboratory findings

We investigated a litter of 7 Landseer dogs. Two male and two female puppies were affected by progressive muscle weakness. One male puppy had been euthanized due to severe clinical signs in the first weeks of life without any ancillary diagnostic tests or necropsy. We examined the remaining 3 affected littermates, which were presented because of reluctance to move and prolonged sleeping episodes compared to their 3 normal siblings during the first few weeks of life.

The male and the two female cases were clinically and neurologically examined at 14 and 18 weeks of age. No abnormalities were found on general physical examination of the male dog. On neurological evaluation, the male affected dog was mentally alert but only able to ambulate for a few steps before lying down. He had a short strided gait in the thoracic and pelvic limbs and while standing his back was arched and the pelvic limbs were placed under the abdomen. Stifle and tarsal joints were kept in a slightly flexed position. On palpation, the muscles of the limbs seemed to be atrophied and the articulations of elbow, carpus, hip and tarsus had a decreased range of motion. Manipulation of the joints seemed to be painful for the dog. Spinal reflexes and proprioceptive testing was normal, as was the remainder of the neurological examination. Both female dogs were examined and were found to be affected in different degrees of severity. While the clinical and neurological status of one female dog was comparable to those of the male dog, the second female was non-ambulatory and had difficulties with swallowing, regurgitation and salivation. This dog had an elevated body temperature of 39.6 °C and auscultation of the lungs revealed harsh sounds. Neurologically, the gag reflex was reduced but no other cranial nerve abnormalities were detected. Compared to her littermates, the muscles of this dog were more profoundly atrophied. The remainder of the neurological tests did not differ from those of her littermates.

In all affected dogs the creatine kinase activity (CK) was elevated (422, 935, 1149 IU/l; reference values: 51-191 IU/l). Hematology of the more severely affected female revealed an elevated leukocyte cell count whereas it was normal in the other dogs. Urinalysis was normal in all dogs. Serology for *Neospora canis* and *Toxoplasma gondii* was tested in 2 dogs and results came back negative.

Thoracic radiographs revealed generalized megaoesophagus and aspiration pneumonia in the severely affected female. In the other two dogs, radiographs did not show evidence of esophageal, lung or cardiac pathology. Radiographic examination of the hips, stifle, elbow and carpal articulation did not reveal any degenerative findings or conformational abnormalities.

We then performed electromyography to measure the nerve conduction velocity. In all 3 affected dogs, appendicular and axial skeletal muscles were examined. There was a dense pattern of fibrillation potentials and complex repetitive discharges in all tested myotomes. The nerve conduction velocity in the tibial nerves was within normal limits along the entire course of the nerve.

All 3 affected dogs were euthanized at 5 months of age after histological diagnosis (see below) and due to the severity of their clinical signs. The three clinically unaffected littermates were examined by means of a physical and neurological examination and routine hematology and biochemical profile. No clinical abnormalities were detected and results of laboratory testing including CK were normal. The dam and one unaffected female sibling were screened for subclinical evidence of the disease by EMG examination and muscular biopsy with both tests yielding negative results.

Similar clinical findings were reported by the owners and their veterinarians in another litter of Landseer dogs with two affected siblings (Video S1). These dogs were euthanized at 8 and 15 months of age due to severe progressive muscle wasting.

Pathological findings

We investigated the histopathology of skeletal muscle samples. All affected dogs showed pathological variation in muscle fiber size and most of the fibers were round to anisomorphic instead of having a (physiological) polygonal shape. Many small fibres were scattered throughout the biopsies, along with some hypercontracted fibres. Sporadically invading phagocytes grouped around degenerating fibers. Staining for acidic phosphatase showed an increase of activity, indicating degeneration to necrosis in a large number of fibres. Some biopsies correlated to an advanced stage of disease with distinct proliferation of the endomysial connective tissue and an increase of adipose tissue. These findings demonstrate different stages of muscular destruction with compensatory hypertrophy and replacement of lost fibers by connective tissue and fat cells. The histhopathological findings were typical for a muscular dystrophy. However, relatively normal immunohistochemistry findings with an anti-dystrophin antibody clearly showed that the muscular dystrophy in the Landseer dogs is different from the Duchenne/Becker type. The unaffected littermates did not show any histopathological alterations (Fig. 1).

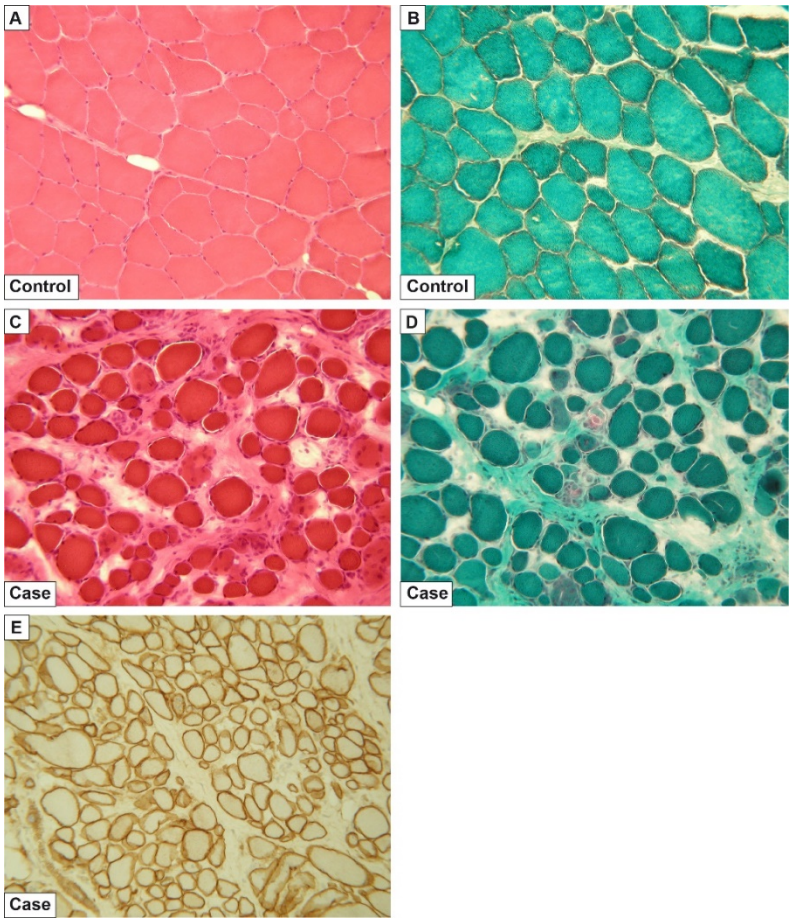

Fig. 1. Skeletal muscle pathology. (A) M. quadriceps of a 4 months old healthy Landseer dog with normal muscle texture without signs of significant pathological changes. Especially, muscle and fiber size configuration and variation, sarcolemma, nuclear structure and distribution as well as endomysial connective tissue are of physiological appearance. HE-staining, 100x. (B) Same muscle in the Gomori-Trichrome-staining according to Engel. Normal muscle tissue and fiber structure without considerable pathological changes are obvious, especially intrasarcoplasmal texture, mitochondria, muscular nuclei and endomysial connective tissue are unobtrusive. (C) M. quadriceps of an affected full-sibling of the control dog shown above. HE staining, 100x. The muscle texture is severely disturbed showing a pathological muscle fiber variation in size and configuration, signs of de- and regeneration, e.g. nuclear proliferation and distribution, as well as a remarkable increase and distortion of the endomysial connective tissue. (D) Same muscle in the Gomori-Trichrome-staining according to Engel points up the pathological muscle changes. Interestingly and in contrast to other forms of muscle dystrophy such as the Duchenne/Becker analogous disease in Golden Retrievers, the sarcoplasm structures and the mitochondria are not markedly altered. (E) Immunohistochemistry with an antibody against dystrophin performed on the same muscle as in (C) and (D). The diseased muscle shows no deficiency and regular expression of dystrophin. Thus this particular muscular dystrophy is markedly different from the Duchenne/Becker types of muscular dystrophy, where dystrophin expression is altered.

Genetic mapping of the causative mutation

The pedigrees of the two Landseer families were consistent with a monogenic autosomal recessive mode of inheritance. The two litters had a common great-great-grand mother in their maternal lineages. On their paternal ancestry we could not detect common ancestors within five generations.

At the time of the initial mapping of the causative locus we had samples from only four reliably diagnosed cases: Three affected littermates from the first family and one isolated case from the second family. We additionally had samples from the three non-affected siblings and both non-affected parents of the first family. We obtained 173,662 SNP chip genotypes for each of these 9 dogs. We then performed a parametric linkage analysis using a fully penetrant, monogenic autosomal recessive model of inheritance in the family with its 10 informative meioses. We obtained positive LOD scores for 5 genome segments containing 74 Mb in total. The highest LOD scores with 1.6 were obtained on chromosomes 24 and 31 (Table S1).

Based on the pedigree records and typical dog breeding practices we hypothesized that all affected dogs most likely were inbred to one single founder animal. Under this scenario the affected individuals were expected to be identical by descent (IBD) for the causative mutation and flanking chromosomal segments. We therefore analyzed the 4 available cases for extended regions of homozygosity with simultaneous allele sharing and found 18 genome regions with a total of 67 Mb that fulfilled our homozygosity search criteria (Table S2).

When we intersected the linked intervals from the family and the homozygous intervals from the four cases, only two intervals on chromosomes 10 and 31 remained. Taken together these two intervals had a size of 4.8 Mb and were considered the minimal critical interval for the subsequent analyses (Fig. 2).

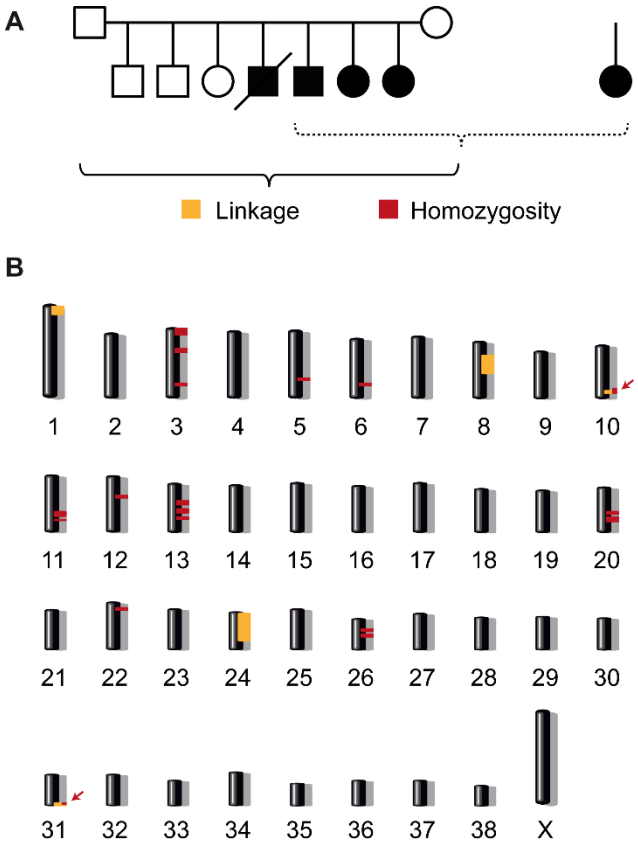

Fig. 2. Mapping strategy. (A) A family comprising 6 offspring with their parents and one isolated case were available for the initial mapping of the disease locus. We performed parametric linkage analysis for a recessive trait in the family and homozygosity analysis across the four cases. (B) The analyses yielded 5 linked genome segments (orange) and 18 homozygous genome segments (red). Only two regions on chromosomes 10 and 31 showed both linkage and homozygosity and were considered the critical intervals (arrows). Specifically, these regions corresponded to Chr10:61,871,450-66,047,210 and Chr31:38,752,158-39,364,930.

Mutation identification

In order to obtain a comprehensive overview of all variants in the critical intervals we re-sequenced the whole genome of one affected Landseer. We collected 330 million 2 x 100 bp paired-end reads corresponding to 14.2x coverage of the genome. We called SNPs and indel variants with respect to the reference genome of a presumably non-affected Boxer. Across the entire genome, we detected ~5.6 million variants of which ~2.8 million were homozygous variants (Table 1). Within the critical intervals there were 10,190 variants, of which 38 were predicted to be non-synonymous.

Table 1. Variants detected by whole genome re-sequencing of an affected Landseer.

Filtering step	Number of variants
Variants in the whole genome[a]	2,808,714
Variants in the critical intervals on CFA 10 and 31	10,190
Variants in the critical intervals that were absent from 170 other dog genomes	18
Non-synonymous variants in the whole genome[a]	11,798
Non-synonymous variants critical intervals on CFA 10 and 31	38
Non-synonymous variants in the critical intervals that were absent from 170 other dog genomes	1

[a] The sequences were compared to the reference genome (CanFam 3.1) from a Boxer. Only variants that were homozygous in the affected Landseer are reported.

We further compared the genotypes of the affected Landseer with 170 dog genomes of various breeds that had been sequenced in the course of other ongoing studies. We hypothesized that the mutant allele at the causative variant should be completely absent from all other dog breeds in our sample set. After this filtering step only a single non-synonymous private variant remained in the critical intervals, Chr31:39,303,964G>T or *COL6A1:c.289G>T*. We confirmed this variant by Sanger sequencing (Fig. 3) and genotyped it in 58 Landseer dogs, 404 Newfoundland dogs, and 473 dogs of 63 diverse other breeds. The genotypes showed perfect co-segregation with the muscular dystrophy in Landseer dogs and we did not find the variant allele outside of the Landseer breed (Table 2; Table S3). The *COL6A1:c.289G>T* variant represents a nonsense mutation in exon 3 of the *COL6A1* gene, encoding the alpha-1 chain of collagen type VI. The variant leads to an early premature stop codon and is predicted to truncate more than 90% from the *COL6A1* open reading frame (p.E97*).

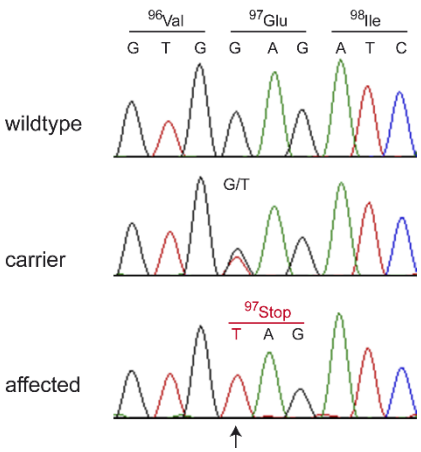

Fig. 3. Electropherograms of the *COL6A1:c.289G>T* variant. A fragment harboring exon 3 and flanking sequences of the *COL6A1* gene was PCR-amplified and sequenced with the Sanger method. The figure

shows representative traces from Landseer dogs with the 3 different genotypes. The position of the variant is indicated by an arrow.

Table 2. Association of the *COL6A1:c.289G>T* genotypes with muscular dystrophy.

Genotype	Landseer cases[a]	Landseer controls[b]	Newfoundland Dogs[c]	Control dogs from other breeds[d]
COL6A1:c.289G>T				
G/G	-	33	404	473
G/T	-	20	-	-
T/T	5	-	-	-

[a] Four cases had been available at the beginning of the project. A fifth case, full sister to the case from the second family, became available only after the genetic mapping of the locus had already been completed.

[b] The Landseer controls consisted of 5 close relatives of the affected dogs and 48 Landseer dogs without known relationships to the two families with muscular dystrophy. Of these 48 dogs, 33 were homozygous G/G and 15 were heterozygous G/T. In this small and not necessarily representative cohort, the carrier frequency was thus 31%.

[c] We genotyped a large number of Newfoundland Dogs as the Landseer breed was originally derived from Newfoundland Dogs. The 404 genotyped samples were collected in the course of other ongoing research projects.

[d] These 473 dogs were specifically genotyped for the *COL6A1:c.289G>T* variant. The genome sequences of 170 independent control dogs were also homozygous G/G at this variant. Therefore, the total number of control dogs adds up to 643.

Discussion

In this study we describe a novel canine muscular dystrophy in Landseer dogs. The clinical and histhopathological findings were similar to the severe Duchenne muscular dystrophy in Golden Retrievers with a genetic defect in the X-linked *DMD* gene encoding dystrophin (Cooper *et al.* 1988; Valentine *et al.* 1988; Valentine *et al.* 1992). However, as male and female Landseer dogs were equally affected in the two studied families, an X-linked genetic defect seemed very unlikely. We therefore applied a positional cloning approach to elucidate the causative genetic defect. With only one complete family and one additional case we were able to map the disease locus to just 0.2% of the dog genome. The increasing availability of "personal genomes" in canine genetics allowed us to quickly identify the only non-synonymous genetic variant in the critical genome regions that was exclusively present in an affected Landseer and absent from 170 control dogs. We observed perfect concordance of the genotypes at this variant with the muscular dystrophy phenotype in more than one thousand dogs. We have to caution here that our genetic data do not prove the causality of the identified variant. Based on these data we cannot formally rule out the possibility that the true causative variant is a non-coding regulatory variant or that the true causative variant was entirely missed by the chosen short read re-sequencing methodology in combination with an incomplete reference genome.

On the other hand, the genetic evidence provides statistical support for the causality of the *COL6A1:c.289G>T* for the observed muscular dystrophy. In addition to the genetic data, the functional knowledge on collagen type VI from other species is fully consistent with this hypothesis. *COL6A1* is actually a very good functional candidate gene for a muscular dystrophy as genetic variants in this gene cause different human myopathies with variable clinical phenotypes ranging from the milder Bethlem myopathy to the severe Ullrich congenital muscular dystrophy.

The reported variant in Landseer dogs introduces an early stop codon into the open reading frame. Although we have not experimentally analyzed the expression of the mutant *COL6A1* gene, it seems likely that the mutant transcripts will be subject to nonsense-mediated decay and that the *COL6A1:c.289G>T* variant leads to a true null allele of the *COL6A1* gene. Unfortunately, we had very limited material from the muscle biopsies and could not analyze the RNA experimentally to confirm this hypothesis.

Based on the combination of the available genetic data, knowledge on *COL6A1* function in other species, and the fact that the identified nonsense variant is likely to represent a full null allele, we think that the causality of the *COL6A1:c.289G>T* variant for the observed muscular dystrophy has been established beyond reasonable doubt.

In humans with collagen VI alterations, the clinical severity is variable. Ullrich's congenital muscular dystrophy is the most severe phenotype with severe weakness and axial und proximal contractures. Affected patients lose the ability to walk within the first or second decade of life. Bethlem congenital muscular dystrophy has milder clinical signs including moderate to mild muscle weakness and distal contractures only. Between Ullrich's and Bethlem muscular dystrophies intermediate phenotypes exist with lesser degrees of weakness and longer periods of unassisted ambulation (Bönnemann 2011). The Landseer dogs showed a severe clinical phenotype, which closely resembles human patients with Ullrich congenital muscular dystrophy. The clinical phenotype of these Landseers was more severe than that seen in a myopathic Labrador Retriever suspected to also have a defect in collagen VI (Marioni-Henry *et al.* 2014). So far, according to our knowledge, no human patients carrying two truncated *COL6A1* alleles have been reported. However, it seems plausible that homozygosity for a *COL6A1* nonsense variant in the Landseer dogs leads to a very severe phenotype with a recessive mode of inheritance. A comparable situation resulting in severe clinical phenotypes has been reported in human patients carrying truncating variants on both alleles of the *COL6A2* gene (Camacho Vanegas *et al. 2001*; Allamand *et al.* 2011).

It is interesting to note that homozygous *Col6a1*$^{-/-}$ knock out mice show a very mild phenotype that has been compared to human Bethlem myopathy although these mice do not express any residual collagen alpha1(VI) chains (Bonaldo *et al.* 1998; Irwin *et al.* 2003). This is in contrast to humans and dogs where the complete *COL6A1* deficiency leads to a much more severe phenotype. The comparatively mild murine phenotype recapitulates the findings from Duchenne muscular dystrophy: Completely *DMD* deficient humans and dogs suffer from a very severe clinical phenotype, whereas *Dmd* deficient mice only have a very mild phenotype (McGreevy *et al.* 2015). Thus, the Landseer dogs with the *COL6A1:c.289G>T* variant are potentially an

interesting animal model for human Ullrich congenital muscular dystrophy as they resemble the severe human clinical phenotype much closer than *Col6a1*$^{-/-}$ knock out mice.

In conclusion, we provide the first comprehensive description of a clinically severe canine muscular dystrophy, most likely caused by a nonsense variant in the *COL6A1* gene. The affected dogs represent a model for human Ullrich congenital muscular dystrophy. Our findings enable genetic testing in dogs, so that the non-intentional breeding of affected Landseers can be avoided in the future.

Materials and Methods

Ethics statement

All animal experiments were performed according to the local regulations. The dogs in this study were examined with the consent of their owners. The collection of blood samples was approved by the "Cantonal Committee For Animal Experiments" (Canton of Bern; permit 23/10). The clinical examinations of affected Landseer dogs and their littermates including the collection of muscle biopsies was performed in the course of standard veterinary diagnostics, with the consent and on behalf of the dogs' owners. The clinical examinations therefore did not constitute an animal experiment in the legal sense and did not require ethical approval.

Clinical examinations

Three affected Landseer siblings were brought to the neurology service at the Vetsuisse Faculty of Zurich for examination and diagnostic testing. A history of all dogs was collected from the breeder. All dogs of the litter underwent both a general physical examination and a detailed neurologic examination, as well as having complete blood counts, serum biochemistry panels and urinalysis. Electromyography and motor nerve conduction velocity (MNCV) were performed on an electrodiagnostic device (Neuroscreen, Toennies AG, Germany) with concentric and monoaxial needle electrodes while the animals were maintained under general gaseous anesthesia. For MNCV the tibial nerve was examined along its entire course using stimulation sites at the trochanteric fossa, the popliteal fossa with recording at the interosseous muscle. Muscle biopsies from *M. triceps*, *M. biceps femoris*, and *M. tibialis cranialis* were collected after electrodiagnostic testing using an open standard surgical approach. Biopsies were placed in a container without any additives and were submitted for histopathologic assessment by overnight express to the Institute of Neuropathology of the Heinrich-Heine-University, Düsseldorf (Germany). Clinical data on two affected Landseer dogs from a second family was kindly contributed by their owners and the attending veterinarians.

Histopathological examinations

Muscle biopsies were frozen at -135 °C in isopentan and stored at -80 °C for further analysis. The frozen biopsies were cut using a cryosat (CM 1900, Leica Biosystems, Germany) into sections of 7 µm thickness and stained using haematoxylin and eosin (HE), modified Gomori trichrome (Engel) and acidic phosphatase (AP). All sections were examined under a transmitted light microscope (Orthoplan, Leica Microsystems, Germany) in 100x, 400x and 1000x magnification. Additional cryosections were used for immunohistochemical stainings using an antibody against the rod domain of dystrophin (1:100, NCL-DYS1, Novocastra Laboratory).

Animal selection for the genetic analysis

This study describes findings from 2 Landseer families. The first family (Fig. 2 A) consisted of the parents and 6 offspring that were available for EDTA blood sample collection. A seventh offspring was affected by muscular dystrophy, but was euthanized before samples could be taken. From the second family we had access to the blood samples of 2 affected female offspring. The sample of the second affected dog from this family became available only after we had already completed the initial genetic mapping of the disease locus. Thus, in total we had samples from 5 affected Landseer dogs, 3 non-affected full-siblings and 2 non-affected parents.

We additionally had samples from 48 other Landseer dogs without known relationships to the two families that had been donated to the biobank of the Institute of Genetics. Finally, we also used 404 samples from Newfoundland dogs and 473 samples from 63 various diverse dog breeds, which are listed in detail in Table S3.

DNA samples and SNP genotyping

We isolated genomic DNA from EDTA blood samples with the Nucleon Bacc2 kit (GE Healthcare). Genotyping was done on illumina canine_HD chips containing 173,662 SNPs by GeneSeek/Neogen. Genotypes were stored in a BC/Gene database version 3.5 (BC/Platforms).

Linkage and homozygosity mapping

For the linkage analysis we had illumina canine_HD SNP chip genotypes from a single family consisting of 8 animals in total. This family comprised both non-affected parents and six offspring, of which three were affected. We used the Merlin software (Abecasis *et al.* 2002) and a fully penetrant, recessive model of inheritance to analyze the data for parametric linkage.

We used PLINK v1.07 (Purcell *et al.* 2007) to search for extended intervals of homozygosity with shared alleles as described (Drögemüller *et al.* 2009). The final critical intervals were defined by visual inspection of all SNP chip genotypes of the four genotyped cases on chromosomes 10 and 31 in an Excel-file.

Gene analysis

We used the dog CanFam 3.1 assembly for all analyses. All numbering within the canine *COL6A1* gene corresponds to the accessions XM_003434001.3 (mRNA) and XP_003434049.2 (protein).

Whole genome sequencing of an affected Landseer dog

We prepared a PCR-free fragment library with 300 bp insert size and collected 330,331,465 illumina HiSeq2500 paired-end reads (2 x 100 bp) or roughly 14.2x coverage. We mapped the reads to the dog reference genome using the Burrows-Wheeler Aligner (BWA) version 0.5.9-r16 (Li and Durbin 2009) with default settings and obtained 380,485,021 uniquely mapping reads. After sorting the mapped reads by the coordinates of the sequence with Picard tools, we labeled the PCR duplicates also with Picard tools (http://sourceforge.net/projects/picard/). We used the Genome Analysis Tool Kit (GATK version v2.3-6, (McKenna *et al.* 2010)) to perform local realignment and to produce a cleaned BAM file. Variant calls were then made with the unified genotyper module of GATK. Variant data for each sample were obtained in variant call format (version 4.0) as raw calls for all samples and sites flagged using the variant filtration module of GATK. Variant calls that failed to pass the following filters were labeled accordingly in the call set: (i) Hard to Validate MQ0 ≥ 4 & ((MQ0 / (1.0 * DP)) > 0.1); (ii) strand bias (low Quality scores) QUAL < 30.0 || (Quality by depth) QD < 5.0 || (homopolymer runs) HRun > 5 || (strand bias) SB > 0.00; (iii) SNP cluster window size 10. The snpEFF software (Cingolani *et al.* 2012) together with the CanFam 3.1 annotation was used to predict the functional effects of detected variants. We considered the following snpEFF categories of variants as non-synonymous: NON_SYNONYMOUS_CODING, CODON_DELETION, CODON_INSERTION, CODON_CHANGE_PLUS_CODON_DELETION, CODON_CHANGE_PLUS_CODON_INSERTION, FRAME_SHIFT, EXON_DELETED, START_GAINED, START_LOST, STOP_GAINED, STOP_LOST, SPLICE_SITE_ACCEPTOR, SPLICE_SITE_DONOR. The critical intervals contained 4,788,534 total and 50,649 coding nucleotides, respectively. In our re-sequencing data, we had ≥4x coverage on 4,671,977 bp of the critical interval (97.6%) and on 49,327 (97.4%) of the coding bases. The genome sequencing data were deposited in the European Nucleotide Archive (ENA) under accession PRJEB9437.

Sanger sequencing

We used Sanger sequencing to confirm the illumina sequencing results and to perform targeted genotyping for selected variants. For these experiments we amplified PCR products using AmpliTaqGold360Mastermix (Applied Biosystems). PCR products were directly sequenced on an ABI 3730 capillary sequencer (Applied Biosystems) after treatment with exonuclease I and shrimp alkaline phosphatase. We analyzed the Sanger sequence data with Sequencher 5.1 (GeneCodes).

Acknowledgments

The authors are grateful to the referring veterinarians Jochen Spennes and the Tierärztliche Praxis am Loh GmbH and to all dog owners and breeders who donated samples and shared pedigree data of their dogs. We acknowledge collaborators around the world for sharing dog genome sequence data from control dogs. We thank Nathalie Besuchet Schmutz, Michèle Ackermann, and Muriel Fragnière for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the whole genome sequencing experiments, and the Vital-IT high-performance computing center of the Swiss Institute of Bioinformatics for performing computationally intensive tasks (http://www.vital-it.ch/). This study was supported by a grant from the Albert-Heim-Foundation.

References

Abecasis, G. R., S. S. Cherny, W. O. Cookson, and L. R. Cardon, 2002 Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat. Genet. 30: 97-101.

Allamand, V., L. Briñas, P. Richard, T. Stojkovic, S. Quijano-Roy, *et al.*, 2011 ColVI myopathies: where do we stand, where do we go? Skelet. Muscle. 1: 30.

Bonaldo, P., P. Braghetta, M. Zanetti, S. Piccolo, D. Volpin, *et al.*, 1998 Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for Bethlem myopathy. Hum. Mol. Genet. 7: 2135-2140.

Bönnemann, C. G., 2011 The collagen VI-related myopathies: muscle meets its matrix. Nat. Rev. Neurol. 7: 379-390.

Briñas, L., P. Richard, S. Quijano-Roy, C. Gartioux, C. Ledeuil C, *et al.*, 2010 Early onset collagen VI myopathies: Genetic and clinical correlations. Ann. Neurol. 68: 511-520.

Camacho Vanegas, O., E. Bertini, R. Z. Zhang, S. Petrini, C. Minosse, *et al.*, 2001 Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. Proc. Natl. Acad. Sci. U. S. A. 98: 7516-7521.

Cingolani, P., A. Platts, L. L.Wang, M. Coon, T. Nguyen, *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6: 80-92.

Cooper, B. J., N. J. Winand, H. Stedman, B. A. Valentine, E. P. Hoffman, *et al.*, 1988 The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. Nature 334: 154-156.

De Palma, S., D. Capitanio, M. Vasso, P. Braghetta, C. Scotton, *et al.*, 2014 Muscle proteomics reveals novel insights into the pathophysiological mechanisms of collagen VI myopathies. J. Proteome Res. 13: 5022-5030.

Drögemüller, C., D. Becker, A. Brunner, B. Haase, P. Kircher, *et al.*, 2009 A missense mutation in the *SERPINH1* gene in Dachshunds with osteogenesis imperfecta. PLoS Genet. 5: e1000579.

Flanigan, K. M., 2012 The muscular dystrophies. Semin. Neurol. 32: 255-263.

Furthmayr, H., H. Wiedemann, R. Timpl, E. Odermatt, and J. Engel, 1983 Electron-microscopical approach to a structural model of intima collagen. Biochem. J. 211: 303-311.

Hart, K. A., S. Hodgson, A. Walker, C. G. Cole, L. Johnson *et al.*, 1987 DNA deletions in mild and severe Becker muscular dystrophy. Hum. Genet. 75: 281-285.

Irwin, W. A., N. Bergamin, P. Sabatelli, C. Reggiani, A. Megighian, *et al.*, 2003 Mitochondrial dysfunction and apoptosis in myopathic mice with collagen VI deficiency. Nat. Genet. 35: 367-371.

Jöbsis, G. J., H. Keizers, J. P. Vreijling, M. de Visser, M. C. Speer, *et al.*, 1996 Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. Nat. Genet. 14: 113-115.

Koenig, M., E. P. Hoffman, C. J. Bertelson, A. P. Monaco, C. Feener *et al.*, 1987 Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 50: 509-517.

Koenig, M., A. H. Beggs, M. Moyer, S. Scherpf, K. Heindrich, *et al.*, 1989 The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. Am. J. Hum. Genet. 45: 498-506.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760.

Marioni-Henry, K., P. Haworth, H. Scott, P. Witte, L. T. Guo, *et al.*, 2014 Sarcolemmal specific collagen VI deficient myopathy in a Labrador Retriever. J. Vet. Intern. Med. 28: 243-249.

McGreevy, J. W., C. H. Hakim, M. A. McIntosh, and D. Duan, 2015 Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. Dis. Model Mech. 8: 195-213.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, *et al.*, 2010 The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297-1303.

Nicholas, F. W., and M. Hobbs, 2014 Mutation discovery for Mendelian traits in non-laboratory animals: a review of achievements up to 2012. Anim. Genet. 45: 157-170.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81: 559-575.

Valentine, B. A., B. J. Cooper, A. de Lahunta, R. O'Quinn, and J. T. Blue, 1988 Canine X-linked muscular dystrophy. An animal model of Duchenne muscular dystrophy: clinical studies. J. Neurol. Sci. 88: 69-81.

Valentine, B. A., N. J. Winand, D. Pradhan, N. S. Moise, A. de Lahunta, *et al.*, 1992 Canine X-linked muscular dystrophy as an animal model of Duchenne muscular dystrophy: a review. Am. J. Med. Genet. 42: 352-356.

Supplementary Material

Table S1. Genome regions that showed positive LOD scores for linkage to muscular dystrophy in a family with 10 informative meioses.

Table S2. Homozygous genome regions with shared alleles among the 4 analyzed cases.

Table S3. *COL6A1:c.289G>T* genotypes of 473 dogs from diverse dog breeds.

Video S1. Clinical phenotype of a 4 month old affected Landseer.